Liquid Biopsy to Identify Actionable Genomic Alterations

Sai-Hong Ignatius Ou, MD, PhD, Misako Nagasaka, MD, and Viola W. Zhu, MD, PhD

OVERVIEW

Liquid biopsy has been used extensively in solid malignancies to detect actionable driver mutations, to monitor treatment response, to detect recurrence, to identify resistance mechanisms, and to prognosticate outcome. Although many liquid biopsy sequencing platforms are being used, only five test kits have received government approval. We review representative literature on these government-approved liquid biopsy kits, which are primarily used to detect *EGFR* mutation in lung cancer and *RAS* (*KRAS, NRAS, BRAF*) mutations in colorectal carcinoma. Another emerging use of single-gene liquid biopsy is to detect *PIK3CA* mutations and to understand resistance to hormonal blockade in breast and prostate cancers. The two most commonly used next-generation sequencing (NGS) liquid biopsy tests (Guardant 360, Guardant Health; FoundationACT, Foundation Medicine Inc.) are discussed. The ability and the applicability of NGS platform to detect tumor mutation burden are also addressed. Finally, the use of circulating tumor DNA (ctDNA) to detect minimal residual disease may be the most important use of ctDNA in the setting of tumor heterogeneity. The ability to identify "shedders" and "nonshedders" of ctDNA may provide important insight into the clinicopathologic characteristics of the tumor and portend important prognostic significance regarding survival.

The use of liquid biopsy (primarily from blood) in solid malignancy provides a convenient and safe way to detect the presence of actionable driver mutations, to assess the resistance mechanisms to actionable driver mutations, to monitor treatment response, to detect early recurrence, to serve as an adjuvant to radiologic imaging as post-treatment surveillance, and to prognosticate the outcome of cancer treatment. Cell-free DNA, including ctDNA, circulating tumor cells, and exosomes containing tumor microRNAs can all be detected by liquid biopsy. The biologic nature of ctDNA, the various sequencing platforms used in liquid biopsy, and the various utilities of liquid biopsy have recently been expertly and comprehensively reviewed by Wan and colleagues.[1] The many sequencing platforms used in liquid biopsy can be broadly summarized as nondigital, digital, and NGS. The performances of these individual platform have been reviewed extensively.[2-10] However, only five liquid biopsy test kits are approved by government agencies.

DETECTION OF SPECIFIC ACTIONABLE GENOMIC ALTERATIONS BY LIQUID BIOPSY

Detection of Activating *EGFR* Mutations in Lung Cancer

Cobas EGFR mutations test version 2 (*EGFR* del 19, L858R, T790M). The Cobas test (Roche Molecular Diagnostics, Pleasonton, CA) is the only U.S. Food and Drug Administration (FDA)–approved liquid biopsy to detect the two most common activating epidermal growth factor receptor (EGFR) mutations (*EGFR* del19 and *EGFR* L858R) for the selection of EGFR tyrosine kinase inhibitor. It was subsequently approved for the detection of the most common acquired resistance mutation, *EGFR* T790M, after progression with first- or second-generation EGFR tyrosine kinase inhibitors for selection of osimertinib to treat patients with *EGFR* T790M–positive non–small cell lung cancer (NSCLC; Table 1).[11-14] ENSURE, a randomized phase III trial comparing erlotinib to platinum/gemcitabine chemotherapy as first-line treatment of *EGFR*-positive NSCLC provided the basis for the approval of detection of *EGFR* del19 and *EGFR* L858R.[15] Additional large-scale real-life prospective trials (ASSESS, Europe and Japan; and IGNITE, Russia and China/South Korea/Taiwan) studying the feasibility and testing the concordance of using Cobas liquid biopsy versus tumor have been completed.[16,17] In the IGNITE study, the concordance between 2,561 matched tissue/cytology and plasma samples was 80.5%, sensitivity was 46.9%, and specificity was 95.6%.[17] In the ASSESS study, the concordance of mutation status in 1,162 matched samples was 89%, sensitivity was 46%, specificity was 97%, positive predictive value was 78%, and negative predictive value was 90%. Two combined single-arm phase II studies of osimertinib provided the basis for the approval for the detection of *EGFR* T790M.[18] A European study (APPLE) investigating

From Chao Family Comprehensive Cancer Center, University of California Irvine School of Medicine, Orange, CA; Karmanos Cancer Center, Wayne State University, Detroit, MI; Hematology/Oncology Section, Veterans Affairs Long Beach Healthcare System, Long Beach, CA.

Disclosures of potential conflicts of interest provided by the authors are available with the online article at asco.org/edbook.

Corresponding author: Sai-Hong Ignatius Ou, MD, PhD, Chao Family Comprehensive Cancer Center, University of California Irvine School of Medicine, 101 The City Dr. S, Building 56, RT81, Room 241, Orange, CA 92868; email: ignatius.ou@uci.edu.

© 2018 American Society of Clinical Oncology

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.1

the use of liquid biopsy to detect *EGFR* T790M mutation is ongoing.[19]

Therascreen EGFR plasma RGQ PCR kit. The Therascreen EGFR Plasma RGQ PCR test (Qiagen Inc., Venlo, the Netherlands) is approved in the European Union (E.U.; Conformité Européene-In Vitro Diagnostics) for detection of the two most common *EGFR* mutations (*EGFR* del19 and *EGFR* L858R) but not T790M (Table 1).[20-23] The IFUM study provided the basis for adoption of this test in European Union.[24,25]

AmoyDx Super-ARMS EGFR mutation test kit. The newest government-approved liquid biopsy, the AmoyDx Super-ARMS EGFR mutation test kit (AmoyDx, Xiamen, China), was approved on January 22, 2018, by the Chinese Food and Drug Administration to detect *EGFR* del19, *EGFR* L858R, and *EGFR* T790M mutations (Table 1).[26-28] AURA17 is a single-arm study of osimertinib in China that provided the supporting evidence for the approval of this test kit.[28]

Role in diagnosis. Although all three tests can detect the three "uncommon" *EGFR* mutations (L861Q, G719X, and S768I)—for which afatinib was approved for use in the United States on January 16, 2018—none of them is officially approved to detect them.[29] Table 1 shows that across all three test kits, the sensitivity for detecting T790M seems to be lower than that for detecting the two common activation mutations. The higher the prevalence of *EGFR* mutations (e.g., approximately 40%–50% in Asian populations), the lower the negative predictive value of the liquid biopsy (Table 1); this is based on the Cobas EGFR version 2 calculations. Thus, a negative *EGFR* plasma test result in a high-endemic region for *EGFR* mutations should be followed by a tumor tissue biopsy.

In addition, the National Institute of Health and Care Excellence in the United Kingdom recently published their technology summary of evidence of seven different liquid biopsy platforms.[30] The four additional tests reviewed, in addition to the three government-approved tests, were the AmoyDx EGFR 29 mutation detection kit, AmoyDx Super-ARMS EGFR T790M mutation detection kit, Droplet Digital PCR (ddPCR) Dx system (Bio-Rad, Hercules, CA), and PANA Mutyper R EGFR kit (Panagene, Daejeon, Korea). The major summary finding is that plasma *EGFR* detection has specificity similar to but sensitivity lower than those of tumor biopsy. The resource impact to the health care system is similar to that for standard of care, but plasma testing could be cost-effective if this led to fewer tissue biopsies[30]

Going forward, increasing evidence suggests that the ability to identify detectable circulating *EGFR* mutations from liquid biopsy ("shedders" compared with "nonshedders") has prognostic implications. In the FLAURA and AURA3 phase III trials (Table 2), although both shedders and nonshedders benefited from osimertinib, progression-free survival (PFS) in the osimertinib arm was numerically better for nonshedders than for shedders; this finding indicates that shedders in general have a higher tumor burden (Table 2).[14,31] Additionally, among patients treated at National Taiwan University Hospital who progressed while receiving osimertinib, a retrospective genomic analysis that used plasma BEAMing technology (Sysmex Inostics, Hamburg, Germany) to detect *EGFR* mutations showed that, again, shedders had significantly shorter median PFS while receiving osimertinib, shorter PFS after progression on osimertinib, and shorter overall survival (OS) after osimertinib treatment (Table 2).[32] Furthermore, 42.5% (17/40) of the shedders compared with only 7.7% (1/13) of nonshedders had brain metastases in this study, again indicating that shedders had more extensive/poor-prognosis disease than nonshedders. Finally, the "disappearance" of *EGFR* T790M at the time of progression during osimertinib treatment was associated with significantly shorter PFS after osimertinib therapy (Table 2), indicating more recalcitrant non–*EGFR*-mediated resistance mechanisms.[32] Shedders of *EGFR* T790M usually have higher tumor burden with late-stage disease. Similarly, shorter PFS has been reported in shedders with *EGFR* L858R mutation than in nonshedders from a retrospective analysis of the EURTAC trial.[33,34]

PRACTICAL APPLICATIONS

- **Despite the widespread use of liquid biopsy in solid malignancy, only five distinct test kits have government approval: Cobas EGFR Mutations Test v2 (Roche Molecular Diagnostics; FDA approval for detection of *EGFR del19, EGFR L858R,* and *EGFR T790M* in lung cancer), Therascreen EGFR RGQ Plasma PCR kit (Qiagen Inc; E.U. approval for detection of *EGFR* del19 and *EGFR* L858R in lung cancer); AmoyDx Super-ARMS EGFR mutation test (AmoyDx; Chinese Food and Drug Administration approval for detection of *EGFR* del19, *EGFR* 858R, and *EGFR* T790M in lung cancer); OncoBEAM RAS CRC Kit (Sysmex Inostics; E.U. approval for detection of *KRAS* and *NRAS* mutations in colorectal cancer); and Idylla ctKRAS Mutation Test and Idylla ctNRAS-BRAF Mutation Test (Biocartis, Inc.; E.U. approval for detection of *KRAS, NRAS,* and *BRAF* mutations in colorectal cancer).**
- **Currently, liquid biopsy is most commonly used to detect actionable *EGFR, RAS* (*KRAS, NRAS*), and *BRAF* V600E mutations, with potentially in the future for detection of *PIK3CA* mutations.**
- **Liquid biopsy based on NGS can detect multiple genomic alterations (mutations, indel, amplification, rearrangement) in multiple actionable genes and also determine tumor mutation burden used to potentially select for immune checkpoint inhibitors. Eventually, liquid biopsy using NGS will supersede use of single/oligo gene liquid biopsy detection kits.**
- **Liquid biopsy to detect nonspecific ctDNA in tumors without clear driver mutations may be most useful in detecting minimal residual disease, which can account for tumor heterogeneity in the adjuvant setting.**
- **Distinguishing "shedders" and "nonshedders" of ctDNA may provide insight into tumor biology and prognostic significance.**

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.2

TABLE 1. Approved Liquid Biopsy Test Kits for Detection of Activating and Resistance *EGFR* Mutations in Non–Small Cell Lung Cancer

	Therascreen EGFR Plasma RGQ PCR Kit	Cobas EGFR Mutation Test v2	AmoyDX Super-ARMS EGFR Mutation Detection Kit
PCR Kit			
Approving Agency/Region	European Union 98/79, Conformité Européene-In Vitro Diagnostics directive	U.S. FDA	Chinese FDA
Date of Approval	January 12, 2015	June 1, 2016 (*EGFR* del 19, *EGFR* L858R) September 28, 2016 (*EGFR* T790M)	January 22, 2018 (*EGFR* del 19, *EGFR* L858R, *EGFR* T790M)
Manufacturer	Qiagen	Roche	AmoyDx
Sequencing Platform	Scorpion Amplification Refractory Mutation System	Scorpion Amplification Refractory Mutation System	Scorpion Amplification Refractory Mutation system
Detectable Technology*	Analog (real-time PCR)	Analog (real-time PCR)	Analog (real-time PCR)
MAF Quantification	Semiquantitative	Semiquantitative	Semiquantitative
No. of *EGFR* Mutations Detected	29	42	41
Major *EGFR* Mutations Detected	*EGFR* del 19 (19 different mutations) *EGFR* L858R *EGFR* T790M *EGFR* G719X (3 different mutations) *EGFR* S761I *EGFR* L861Q *EGFR* exon 20 insertions (3 insertions)	*EGFR* del 19 (29 different mutations) *EGFR* L858R (2 different mutations) *EGFR* T790M *EGFR* G719X (3 different mutations) *EGFR* S761I *EGFR* L861Q *EGFR* exon 20 insertions (5 insertions)	*EGFR* del 19 (29 different mutations) *EGFR* L858R *EGFR* T790M *EGFR* G719X (3 different mutations) *EGFR* S761I *EGFR* L861Q *EGFR* exon 20 insertions (5 insertions)
Indication	*EGFR* del 19, *EGFR* L858R *EGFR* Del 19 and *EGFR* L858R	*EGFR* del 19 *EGFR* L858R *EGFR* T790M	*EGFR* del 19 *EGFR* L858R *EGFR* T790M
Study(ies) Supporting Approval	IFUM	ENSURE (YO25121)	Single-center, single-arm study (First Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang, China)
Comparator Tissue Test	Therascreen *EGFR* test (tissue use)	Cobas *EGFR* mutation test v2 (tissue test) NGS on an Illumina MiSeq	AmoyDx *EGFR* 29 mutation detection kit (tissue)
No. of Patients	859 with tumors successfully screened 652 with successfully paired tumor/plasma analyzed	601 with tumors successfully screened 431 with successfully paired tumor/plasma analyzed and validated 217 enrolled	109 screened 61 with tissue positive for all *EGFR* mutations 50 with plasma positive for all *EGFR* mutations

Continued

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TABLE 1. Approved Liquid Biopsy Test Kits for Detection of Activating and Resistance Mutations in Non–Small Cell Lung Cancer (Cont'd)

	Therascreen EGFR Plasma RGQ PCR Kit	Cobas EGFR Mutation Test v2	AmoyDX Super-ARMS EGFR Mutation Detection Kit
Primary Data	Screening for all patients: Of 105 patients positive for *EGFR* mutations from tumor, 69 (65.7%) were detected by plasma Of 547 patients negative for *EGFR* mutations from tumor, 546 (99.8%) were negative for *EGFR* mutation by plasma	Screening for all patients Of 210 patients positive for *EGFR* mutations from tumor, 161 (76.7%) were detected by plasma Of 221 patients negative for *EGFR* mutations from tumor, 217 (98.2%) were negative for *EGFR* mutation by plasma Screening for *EGFR* del 19 patients Of 120 patients positive for *EGFR* del 19 mutation from tumor, 97 (80.8%) were detected by plasma Of 311 patients negative for *EGFR* del 19 mutation from tumor, 307 (98.7%) were negative for *EGFR* mutation by plasma Screening for *EGFR* L858R mutation Sensitivity: Of 90 patients positive for *EGFR* L858R from tumor, 61 (67.8%) were detected by plasma Specificity: Of 341 patients negative for *EGFR* L858R from tumor, 338 (99.1%) were negative for *EGFR* L858R by plasma	Of 61 patients positive for *EGFR* mutations from tumor, 50 were also detected by plasma Of 48 patients negative for *EGFR* mutations from tumor, 48 were also negative for *EGFR* mutations by plasma
Performance Characteristics**	Sensitivity: 65.7% (95% CI, 55.8%–74.7%) Specificity: 99.8% (95% CI, 99%–100%) Concordance: 94.3% (95% CI, 92.3%–96.0%) PPV: 98.6% (95% CI, 92.3%–100%) NPV: 93.8% (95% CI, 91.5%–95.6%)	Prevalence: 15% PPV: 88.6% (95% CI, 79.7%–96.6%) NPV: 96.0% (95% CI, 94.3%– 97.6%) Prevalence: 40% PPV: 96.8% (95% CI, 93.0%–99.4%) NPV: 86.4% (95% CI, 83.3%–89.4%) Prevalence: 50% PPV: 97.8% (95% CI, 95.0%–100%) NPV: 80.8% (95% CI, 77.4%–84.8%)	Sensitivity: 82% (95% CI, 72.3%–91.6%) Specificity: 100% PPV: 81.4% (95% CI, 71.4%–91.3%) NPV: 100% Concordance: 89.9% (95% CI, 84.3%–95.6%)
Clinical Outcome Data	ORR (*EGFR* del 19 + L858R) Tumor+/plasma+: 72.5% (95% CI, 61.0%–81.6%); tumor+/plasma–: 63.6% (95% CI, 46.6%–77.8%) ORR (*EGFR* del 19) Tumor+/plasma+: 82.2% (95% CI, 68.7%–90.7%); tumor+/plasma–: 65.0% (95% CI, 43.5%–73.7%) ORR (*EGFR* L858R) Median PFS (all enrolled patients) All enrolled patients: 9.7 mo (95% CI, 8.5–11.0 mo) Tumor+/plasma+: 10.2 mo (95% CI, 8.5–12.5 mo) *EGFR* del 19 All: 9.6 mo (95% CI, 8.0–11.0 mo) Tumor+/plasma+: 10.3 mo (95% CI, 8.5–12.4 mo)	PFS benefit of erlotinib over chemotherapy (ITT population, n = 217) HR, 0.33 (95% CI, 0.23–0.47); log-rank test p < .0001 PFS benefit of erlotinib over chemotherapy (tumor+/plasma+; n = 137) HR, 0.29 (95% CI, 0.19–0.45); log-rank test p < .0001 PFS benefit of erlotinib over chemo (tumor+/plasma–; n = 42) HR, 0.37 (95% CI, 0.50–0.90)	ORR All patients: 64.3% (95% CI, 48%–78%; n = 42) Tumor+/plasma+: 65.7% (95% CI, 48%–81%; n = 35) Tumor+/plasma–: 57.15% (95% CI, 18%–90%)
T790M			

Continued

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TABLE 1. Approved Liquid Biopsy Test Kits for Detection of Activating and Resistance Mutations in Non–Small Cell Lung Cancer (Cont'd)

	Therascreen EGFR Plasma RGQ PCR Kit	Cobas EGFR Mutation Test v2	AmoyDX Super-ARMS EGFR Mutation Detection Kit
Validation Study	NA	Combined AURA extension study and AURA2 (*EGFR* T790M)	AURA17
Primary Data	NA	551 of 710 (78%) patients enrolled in both trials had paired tissue and plasma samples 416 of 551 (75.5%) patients were plasma positive for T790M, 127 (23.0%) were plasma negative for T790M, and 8 patients had invalid plasma tests	240 patients
Comparator Tissue Test		Cobas *EGFR* mutation test v2 (tissue test)	Cobas *EGFR* mutation test v2 (tissue test)
Performance Parameters	NA	*EGFR* T790M Sensitivity: 61% (95% CI, 57%–66%) Specificity: 79% (95% CI, 70%–85%) Concordance: 65% (95%: 61%–69%) *EGFR* del 19 Sensitivity: 85% (95% CI, 81%–89%) Specificity: 98% (95% CI, 95%–100%) Concordance: 90% (95% CI, 87%–92%) *EGFR* L858R Sensitivity: 76% (95% CI, 69%–82%) Specificity: 98% (95% CI, 96%–99%) Concordance: 91% (95% CI, 88%–93%)	AmoyDx Super-ARMS *EGFR* T790M mutation detection kit only *EGFR* T790M Sensitivity: 49% (95% CI, 41%–57%) Specificity: 78% (95% CI, 67%–86%)
Clinical Outcome	NA	ORR T790M tumor+/plasma+: 64% (95% CI, 57%–70%; n = 235) T790M tumor+/plasma-: 70% (95% CI, 62%–77%; n = 155)	
References	20-25	11-13,18,23	26-28

*Analog detection (also known as real-time PCR) uses fluorescent markers that attach to specific mutation sites, making them detectable; fluorescence in the sample is detected as a whole. In digital PCR, the sample is separated into many partitions, each tested individually, providing a lower threshold of detection.

**Can detect the level of ctDNA at or above threshold limit, fully quantitative can measure the level ctDNA in a sample.

Abbreviations: FDA, U.S. Food and Drug Administration; PCR, polymerase chain reaction; MAF, mutant allele frequency; IFUM, Iressa Follow up Measure Study, NGS, next-generation sequencing; ORR, overall response rate; NPV, negative predictive value; PPV, positive predictive value; ITT, intention-to-treat; HR, hazard ratio; PFS, progression-free survival; NA, not applicable.

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TABLE 2. Survival Outcome (Progression-Free Survival, Overall Survival, and Survival After Osimertinib Progression) of *EGFR* T790M Shedders Versus Nonshedders

	FLAURA [31]								
	Overall (n = 556)			Shedders (n = 359)			Nonshedders (n = 124)		
EGFR Variable	Osiemrtinib	Gefitinib/ Erlotinib	HR (95% CI); p Value	Osimertinib	Gefitinib/ Erlotinib	HR (95%CI); p Value	Osimertinib	Gefitinib/ Erlotinib	HR (95% CI); p Value
Patients, n	279	277		183	176		60	64	
Median PFS (95% CI)	18.9 (15.2–21.4) mo	10.2 (9.6–11.1) mo	0.46 (0.37–0.57); p < .0001	15.2 (13.7–20.7) mo	9.7 (8.4–11.1) mo	0.44 (0.34–0.57); p < .0001	23.5 (17.8–24.3) mo	15.0 (9.7–18.3) mo	0.48 (0.28–0.80); p = .0047
	AURA3								
	Overall (n = 419)			Shedders (n = 272)			Nonshedders		
	Osimertinib	Chemotherapy	HR (95% CI); p Value	Osimertinib	Chemotherapy	HR (95% CI); p Value	Osimertinib	Chemotherapy	HR (95% CI); p Value
Patients, n	279	140		116	56		NR	NR	
Median PFS (95% CI)	10.1 (8.3–12.3) mo	4.4 (4.2–5.6) mo	0.30 (0.23–0.41); p < 0.0001	8.2 (6.8–97) mo	4.2 (41.–5.1) mo	0.42 (0.29–0.61)	NR	NR	NR
	AURA Phase I and AURA Extension Study (National Taiwan University Hospital study)[32]								
	Shedder	Nonshedder	p Value	Shedder Without Brain Metastases	Nonshedder Without Brain Metastases	p Value	T790M "Drop Out" at Osimertinib Progression	Nonshedders at Osimertinib Progression	Shedders (New or Continual) at Osimertinib Progression
Patients, n	40	13		23	12		11	12	17
Median PFS (during osimertinib treatment)	9.4 (7.1–13.8) mo	15.4 (9.6–NR) mo	.055	12.4 (8.4–20.9) mo	15.4 (11.1–NR) mo	.26	2.6 (1.3–NR) mo	11.1 (9.4–NR) mo	12.6 (8.4–13.9) mo
Median OS (during osimertinib treatment)	13.5 (9.7–21.7) mo	36.5 (33.3–NS) mo	.0157	19.8 (11.7–NR) mo	36.5 (28.4–NR) mo	.089	9.3 (5.3–NA) mo	22.4 (15.6–NA) mo	15.4 (11.7–36.5) mo
Patients, n	38	9		21	8		11	12	17
Median survival after osimertinib progression	5.0 (3.8–9.0) mo	14.5 (7.2–NR) mo	.008	7.7 (5.0–NR) mo	17.8 (12.2–NR) mo	.031	5.0 (3.0–NA) mo	10.8 (7.2–NA) mo	5.0 (3.1–16.5) mo

Abbreviations: HR, hazard ratio; PFS, progression-free survival; NR, not reached; NA, not available; OS, overall survival.

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.6

Detection of *RAS* and *BRAF* Mutations in Colorectal Cancer

Retrospective analysis of two randomized phase III trials in metastatic colorectal cancer (mCRC; PRIME, CRYSTAL) comparing the addition of anti-EGFR antibody to chemotherapy versus chemotherapy alone indicated that multiple *KRAS, NRAS*, and *BRAF* mutations modulate the clinical response to anti-EGFR antibodies.[35,36] In particular, an additional 14.7% (CRYSTAL) to 17% (PRIME) of non–exon 2 *KRAS* mutations were detected, which also did not benefit from EGFR monoclonal treatment, hence the detection of KRAS mutation should be extended to non–exon 2 *KRAS* mutations. The E.U. has approved three sets of liquid biopsy kits from two different companies that detect an extended spectrum of *RAS* mutations to select for patients with mCRC who will potentially benefit from anti-EGFR therapy and avoid the unnecessary expense and additional side effects of anti-EGFR antibody treatment.

OncoBEAM RAS CRC kit (*KRAS* and *NRAS*). Several studies using different tissue comparator tests provided the evidence for approval of OncoBEAM RAS detection kit (Sysmex Inostics; Table 3) for the detection of *KRAS* and *NRAS* mutations in mCRC.[37-42] Not surprisingly, when a much more sensitive tissue comparator test, such as NGS, is used, the sensitivity and specificity of BEAMing technology are not as high (e.g., with the Sanger sequencing method as the comparator; Table 3).[43] The additional use of BEAMing to detect *KRAS* mutations includes real-time monitoring of treatment, detection of resistance mechanisms, and early detection of relapse and/or residual disease; these have been concisely reviewed by an expert task force panel.[44]

Idylla ctKRAS mutation test and Idylla ctNRAS-BRAF mutation test. The RASANC study provided the supporting evidence for the approval of the Idylla tests (Biocartis, Inc., Jersey City, NJ) for the detection of *KRAS, NRAS,* and *BRAF* in mCRC.[45,46] Primary tumor removal, metachronous status, absence of liver metastases, and peritoneal carcinomatosis were significantly associated with mutant *RAS* tumor and negative plasma status, indicating the difference between shedders and nonshedders.[45]

In head and neck cancer, additional use of liquid biopsy to detect *RAS* mutations has been reported to identify resistance to cetuximab (anti-EGFR antibody)-based therapy.[47] However, the resistance to anti-EGFR antibody therapy is not limited to *RAS* mutations based on liquid biopsy.[48-50]

BRAF V600E

BRAF V600E mutation is a validated and actionable driver mutation in melanoma and NSCLC.[51,52] Surprisingly, to date there is little published literature on the clinical use of liquid biopsy to detect *BRAF* V600E mutation in melanoma and NSCLC.[53] Most of the literature on liquid biopsy for *BRAF* V600E is in mCRC, especially when used in extended-spectrum *RAS* mutational analysis to select patients for anti-EGFR treatment (Table 3). Not surprisingly, a greater amount of *BRAF* V600E detected at liquid biopsy was associated with worse prognosis for survival in mCRC (hazard ratio [HR], 7.33; 95% CI, 1.04–2.89; p = .002) in a multivariate analysis that included age, tumor location, carcinoembryonic antigen (CEA), *KRAS* mutational status, and amount of circulating cell-free DNA.[54]

PIK3CA Mutations

Although no approved drug specifically targets *PIK3CA* mutations, and *PIK3CA* mutations have not been conclusively shown to be a driver mutation, *PIK3CA* mutations may confer sensitivity to PIK3CA inhibitors. Two randomized phase III trials have investigated the potential clinical benefit of the addition of buparlisib (BKM210), a class 1 pan-PIK3CA inhibitor, to fulvestrant in hormone receptor–positive, HER2-negative patients whose disease is refractory to aromatase inhibitors (AIs; BELLE-2) and additionally refractory to mTOR inhibition (BELLE-3).[55,56] Both trials investigated the presence of plasma *PIK3CA* mutations, established by using BEAMing technology, and its implication for clinical outcome. Overall, the most common *PIK3CA* mutations identified from ctDNA from BELLE-3 were H1047R (39.7%), E545K (36.8%), E542K (18.4%), and H1047L (5.1%).

The sensitivity of the plasma *PIK3CA* mutations compared with tissue *PIK3CA* mutations in BELLE-2 was 71.2% (99/139), specificity was 79.2% (244/307), and overall concordance was 76.7% (342/446). Similarly, for BELLE-3, sensitivity was 80% (70/87), specificity was 87% (142/163), and concordance was 84.8% (212/250). The lower concordance in BELLE-2 probably results from the fact that detection of *PIK3CA* mutation in tumor is performed by using Sanger sequencing, which is not as sensitive as BEAMing technology. The presence of *PIK3CA* mutations in plasma corresponded to improved PFS in patients who received buparlisib in addition to fulvestrant compared with patients who received only fulvestrant in both BELLE-2 and BELLE-3 trials (Table 4).[55,56] However, neither trial analyzed the prognostic significance of shedders versus nonshedders of *PIK3CA* mutations.

The role of *PIK3CA* mutation in breast cancer is being further evaluated in two phase III trials: SOLAR-1 (NCT02437318) is comparing alpelisib (a PI3KA inhibitor) and fulvestrant versus fulvestrant alone, with a prospective analysis of *PIK3CA* mutations in ctDNA, and SANDPIPER (NCT02340221) is comparing taselisib (GDC-0032) and fulvestrant versus fulvestrant alone in *PIK3CA*-mutant breast cancer as determined by Cobas testing. Otherwise, the eligibility criteria for both trials are similar to those of BELLE-2 (disease that is refractory to AI). If these trials will lead to the eventual approval of PIK3CA inhibitors, then *PIK3CA* mutations detection will become very important.

The prognostic significance of pretreatment plasma *PIK3CA* in cervical cancer was also investigated in 117 patients with cervical cancer treated in Hong Kong over a period of 10 years (1997 to 2007). ddPCR was used to detect *PIK3CA* E542K and E545K mutations.[57] Overall, 26 of 117 (22.2%) patients were identified (23.7% in squamous cell carcinoma of

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.7

TABLE 3. Performance Characteristics of OncoBEAM RAS CRC Kit, Idylla ctKRAS, and Idylla ctNRAS-BRAF Mutation Tests

	OncoBEAM RAS CRC Kit	Idylla ctKRAS Mutation Test	Idylla ctNRAS-BRAF Mutation Test
CE Mark Approval Date	April 2016	November 2017	November 2017
Technology	BEAMing digital PCR	Idylla real-time PCR	Idylla real-time PCR
Amount of Plasma Needed, mL	3–4	1	1
Hands-on Time, min	Dedicated technician	< 1	< 1
Turnaround Time	1 wk	130 min	110 min
Gene Coverage	34 mutations: 16 *KRAS* and 18 *NRAS* Exon 2: codons 12 and 13 Exon 3: codons 59 and 61 Exon 4: codons 117 and 146	21 *KRAS* Exon 2: codons 12 and 13 Exon 3: codons 59 and 61 Exon 4: codons 117 and 146	18 *NRAS* Exon 2: codons 12 and 13 Exon 3: codons 59 and 61 Exon 4: codons 117 and 146 5 *BRAF* Exon 15: codon 600
Sensitivity in Plasma, %	0.02–0.04	0.5	0.5
MAF Quantification	Quantitative	Semiquantitative	Semiquantitative
Clinical Validity Measures	OncoBEAM RAS CRC Kit, Instructions for Use OBMRASIVD[37] Tissue comparator test: standard of care tissue testing Sensitivity: 92.6% (112/121) Specificity: 94.0% (110/117) Concordance: 93.3% (222/238) Grasselli et al, 2017 (n = 146)[38] Tissue comparator test: PCR Sensitivity: 89% Specificity: 90% Concordance: 89.7% PPV: 84% NPV: 93% Vidal et al, 2017 (n = 115)[39] Tissue comparator test: comparison: standard of care Sensitivity: 92.6% (112/121) Specificity: 94% (110/117) Concordance: 93.3% (222/238) Schmiegel et al, 2017 (n = 98)[40] Tissue comparator test: Sanger sequencing Sensitivity: 90.4% (47/52) Specificity: 93.5% (43/46) Concordance: 91.8% (90/98) Normanno et al, 2017 (n = 92)[43] Tissue comparator test: NGS Sensitivity: 69.7% Specificity: 83.1% Concordance: 78.3% (72/92) PPV: 69.7%; 83.1% Hahn et al, 2015 (n = 46)[42] Comparator tumor test: Sanger sequencing or pyrosequencing Sensitivity: 92% (23/25) Specificity: 100% (21/21) Concordance: 96% (44/46)	RASANC study (n = 198)[45,46] Comparator test: plasma-based NGS analysis with sensitivity of 0.2% *KRAS*+ (n = 84) *NRAS*+ (n = 6) *BRAF*+ (n = 13) *RAS:* Sensitivity: 97.8% (90/92) *RAS:* Specificity: 95.3% (101/106) *RAS:* Concordance: 96.5% (191/198) *KRAS:* Concordance: 96% *NRAS:* Concordance: 100% *BRAF:* Concordance: 99.5%	
Clinical Outcome	Grasselli et al, 2017 (n = 146)[38] Median OS: Tissue *KRAS*-wild-type: 39.1 mo Tissue *KRAS*-mutated: 28.7 mo Plasma *KRAS*-wild-type: 42.9 mo Plasma *KRAS* mutated: 27.8 mo	NR	

Abbreviations: CE, Conformité Européene; PCR, polymerase chain reaction; MAF, mutant allele frequency; PPV, positive predictive value; NPV, negative predictive value; NGS, next-generation sequencing; NR, not reported.

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.8

TABLE 4. Clinical Outcomes of BELLE-2 and BELLE-3 According to *PKI3CA* Mutation Status in Tumor and Plasma

	Overall Population		ctDNA *PIK3CA*+		ctDNA *PIK3CA*–		Tumor *PIK3CA*+		Tumor *PIK3CA*–	
Variable	Buparlisib and Fulvestrant	Placebo and Fulvestrant	Buparlisib and Fulvestrant	Placebo and Fulvestrant	Buparlisib and Fulvestrant	Placebo and Fulvestrant	Buparlisib and Fulvestrant	Placebo and Fulvestrant	Buparlisib and Fulvestrant	Placebo and Fulvestrant
BELLE-2*										
Patients, n	576	571	87	113	199	188	21	19	123	120
ORR, % (95% CI)	11.8 (9.3–14.7)	7.7% (5.7–10.2)	18.4 (10.9–28.1)	3.5 (1.0–8.8)	11.6 (7.5–16.8)	10.6% (6.6–16.0)	NR	NR	NR	NR
Median PFS, mo (95% CI)	6.9 (6.8–7.8)	5.0 (4.0–5.2)	7.0 (5.0–10.0)	3.2 (2.0–5.1)	6.8 (4.7–8.5)	6.8 (4.7–8.6)	4.4 (1.6–NE)	10.7 (3.0–NE)	5.1 (3.5–8.5)	4.7 (3.3–8.5)
HR (95% CI)	0.78 (0.67–0.89)		0.58 (0.41–0.82)		1.02 (0.79–1.30)		1.18 (0.4 9–2.85)		0.98 (0.72–1.32)	
p Value	.00021 (1-sided)		.001 (1-sided)		.557 (1-sided)		NR		NR	
BELLE-3**										
Patients, n	289	143	100	35	132	81	75	34	135	69
ORR, % (95% CI)	3.9 (2.8–4.2)	1.8 (1.5–2.8)	4.2 (2.8–6.7)	1.6 (1.4–2.8)	3.9 (2.8–4.3)	2.7 (1.5–3.6)	4.7 (2.9–6.7)	1.4 (1.4–2.2)	2.8 (1.9–3.4)	1.7 (1.4 -2.9)
Median PFS, mo (95% CI)	0.67 (0.53–0.84)		0.46 (0.29–0.73)		0.73 (0.53–1.00)		0.39 (0.23–0.65)		0.81 (0.59–1.12)	
HR (95% CI)	.0030		.0031		.026		< .0001		.099	

*BELLE-2: Women who were estrogen receptor–positive or progesterone receptor–positive, human epidermal growth factor receptor 2–negative, aromatase inhibitor refractory. ctDNA *PKI3CA* determined by BEAMing polymerase chain reaction; tumor *PIK3CA* determined by Sanger sequencing.

**BELLE-3: Women who were estrogen receptor-positive or progesterone receptor–positive, human epidermal growth factor receptor 2–negative, aromatase inhibitor refractory, and refractory to mTOR inhibition. ctDNA *PIK3CA* determined by Inostics BEAMing polymerase chain reaction, covering exons 9 and 20; tumor *PIK3CA* determined by Cobas PIK3CA assay covering exons 7, 9, and 20.

Abbreviations: ORR, overall response rate; NR, not reported; NE, not evaluable; PFS, progression-free survival; HR, hazard ratio.

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.9

the cervix and 15.0% in adenocarcinoma of the cervix). The presence of plasma *PIK3CA* mutations was associated with high pathologic grade and large tumor size (mean tumor size, 4 cm) compared with 3 cm for patients negative for plasma *PIK3CA* mutations but was not associated with age, clinical stage, presence or absence of lymphovascular invasion, or pelvic lymph node metastasis. Although *PIK3CA* mutation was associated with shortened disease-free survival and OS on univariate analysis, on multivariate analysis *PIK3CA* mutation did not affect disease-free survival (HR 1.449; p = .464) or OS (HR 1.261; p = .643). Going forward, in analyses of larger cohorts of patients with cervical cancer who have uniform stage, histology, histologic grade, and treatment modality and are from a more contemporaneous treatment period, plasma *PIK3CA* mutations may turn out to be an important prognostic factor in cervical cancer. Liquid biopsy for *PIK3CA* mutation has also been investigated in small numbers of patients with bladder cancer and mCRC.[58,59]

HER2 Amplification

HER2 amplification was validated as an actionable target in breast cancer in 2001 and in gastric adenocarcinoma in 2010.[60,61] The gold standard of detecting *HER2* amplification in breast cancer is by immunohistochemistry or florescence in situ hybridization. Given the generally ready availability of tumor tissue from patients with breast cancer, there is scant literature on the use of liquid biopsy to detect or to monitor *HER2* amplification at diagnosis or during treatment, respectively. Small studies have demonstrated that *HER2* amplification detected in plasma correlates well with tumor *HER2* amplification and that the level of *HER2* amplification can be used to monitor treatment and recurrence in gastric cancer.[62,63] *HER2* amplification was successfully identified by using a liquid biopsy NGS panel to identify driver mutations to enroll in a basket targeted therapy trial that included HER2 inhibition.[64]

USE OF LIQUID BIOPSY TO DETERMINE RESISTANCE MECHANISMS TO HORMONAL BLOCKADE IN SOLID TUMORS

Estrogen Receptor (*ESR1*) Mutation

Detection of *ESR1* mutation (primarily in the ligand-binding domain) in plasma has been investigated as a resistance mechanism to an AI in two randomized phase III trials, SoFEA (Study of Faslodex Versus Exemestane With or Without Arimidex) and PALOMA3 (Palbociclib Combined With Fulvestrant in Hormone Receptor–Positive HER2-Negative Metastatic Breast Cancer After Endocrine Failure), using the Bio-Rad QX-200 ddPCR system. In the SoFEA analysis, patients with mutant *ESR1* in the exemestane treatment arm had significantly worse PFS (p = .01) than those without mutated *ESR1*. Similarly, PFS was significantly improved (p = .02) in patients with mutant *ESR1* when they received fulvestrant-containing regimens but not (p = .77) in patients with wild-type *ESR1* (Table 5). On the other hand, not surprisingly, the retrospective analysis of the PALOMA3 showed that both *ESR1*-mutated (HR 0.43; 95% CI, 0.25–0.74; p = .002) and *ESR1* wild-type (HR 0.49; 95% CI, 0.35–0.70; p < .001) patients benefited from the addition of palbociclib to fulvestrant compared with those receiving placebo and fulvestrant because palbociclib is a CDK4/6 inhibitor (Table 6).[65]

Together, these two trials suggest that for patients who progressed while receiving a nonsteroidal AI and have detectable *ESR1* mutations, it is best to switch to fulvestrant-containing regimens, such as fulvestrant and palbociclib. For patients who progressed while receiving a nonsteroidal AI and have no detectable *ESR1* mutations, switching to an irreversible AI, such as exemestane, remains a treatment option before proceeding to fulvestrant-contain regimens. These are hypothesis-generating observations because only 161 of 723 (22.3%) patients enrolled in the SoFEA and only 360 of 551 (65.3%) patients

Table 5. Effect of *ESR1* Mutations on Clinical Outcome in Patients Receiving Endocrine Therapies in SoFEA Trial

	Exemestane		
Variable	***ESR1* Wild-Type**	***ESR1*-Mutated**	**Hazard Ratio (95% CI); p Value**
Patients (n)	39	18	
Median PFS, mo (95% CI)	8.0 (3.0–11.5)	2.6 (2.4–6.2)	2.12 (1.18–3.81); .01
Median OS, mo (95% CI)	22.8 (17.6–32.4)	12.8 (5.7–27.0)	1.65 (0.81–3.38); .16
	***ESR1*-Mutated**		
	Fulvestrant Regimens	**Exemestane**	
Patients (n)	45	18	
Median PFS, mo (95% CI)	5.7 (3.0–8.5)	2.6 (2.4–6.2)	0.52 (CI, 0.30–0.92); .02
	***ESR1* Wild-Type**		
	Fulvestrant Regimens	**Exemestane**	
Patients (n)	59	39	
Median PFS, mo (95% CI)	8.0 (3.0–11.5)	5.4 (3.7–81.)	1.07 (0.68–1.67); .77

Abbreviation: PFS, progression-free survival.

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.10

enrolled in PALOMA3 had plasma available and successfully analyzed.

In the PALOMA3 trial, logistic multivariate analysis showed that the presence of plasma *ESR1* mutations was negatively associated with tamoxifen treatment only (odds ratio [OR], 0.06; 95% CI, 0.01–0.45; p = .01) and positively associated with sensitivity to endocrine therapy (OR, 3.95; 95% CI, 1.59–9.78; p = .003), bone metastasis (OR, 3.31; 95% CI, 1.54–6.39; p = .002), and visceral metastasis (OR, 1.74; 95% CI, 1.02–2.98; p = .04). Importantly, multivariate analysis of the PFS benefit in PALOMA3 showed a negative association with plasma *ESR1*-mutated genotype (HR 1.4; 95% CI, 1.07–2.08; p = .02), even after factoring in the significant benefit from palbociclib.[65]

Recent analysis of a large trial of first-line AI therapy in breast cancer indicated that *ESR1* mutations may play a role in AI resistance, but other resistance mechanism, such as *KRAS,* have been identified.[66] Finally, liquid biopsy using NGS indicated multiple resistance mechanisms to AI in addition to *ESR1* mutations.[67]

Androgen Receptor Variant 7

A splice variant of the androgen receptor variant 7 (*AR-V7*) that deletes the ligand domain of *AR* isolated from circulating tumor cells in the plasma is involved in de novo and acquired resistance to abiraterone (an androgen synthesis inhibitor) and enzalutamide (an *AR* ligand–binding domain antagonist), both of which have been approved for the treatment of castration-resistant prostate cancer (Table 7). In multivariable Cox regression analysis stratified by treatment type, *AR-V7* detection remained independently predictive of prostate-specific antigen–defined PFS (HR 8.2; 95% CI, 2.7–24.9; p < .001), clinical/radiographic PFS (HR 4.9; 95% CI, 1.7–13.8; p = .003), or OS (HR 5.0; 95% CI, 1.3–19.8; p = .021; Table 7).[68] Further study has indicated that *AR* amplification was associated with resistance to enzalutamide, and mutations (*H874Y, T877A, D879E, L881I, E893K,* and *M895V)* in the ligand-binding domain of *AR* are more associated with resistance to abiteratone.[69] More recently, a longitudinal multiplex targeting sequencing of plasma from a study of 65 patients with castration-resistant prostate cancer receiving enzalutamide indicated that besides *AR* amplification of *AR* mutations, *RB1* loss (p = .01) and *MET* copy number gain/amplification (p = .2) were significantly associated with shorter PFS after adjustment for the presence of ctDNA.[70] Overall mutations in the direct target of hormonal blockade can partially explain the resistance to hormonal treatment of breast and prostate cancer. However, the advent of NGS liquid biopsy has allowed the discovery a much more complex pattern of resistance mechanisms.

NEXT-GENERATION SEQUENCING PLATFORMS

Liquid biopsy using NGS will be the dominant platform, with its increased depth of coverage, simultaneous multigene sequencing, and ability to simultaneously detect all modes of genomic alterations (point mutations, insertion/deletion [indel], amplification, and gene rearrangements) to account for both tumor heterogeneity and multiple resistance mechanisms and determination of tumor mutation burden (TMB). The two most common commercially used liquid biopsies are Guardant360 (Guardant Health, Redwood City, CA) and FoundationACT (Foundation Medicine Inc., Cambridge, MA). As listed in Table 8, the limit of detection and

Table 6. Effect of *ESR1* Mutations on Clinical Outcome in Patients Receiving Endocrine Therapies in PALOMA3 Trial

	ESR1-Mutated		
Variable	**Palbociclib + Fulvestrant**	**Placebo + Fulvestrant**	**Odds Ratio or Hazard Ratio (95% CI); p Value Statistical Consideration**
Patients, n	63	28	
ORR, %	17.5	3.6	5.94* (0.78–270.5); .06
Median PFS, mo (95% CI)	9.4 (5.3–11.1)	3.6 (2.0–5.5)	0.43** (0.25–0.74); .002
	***ESR1* Wild Type**		
	Palbociclib + Fulvestrant	**Placebo + Fulvestrant**	
Patients, n	177	92	
ORR,%	21.5	14.1	1.66* (0.81–3.61); .14
Median PFS, mo (95% CI)	9.5 (9.2–NE)	5.4 (3.5–7.4)	0.49** (0.35–0.70); ≤ .001
	***ESR1* Status Available**		
	***ESR1* Wild Type**	***ESR1*-Mutated**	
Patients, n	261	91	
Median PFS (95% CI)	9.2 (7.4–10.9)	7.3 (3.7–9.1)	1.46** (1.06–2.02); .02

*Odds ratio (univariate analysis).

**Hazard ratio (univariate analysis).

Abbreviations: ORR, overall response rate; NE, not evaluable.

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.11

TABLE 7. Androgen Receptor Variant 7 Detected by Quantitative Reverse Transcriptase Polymerase Chain Reaction in Circulating Tumor Cells as Both Primary and Acquired Resistance to Abiraterone and Enzalutamide

	Abiraterone (n = 31)		Enzalutamide (n = 31)	
	Pretreatment AR-V7–Positive	Pretreatment AR-V7–Negative	Pretreatment AR-V7–Positive	Pretreatment AR-V7–Negative
Patients, n	6	25	12	19
PSA ORR, % (95% CI)	0 (0–46)	68 (46–85)	0 (0–26)	53 (29–76)
p Value	.004		.004	
PSA PFS, mo (95% CI)	1.3 (0.9–NR)	5.3 (5.3–NR)	1.4 (0.9–NR)	6.0 (3.8–NR)
HR (95% CI)	16.1 (3.9–66.0)		7.4 (2.7–20.6)	
p Value	< .001		< .001	
PFS*, mo (95% CI)	2.3 (1.4–NR)	6.3 (6.3–NR)	2.1 (2.0–NR)	6.1 (4.7–NR)
HR (95% CI)	16.5 (3.3–82.9)		8.5 (2.8–25.5)	
p Value	< .001		< .001	
OS, mo	10.6	NR	5.5	NR
HR (95% CI)	12.7 (1.3–125.3)		6.9 (1.7–28.1)	
p Value	.006		.002	

	Pretreatment (n = 62)		Pretreatment AR-V7–Negative (n = 42)	
	AR-V7–Positive	AR-V7–Negative	Post-Treatment AR-V7–Negative	Post-Treatment AR-V7–Positive
Patients, n	18	44	36	6
ORR, % (95% CI)	0 (0–19)	61 (45–76)	68 (52–81)	17 (4–58)
p Value	< .0001		NA	
PSA PFS, mo (95% CI)	1.4 (0.9–2.6)	6.1 (5.3–NR)	6.1 (5.9–NA)	3.0 (2.2–NA)
HR (95% CI)	10.5 (4.7–23.6)		NA	
p Value	< .001		NA	
PFS*, mo (95% CI)	2.1 (1.9–3)	6.4 (6.1–NR)	6.5 (6.1–NA)	3.2 (3.1–NA)
HR (95% CI)	12.7 (5.1–31.9)		NA	
p Value	< .001		NA	
OS (95% CI)	9.2 (4.5–NR)	>11.9 (11.9–NR)	NA	NA
HR (95% CI)	8.3 (2.5–27.4)		NA	
p Value	< .001		NA	

*Clinical or radiographic; PSA response is defined as ≥50% decline in baseline PSA maintained for ≥4 weeks.

Abbreviations: AR-V7, androgen receptor-variant 7; PSA, prostate-specific antigen; ORR, overall response rate; PFS, progression-free survival; NR, not reached; HR, hazard ratio; OS, overall survival; NA, not available.

the gene panel of Guardant360 has evolved over time. It is important to be aware of the number of exact genes included in the NGS sequencing platform by consulting with the company websites because technology advances will lead to increased complexities of the NGS panel being offered commercially. It is also important to note how the interpretation of the literature depends on the various versions of Guardant360.

Among the many other NGS liquid biopsy platforms that are available or under development are Oncomine Lung cfDNA Assay (ThermoFisher Scientific, Waltham, MA),[88] Archer Reveal ctDNA (ArcherDX, Boulder, CO),[89] Oncotype SEQ (Genomic Health, Redwood City, CA),[90] LiquidDx (MolecularMD, Portland, OR),[91] CancerIntercept Detect and CancerIntercept Monitor (Pathway Genomics, San Diego, CA),[92] OptiSeq NGS Pan-Cancer Panel (DiCarta, Richmond, CA),[93] and PlasmaSELECT (Personal Genome Diagnostics, Baltimore, MD).[94]

USE OF LIQUID BIOPSY TO DETERMINE TUMOR MUTATION BURDEN

TMB is usually expressed as number of mutations per megabase of DNA (MB)[95] and has been shown to be predictive to response to anti–PD-L1 therapy.[96] In the CheckMate026 trial, which compared nivolumab to platinum-based chemotherapy as first-line treatment of NSCLC, OS did not differ between nivolumab and chemotherapy. A subgroup analyzed with TMB, as determined by whole-exome sequencing, indicated

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.12

TABLE 8. Comparison of Various and Most Current Versions of Guardant360 and Foundation Medicine FoundationACT Liquid Biopsy Platforms (as of December 31, 2017)

	Guardant360 54-Gene Panel	Guardant360 68-Gene Panel	Guardant360 70-Gene Panel	Guardant360 73-Gene Panel*[71]	Foundation Medicine FoundationACT 62-Gene Panel*[72]
Commercial Launch Date	June 2014	August 2015	October 2015	November 2016	June 2016
Amount of Blood Needed	Two 10-mL Streck tubes	Two 10-mL Streck tubes	Two 10-mL Streck tubes	Two 10-mL Streck tubes	Two 10-mL tubes (Foundation ACT Specimen Collection Kit)
Minimal Amount of Cell-Free DNA	> 5 ng of cell-free DNA1	> 5 ng of cell-free DNA1	> 5 ng of cell-free DNA1	> 5 ng of cell-free DNA1	≥ 20 ng of cell-free DNA
Methods for Isolating ctDNA	QIAamp circulating nucleic acid kit (Qiagen)	QIAamp circulating nucleic acid kit (Qiagen)	QIAamp circulating nucleic acid kit (Qiagen)	QIAamp circulating nucleic acid kit (Qiagen)	
Types of Alterations Analyzed and No. of Genes	Complete exons (18 genes)	Complete exons (29 genes)	Complete exons (30 genes)	Complete exons (19 genes)**	Complete exons (27 genes)
	Critical exons (36 genes)	Critical exons (39 genes)	Critical exons (40 genes)	Critical exons (54 genes)	Critical exons (33 genes)
	CNV (3 genes)	Indel (1 gene)	Indels (3 genes)	Indels (23 genes)	Indel (27genes)
		CNV (16 genes)	CNV (18 genes)	CNV (18 genes)	Amplification (27 genes)
		Fusions (4 genes)	Fusions (6 genes)†	Fusions (6 genes)†	Fusions (6 genes)†‡ Promoter region only (TERT)
Sequencing Platform	HiSEquation 2500 (Illumina)1,2	HiSeq 2500 or NextSEquation 500 (Illumina)	HiSeq 2500 or NextSEquation 500 (Illumina)	HiSeq 2500 or NextSEquation 500 (Illumina)	HiSeq 2500 or HiSeq 4000 (Illumina)
Depth of Coverage	?10,000× (average raw)			~15,000x (average raw)	7,503× (median, unique, actual depth of coverage reported)
Single-Nucleotide Variation	Varies between < 0.1% and 0.1%; depends on particular gene	Varies between < 0.1% and 0.1%; depends on particular gene	Varies between < 0.1% and 0.1%; depends on particular gene	≥ 0.04%; ≥ 0.25%, sensivitiy 100%, PPV 99.2%; 0.05-0.25% sensitivity 63.8%, PPV 96.3%	≥ 0.1% (0.1%–0.5%; sensitivity, 67.3%; PPV, 93.6%) ≥ 0.05% (sensitivity > 98.9%; PPV > 99.9%)
Indel	Varies between < 0.1% and 0.1%; depends on particular gene	Varies between < 0.1% and 0.1%; depends on particular gene	Varies between < 0.1% and 0.1%; depends on particular gene	≥ 0.02%; ≥ 0.02%, sensitivity 100%, PPV 98.2%; 0.05-0.2% sensitivity 67.8%, PPV 98.2%	≥ 1% (sensitivity, 99.9%; PPV, 98.8%)
CNV/Amplification	Varies between < 0.1% and 0.1%; depends on particular gene	Varies between < 0.1% and 0.1%; depends on particular gene	Varies between < 0.1% and 0.1%; depends on particular gene	≥ 2.12 copies	≥ 20% (sensitivity, 95.3%; PPV, 97.5%)
Fusions/Rearrangement		Varies between < 0.1% and 0.1%; depends on particular gene	Varies between < 0.1% and 0.1%; depends on particular gene	≥ 0.04%; ≥ 0.2% sensitivity 95%, PPV 100%; 0.05–0.2% sensitivity 83%, PPV 100%	≥ 1% (sensitivity > 99.9%; PPV, 98.8%) < 1% (sensitivity, 86.8%; PPV > 99%)
Turnaround Time for Results			Within 2 wk of receiving blood sample	Median of 7 calendar days of receiving blood sample	Within 2 wk of receiving blood sample

Continued

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TABLE 8. Comparison of Various and Most Current Versions of Guardant360 and Foundation Medicine FoundationACT Liquid Biopsy Platforms (as of December 31, 2017) (Cont'd)

	Guardant360 54-Gene Panel	Guardant360 68-Gene Panel	Guardant360 70-Gene Panel	Guardant360 73-Gene Panel*[71]	Foundation Medicine FoundationACT 62-Gene Panel*[72]
Difference in Gene Panel		ABL, CSF1R, ERBB4, FLT3, KDR (VEGFR2), PROC, RB1, and SMARCB1 genes were deleted and replaced with ARAF, ARID1A, BRCA1, BRCA2, CCND1, CCND2, CCNE1, CDK4, CDK6, CDKN2B, ESR1, GATA3, MAP2K1, MAP2K2, NF1, NTRK1, NTRK3, NFE2L2, RHEB, RHOA, RIT1, RAF1, and ROS1 genes. Copy number amplification calling increased from 3 to 16 genes (added AR, BRAF, CCND1, CCND2, CCNE1, CDK4, CDK6, FGFR1, FGFR2, KIT, KRAS, MYC, PDGFRA, PIK3CA, RAF1) Fusing calling added for 4 genes (ALK, ROS1, RET, NTRK1). Indel calling added for EGFR (exon 19/20)	RB1 and TSC1 genes were added; number of fusions detected increased from 4 to 6 (additional FGFR2 and FGFR3 fusion detection). Indel coverage for ERBB2 (exon19/20) and MET exon 14 splice site variants was added	SRC and CDKN2B were deleted from panel and DDR2, MAPK1(ERK2), MAPK3 (ERK1), MTOR, and NTRK3 were added. Copy number amplification calling was increased to 18 genes (CCND1, CCND2). Indel coverage expanded from 3 to 23 genes (ATM, APC, ARID1A,BRCA1, BRCA2, CDH1,CDKN2A, GATA3, KIT, MLH1,MTOR, NF1, PDGFRA,PTEN,RB1,SMAD4,STK11,TP53,TSC1,VHL)	NA
Cost	NA	NA	NA	$5,800 per test ($5,000 suggested contracted price per test)	$5,800 USD per test (contracted price may vary)
References (Selected Clinical Use of One Particular Platform)	73-76	75-79	64,75,76,80,81	82,83	84-87

*Although the difference in the number of genes is 11 between the current Guardant360 and Foundation Medicine FoundationACT, the actual difference in the number of genes analyzed is 31 because of unique genes analyzed by each liquid biopsy. Twenty-one genes were unique to Guardant360: AR, ARID1A, ATM, FBXW7, GATA3, HNF1A, MAPK3, MLH1, MPL, NFE2L2, NOTCH1, NTRK1, NTRK3, RB1, RHEB, THOA, RIT1, SMAD, STK11, TSC1, VHL. Ten genes were unique to Foundation Medicine FoundationACT: ABL, BTK, CRKL, ERRFI1, FOXL2, MDM2, MYD88, PDGFR-beta, PD-L2, VEGF-alpha.

**The 19 genes that have complete exon coverage in the Guardant 73-gene panel are as follows: APC, AR, BRAF, BRAC1, BRCA2, EGFR, ERBB2, HRAS, KRAS, NRAS, KIT, MAPK1(ERK2), MAPK2 (ERK1), MET, MYC, PIK3CA, RB1, TP53, STK11. Underlined genes have complete exon coverage by Foundation Medicine FoundationACT.

†Both Guardant and Foundation Medicine FoundationACT report four common fusions: ALK, ROS1, RET, and FGFR3 fusions. Guardant360 also reports additional NTRK1 and FGFR2 fusions, and Foundation Medicine FoundationACT reports additional EGFR and PDGFR-alpha fusions.

‡ROS1 rearrangement is the only ROS1 alteration detected by the current version of Foundation Medicine FoundationACT.

Abbreviations: PPV, positive predictive value; NA, not available; USD, US dollars.

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

that patients with NSCLC and high TMB had a marginally significantly improved PFS compared with those receiving platinum-based chemotherapy (9.7 months [95% CI, 5.1 months to not reached] versus 5.8 months [95% CI, 4.2–8.5 months]; HR 0.62 [95% CI, 0.38–1.00]). OS did not differ between nivolumab and chemotherapy regardless of high or low TMB.[96] Importantly, there is excellent concordance between TMB as determined by a commercially available (Foundation Medicine Inc.) tissue-based hybrid-capture NGS method and whole-exome sequencing of the tumor but essentially no correlation between TMB and PD-L1 expression.[96]

It is now feasible to determine TMB by liquid biopsy using whole-exome sequencing or NGS hybrid-capture method.[97,98] The blood TMB (bTMB) assay developed by Foundation Medicine Inc. uses 10 mL of plasma to deliver an account of somatic base substitutions down to 0.5% allele frequency across 394 genes from as little as 1% tumor content in a cell-free DNA sample. Computational methods differ: bTMB analyzes only single-nucleotide variants (SNVs), whereas tissue TMB also includes analysis of indels and fusions. Sensitivity, specificity, and positive predictive value were 93.9%, 100.0%, and 100.0%, respectively, when compared with tissue TMB.[98] Importantly, the bTMB assay demonstrates strong sensitivity and specificity between the cut-points of 8–20 mutations/MB. The clinical utility of the bTMB assay was retrospectively analyzed in 783 patients with NSCLC receiving second-line therapy (OAK and POPLAR); PFS and OS were compared between atezolizumab (an anti–PD-L1 antibody) and single-agent docetaxel as second-line treatment. At a bTMB cutoff of 16 or greater, PFS and OS significantly improved in the phase II randomized POPLAR study.[99] Among the 850 patients enrolled in the phase III OAK trial, 583 patients were considered as the "biomarkers evaluable population" (BEP) for bTMB. A bTMB of 16 or greater represented 27% of the BEP. There was an OS benefit for patients with both bTMB of 16 or greater (HR 0.64; 95% CI, 0.44–0.92) and bTMB less than 16 (HR 0.65; 95% CI, 0.52–0.81). Indeed, PFS (HR 0.73; 95% CI, 0.56–0.95) and OS (HR 0.69; 95% CI, 0.52–0.93) benefits were observed in patients with bTMB of 10 or less. A bTMB of 10 or less represents 43% of the BEP, and the sensitivity and specificity were both 100% when compared with tissue TMB.[99] The bTMB assay is being validated in two trials. B-F1RST (NCT02848651) is a phase II single-arm study of atezolizumab as first-line treatment of NSCLC while prospectively collecting plasma to evaluate whether bTMB can predict for improved clinical outcome with atezolizumab. B-FAST (NCT03178552) is prospectively enrolling patients with NSCLC receiving first-line therapy in a global, multicenter, multicohort, randomized phase II/III trial without certain targetable driver mutations (*RET* or *ALK* fusion) detected based on blood-based assays; additionally, the trial is comparing single-agent atezolizumab therapy to standard of care based on bTMB results for those patients without RET or ALK rearrangement.[100]

USE OF CTDNA IN LIQUID BIOPSY TO DETECT MINIMAL RESIDUAL DISEASE IN SOLID TUMORS

Most solid tumors do not have a single truncal actionable driver mutation. There is vast literature on the use of nonspecific ctDNA in solid malignancies ranging from detection to monitor treatment response to detection of recurrence, and even for survival prognostication, all of which are beyond the scope of this review. However, ctDNA detection by liquid NGS has been used to identify minimal residual disease (MRD); we review two important studies here.[101,102]

In a study of 230 patients with resected stage II colon cancer conducted by Tie and colleagues, 20 (8.7%) patients had ctDNA detected by massive parallel sequencing. Most of these mutations were *TP53*, *APC*, and *KRAS*. There was no difference in clinicopathologic difference (age, sex, right- versus left-sided colon, tumor differentiation, T3 versus T4, number of lymph node resected < 12 versus ≥ 12, mismatch repair status, presence or absence of lymphovascular invasion, with or without adjuvant chemotherapy) between patients with and without detectable ctDNA. Among the 178 patients not treated with adjuvant chemotherapy, postoperative recurrence occurred in 11 of 14 (78.6%) ctDNA-positive patients compared with 16 of 164 (9.8%) ctDNA-negative patients. The 3-year relapse-free survival rate was 0% for ctDNA-positive patients compared with 90% for ctDNA-negative patients (HR 18; 95% CI, 7.9–40; $p = 2.5 \times 10^{-12}$). By multivariate analysis, ctDNA positivity (HR 28; 95% CI, 11–68; p = .0001) and T stage (T4 versus T3; HR 8.1; 95% CI, 3.1–21; p < .0001) were the two independent predictive factor for RFS after factoring in the above clinicopathologic factors. In addition, among the 230 patients, ctDNA positivity (HR 14; 95% CI, 6.8–28; p < .0001) and T stage (HR 2.6; 95% CI, 3.1–5.5; p = .001) remained the two independent factors for PFS after factoring in adjuvant chemotherapy. More importantly, postoperative elevation of CEA level was not a predictive factor for relapse-free survival, even by univariate analysis with or without adjuvant chemotherapy.[101] Furthermore, ctDNA positivity immediately after completion of postoperative chemotherapy was significantly associated with poor relapse-free survival (HR 11; 95% CI, 1.8–6.8; p = .001). However, the number of patients analyzed was 44, and only three patients had ctDNA positivity immediately after completion of postoperative chemotherapy.

Finally, during the follow-up period, ctDNA was significantly more frequently positive (23/27 [85%]) than was CEA elevation (11/27 [41%]) at the time of radiographic recurrence (p = .002). The time between ctDNA positivity (median, 167 days) was significantly longer than that between CEA elevation (median, 61 days) to radiographic recurrence (p = .04). This retrospective analysis of prospective collected data indicated that ctDNA monitoring has the ability to detect MRD and hence early recurrence, leading to the potential initiation and/or lengthening of adjuvant chemotherapy based on the presence or absence of ctDNAs. There is a randomzied ctDNA driven adjuvant chemotherapy trial in resected stage II CRC in Australia/New Zealand (DYNAMIC,

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.15

Trial ID number: ACTRN12615000381583). In the standard-of-care arm, adjuvant chemotherapy wil be given as determined by the clinicians, while in the investigational arm, adjuvant chemothrapy will be guided by the presence (adjuvant chemotherapy) or absence (no adjuvant chemotheray) of postoperative ctDNA in resected stage II CRC. The primary endpoint is to evaluate whether an adjuvant therapy strategy based on ctDNA results may affect the number of patients treated with chemotherapy and recurrence-free survival.

TRACERx is a U.K. study that developed proprietary NGS technology to link the individual patient's tumor DNA to plasma ctDNA by identifying two distinct tumor-associated mutations (usually SNV).[102] The resected tumor was further microdissected at several different areas, and all "subtumors" underwent NGS sequencing to detect tumor heterogeneity.[103] The TRACERx study group used this phylogenetic ctDNA (at least two SNVs) tracking to detect recurrence in 96 patients with resected early-stage lung cancer. A group of 24 patients were followed longitudinally with pre- and postoperative plasma ctDNA; of these, 10 had been relapse free for a median of 775 days, whereas 14 had confirmed relapse. Thirteen of the 14 patients who relapsed had ctDNA detected before or at clinical relapse compared with only one of 10 patients in remission (two SNVs). Overall, the median duration between ctDNA detection and radiographic evidence of relapse was 70 days. Relapse can involve one or more subclones. This analysis has some limitations due to lack of regularly scheduled imaging to account for potential lead-time bias of regular liquid biopsy compared with detection of recurrence by clinical practice.[104] Regardless, the TRACERx approach remains an important technology that accounts for tumor heterogeneity in the primary tumor, especially because most lung cancers do not harbor a signature actionable driver alteration.

Detection of MRD may turn out to be the most important aspect of use of liquid biopsy because it may help guide extent of adjuvant treatment,[105] similar to the use of Epstein-Barr virus load in the treatment of early-stage nasopharyngeal carcinoma.[106]

SHEDDERS VERSUS NONSHEDDERS OF ACTIONABLE GENOMIC ALTERATIONS

In the TRACERx study, 46 of 96 patients had at least two SNVs that allowed phylogenetic ctDNA tracking, including 30 of 31 (97%) lung squamous cell carcinomas, 11 of 58 (19%) lung adenocarcinomas, and five of seven other NSCLCs. Among stage I NSCLC, 16 of 17 squamous cell carcinomas (94%) compared with five of 39 (13%) adenocarcinomas had detectable ctDNA. Of note, the median necrosis rate in lung squamous cell carcinoma (40%; n = 31) was significantly higher than in ctDNA-positive lung adenocarcinomas (15%; n = 11) and ctDNA-negative adenocarcinomas (2%; n = 47; p < .0001). Multivariate analysis indicated that nonadenocarcinoma histology (p = .001), lymphovascular invasion (p = .042), and high Ki-67 proliferation index (p = .022) are independent predictors of ctDNA positivity but not tumor necrosis (10% increment increase; p = .862), tumor size (1-cm increment increase; p = .134), or amount of cfDNA (p = .229). Importantly, the mean plasma variant allele frequency positively corresponded to tumor size. The TRACERx team was able to estimate that a mean clonal plasma variant allele frequency of 0.1% corresponds to a tumor burden of 10 cm^3 or a burden of 302 million tumor cells.[102]

A large-scale systemic review of 39 studies that included 4,052 patients, the detection of ctDNA was associated with a significantly worse OS in multivariable analyses [HR 2.70; 95% CI, 2.02–3.61; p < .001).[107] There was also a statistically significant association between high total cell-free DNA and worse OS at 3 years by multivariate analysis (HR 1.91; 95% CI, 1.59–2.29; p < .001).[107] Thus, the use of liquid biopsy to identify shedders and nonshedders has important clinical implication and has been proposed to be included as part of the staging workup (tumor, node, metastasis, biopsy).[10]

CONCLUSION

The advantages of liquid biopsy not only include the possibility to circumvent the need to achieve adequate tumor tissue biopsy specimens, it also reflects an aggregate of ctDNA output from potentially both primary and all metastatic sites accounting for tumor heterogeneity that cannot be evaluated by a single core tumor needle biopsy. Liquid biopsy may provide prognostic implication and help guide initiation and/or duration of adjuvant treatment by detecting MRD. Although all the current government-approved liquid biopsy test kits are designed to detect only mutations in one or a few specific genes, the future of liquid biopsy will be using NGS with increased sensitivity and ability to detect all modes of genetic alteration and TMB. The gene panel and limit of detection used by commercial NGS liquid biopsy kits change over time. It is incumbent upon the ordering physician to understand the current NGS platforms and performance characteristics and upon readers to interpret the literature with the comprehension of the changes in the gene panel and performance characteristics of commercial NGS tests as time goes by.

References

1. Wan JCM, Massie C, Garcia-Corbacho J, et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nat Rev Cancer*. 2017;17:223-238.
2. Bettegowda C, Sausen M, Leary RJ, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med*. 2014;6:224ra24.
3. Webb S. The cancer bloodhounds. *Nat Biotechnol*. 2016;34:1090-1094.
4. Newman AM, Lovejoy AF, Klass DM, et al. Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat Biotechnol*. 2016;34:547-555.

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.16

5. Oxnard GR, Paweletz CP, Sholl LM. Genomic analysis of plasma cell-free DNA in patients with cancer. *JAMA Oncol*. 2017;3:740-741.

6. Sacher AG, Komatsubara KM, Oxnard GR. Application of plasma genotyping technologies in non-small cell lung cancer: a practical review. *J Thorac Oncol*. 2017;12:1344-1356.

7. Huang WL, Chen YL, Yang SC, et al. Liquid biopsy genotyping in lung cancer: ready for clinical utility? *Oncotorget*. 2017;8:18590-18608.

8. Bernabé R, Hickson N, Wallace A, et al. What do we need to make circulating tumour DNA (ctDNA) a routine diagnostic test in lung cancer? *Eur J Cancer*. 2017;81:66-73.

9. Thierry AR, El Messaoudi S, Mollevi C, et al. Clinical utility of circulating DNA analysis for rapid detection of actionable mutations to select metastatic colorectal patients for anti-EGFR treatment. *Ann Oncol*. 2017;28:2149-2159.

10. Yang M, Forbes ME, Bitting RL, et al. Incorporating blood-based liquid biopsy information into cancer staging: time for a TNMB system? *Ann Oncol*. 2018;29:311-323.

11. U.S. Food and Drug Administration. Cobas EGFR Mutation Test v2. 2016. https://www.fda.gov/drugs/informationondrugs/approveddrugs/ucm504540.htm. Accessed January 15, 2018.

12. U.S. Food and Drug Administration. Cobas EGFR mutation test v2. http://www.accessdata.fda.gov/cdrh_docs/pdf12/P120019S007c.pdf. Accessed January 28,2018.

13. Malapelle U, Sirera R, Jantus-Lewintre E, et al. Profile of the Roche cobas® EGFR mutation test v2 for non-small cell lung cancer. *Expert Rev Mol Diagn*. 2017;17:209-215.

14. Mok TS, Wu Y-L, Ahn M-J, et al. Osimertinib or platinum-pemetrexed in EGFR T790M-positive lung cancer. *N Engl J Med*. 2017;376:629-640.

15. Wu YL, Zhou C, Liam CK, et al. First-line erlotinib versus gemcitabine/cisplatin in patients with advanced EGFR mutation-positive non-small-cell lung cancer: analyses from the phase III, randomized, open-label, ENSURE study. *Ann Oncol*. 2015;26:1883-1889.

16. Reck M, Hagiwara K, Han B, et al. ctDNA Determination of EGFR mutation status in European and Japanese patients with advanced NSCLC: the ASSESS Study. *J Thorac Oncol*. 2016;11:1682-1689.

17. Han B, Tjulandin S, Hagiwara K, et al. EGFR mutation prevalence in Asia-Pacific and Russian patients with advanced NSCLC of adenocarcinoma and non-adenocarcinoma histology: the IGNITE study. *Lung Cancer*. 2017;113:37-44.

18. Jenkins S, Yang JC, Ramalingam SS, et al. Plasma ctDNA analysis for detection of the EGFR T790M mutation in patients with advanced non-small cell lung cancer. *J Thorac Oncol*. 2017;12:1061-1070.

19. Remon J, Menis J, Hasan B, et al. The APPLE trial: feasibility and activity of AZD9291 (osimertinib) treatment on positive plasma T790M in EGFR-mutant NSCLC patients. EORTC 1613. *Clin Lung Cancer*. 2017;18:583-588.

20. European Medicines Agency. Iressa: public assessment report — product information. 2016. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR_-_Product_Information/human/001016/WC500036358.pdf.

21. European Medicines Agency. Tagrisso: public assessment report — product information. http://www.ema.europa.eu/docs/en_GB/document_library/EPAR__Product_Information/human/004124/WC500202022.pdf.

22. U.S. Food and Drug Administration. Therascreen® EGFR RGQ PCR Kit Instructions for Use (Handbook). https://www.accessdata.fda.gov/cdrh_docs/pdf12/P120022c.pdf. Accessed January 28, 2018.

23. Thress KS, Brant R, Carr TH, et al. EGFR mutation detection in ctDNA from NSCLC patient plasma: a cross-platform comparison of leading technologies to support the clinical development of AZD9291. *Lung Cancer*. 2015;90:509-515.

24. Douillard JY, Ostoros G, Cobo M, et al. First-line gefitinib in Caucasian EGFR mutation- positive NSCLC patients: a phase-IV, open-label, single-arm study. *Br J Cancer*. 2014;110:55-62.

25. Douillard JY, Ostoros G, Cobo M, et al. Gefitinib treatment in EGFR mutated Caucasian NSCLC: circulating-free tumor DNA as a surrogate for determination of EGFR status. *J Thorac Oncol*. 2014;9:1345-1353.

26. AmoyDx. AmoyDx Super-BEAM EGFR detection kit. http://www.amoydiagnostics.com/China-FDA-Grants-Liquid-Biopsy-Approval-to-AmoyDx-Super-ARMS-EGFR-Mutation-Test_n114. Accessed February 1,2018.

27. Li Y, Xu H, Su S, et al. Clinical validation of a highly sensitive assay to detect EGFR mutations in plasma cell-free DNA from patients with advanced lung adenocarcinoma. *PLoS One*. 2017;12:e0183331.

28. Zhou C, Wang M, Cheng Y, et al. Detection of EGFR T790M in Asia-Pacific patients (pts) with EGFR mutation-positive advanced non-small cell lung cancer (NSCLC): circulating tumour (ctDNA) across 3 platforms. *Ann Oncol*. 2017;28(suppl_5):v460-v496.

29. Yang JC, Sequist LV, Geater SL, et al. Clinical activity of afatinib in patients with advanced non-small-cell lung cancer harbouring uncommon EGFR mutations: a combined post-hoc analysis of LUX-Lung 2, LUX-Lung 3, and LUX-Lung 6. *Lancet Oncol*. 2015;16:830-838.

30. National Institute for Health and Care Excellence. Plasma EGFR mutation tests for adults with locally advanced or metastatic non-small-cell lung cancer. Medtech Innovation Briefing. January 18, 2018. https://www.nice.org.uk/advice/mib137/resources/plasma-egfr-mutation-tests-for-adults-with-locally-advanced-or-metastatic-nonsmallcell-lung-cancer-pdf-2285963400737221. Accessed January 28, 2018.

31. Gray JE, Okamoto I, Sriurapong V, et al. Osimertinib vs SoC EGFR-TKI as first-line treatment in patients with EGFRm advanced NSCLC (FLAURA): plasma ctDNA analysis. *J Thorac Oncol*. 2017;12:S1754-S1755.

32. Lin CC, Shih JY, Yu CJ, et al. Outcomes in patients with non-small-cell lung cancer and acquired Thr790Met mutation treated with osimertinib: a genomic study. *Lancet Respir Med*. 2018;6:107-116.

33. Karachaliou N, Mayo-de las Casas C, Queralt C, et al. Association of EGFR L858R mutation in circulating free DNA with survival in the EURTAC trial. *JAMA Oncol*. 2015;1:149-157.

34. Rosell R, Carcereny E, Gervais R, et al. Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. *Lancet Oncol*. 2012;13:239-246.

35. Douillard JY, Oliner KS, Siena S, et al. Panitumumab-FOLFOX4 treatment and RAS mutations in colorectal cancer. *N Engl J Med*. 2013;369:1023-1034.

36. Van Cutsem E, Lenz HJ, Köhne CH, et al. Fluorouracil, leucovorin, and irinotecan plus cetuximab treatment and RAS mutations in colorectal cancer. *J Clin Oncol*. 2015;33:692-700.

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.17

37. Sysmex Inostics. OncoBEAMTM RAS CRC Kit Instructions for Use. https://www.sysmex-inostics.com/products/product-single-view/oncobeamTM-ras-crc-kit-ivd-3855.html. Accessed February 1, 2018.

38. Grasselli J, Elez E, Caratu G, et al. Concordance of blood- and tumor-based detection of RAS mutations to guide anti-EGFR therapy in metastatic colorectal cancer. *Ann Oncol*. 2017;28:1294-1301.

39. Vidal J, Muinelo L, Dalmases A, et al. Plasma ctDNA RAS mutation analysis for the diagnosis and treatment monitoring of metastatic colorectal cancer patients. *Ann Oncol*. 2017;28:1325-1332.

40. Schmiegel W, Scott RJ, Dooley S, et al. Blood-based detection of RAS mutations to guide anti-EGFR therapy in colorectal cancer patients: concordance of results from circulating tumor DNA and tissue-based RAS testing. *Mol Oncol*. 2017;11:208-219.

41. Normanno N, Abate RE, Lambiase M, et al. RAS testing of liquid biopsy correlates with the outcome of metastatic colorectal cancer patients treated with first-line FOLFIRI plus cetuximab in the CAPRI-GOIM trial. *Ann Oncol*. 2018;29:112-118.

42. Hahn S, Tannapfel A, Friess H, et al. Clinical performance comparison of RAS mutation detection in plasma and tissue samples from stage IV colorectal cancer patients from first-line anti-EGFR therapy selection. *Eur J Cancer*. 2015;51:S79.

43. Ma W. Significant improvement in detecting BRAF, KRAS, and EGFR mutations using next-generation sequencing as compared with FDA-cleared kits. *Mol Diagn Ther*. 2017;21:571-579.

44. Garcia-Foncillas J, Alba E, Aranda E, et al. Incorporating BEAMing technology as a liquid biopsy into clinical practice for the management of colorectal cancer patients: an expert taskforce review. *Ann Oncol*. 2017;28:2943-2949.

45. Jacobs B, Claes B, Bachet JB, et al. Evaluation of a fully automated extended RAS-BRAF test on prospectively collected plasma samples from patients with metastatic colorectal cancer. *J Clin Oncol*. 2017;35(15_suppl):e15127.

46. Bachet JB, Bouche O, Taïeb J, et al. RAS mutations concordance in circulating tumor DNA (ctDNA) and tissue in metastatic colorectal cancer (mCRC): RASANC, an AGEO prospective multicenter study. *J Clin Oncol*. 2017;35(15_suppl):11509.

47. Braig F, Voigtlaender M, Schieferdecker A, et al. Liquid biopsy monitoring uncovers acquired RAS-mediated resistance to cetuximab in a substantial proportion of patients with head and neck squamous cell carcinoma. *Oncotarget*. 2016;7:42988-42995.

48. Siravegna G, Mussolin B, Buscarino M, et al. Clonal evolution and resistance to EGFR blockade in the blood of colorectal cancer patients. *Nat Med*. 2015;21:795-801.

49. Van Emburgh BO, Arena S, Siravegna G, et al. Acquired RAS or EGFR mutations and duration of response to EGFR blockade in colorectal cancer. *Nat Commun*. 2016;7:13665.

50. Pietrantonio F, Vernieri C, Siravegna G, et al. Heterogeneity of acquired resistance to anti-EGFR monoclonal antibodies in patients with metastatic colorectal cancer. *Clin Cancer Res*. 2017;23:2414-2422.

51. Kim G, McKee AE, Ning YM, et al. FDA approval summary: vemurafenib for treatment of unresectable or metastatic melanoma with the BRAFV600E mutation. *Clin Cancer Res*. 2014;20:4994-5000.

52. Planchard D, Smit EF, Groen HJM, et al. Dabrafenib plus trametinib in patients with previously untreated BRAFV600E-mutant metastatic non-small-cell lung cancer: an open-label, phase 2 trial. *Lancet Oncol*. 2017;18:1307-1316.

53. Gonzalez-Cao M, Mayo-de-Las-Casas C, Molina-Vila MA, et al. BRAF mutation analysis in circulating free tumor DNA of melanoma patients treated with BRAF inhibitors. *Melanoma Res*. 2015;25:486-495.

54. El Messaoudi S, Mouliere F, Du Manoir S, et al. Circulating DNA as a strong multimarker prognostic tool for metastatic colorectal cancer patient management care. *Clin Cancer Res*. 2016;22:3067-3077.

55. Baselga J, Im SA, Iwata H, et al. Buparlisib plus fulvestrant versus placebo plus fulvestrant in postmenopausal, hormone receptor-positive, HER2-negative, advanced breast cancer (BELLE-2): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol*. 2017;18:904-916.

56. Di Leo A, Johnston S, Lee KS, et al. Buparlisib plus fulvestrant in postmenopausal women with hormone-receptor-positive, HER2-negative, advanced breast cancer progressing on or after mTOR inhibition (BELLE-3): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol*. 2018;19:87-100.

57. Chung TKH, Cheung TH, Yim SF, et al. Liquid biopsy of PIK3CA mutations in cervical cancer in Hong Kong Chinese women. *Gynecol Oncol*. 2017;146:334-339.

58. Christensen E, Birkenkamp-Demtröder K, Nordentoft I, et al. Liquid biopsy analysis of FGFR3 and PIK3CA hotspot mutations for disease surveillance in bladder cancer. *Eur Urol*. 2017;71:961-969.

59. Kidess-Sigal E, Liu HE, Triboulet MM, et al. Enumeration and targeted analysis of KRAS, BRAF and PIK3CA mutations in CTCs captured by a label-free platform: comparison to ctDNA and tissue in metastatic colorectal cancer. *Oncotarget*. 2016;7:85349-85364.

60. Slamon DJ, Leyland-Jones B, Shak S, et al. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *N Engl J Med*. 2001;344:783-792.

61. Bang YJ, Van Cutsem E, Feyereislova A, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet*. 2010;376:687-697.

62. Shoda K, Ichikawa D, Fujita Y, et al. Monitoring the HER2 copy number status in circulating tumor DNA by droplet digital PCR in patients with gastric cancer. *Gastric Cancer*. 2017;20:126-135.

63. Shoda K, Masuda K, Ichikawa D, et al. HER2 amplification detected in the circulating DNA of patients with gastric cancer: a retrospective pilot study. *Gastric Cancer*. 2015;18:698-710.

64. Kim ST, Banks KC, Lee SH, et al. Prospective feasibility study for using cell-free circulating tumor DNA–guided therapy in refractory metastatic solid cancers: an interim analysis. *Precision Oncol*. 2017;1:1-15.

65. Fribbens C, O'Leary B, Kilburn L, et al. Plasma ESR1 mutations and the treatment of estrogen receptor-positive advanced breast cancer. *J Clin Oncol*. 2016;34:2961-2968.

66. Fribbens C, Garcia Murillas I, Beaney M, et al. Tracking evolution of aromatase inhibitor resistance with circulating tumour DNA analysis in metastatic breast cancer. *Ann Oncol*. 2018;29:145-153.

67. Chung JH, Pavlick D, Hartmaier R, et al. Hybrid capture-based genomic profiling of circulating tumor DNA from patients with estrogen receptor-positive metastatic breast cancer. *Ann Oncol*. 2017;28:2866-2873.

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.18

68. Antonarakis ES, Lu C, Wang H, et al. AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. *N Engl J Med*. 2014;371:1028-1038.

69. Azad AA, Volik SV, Wyatt AW, et al. Androgen receptor gene aberrations in circulating cell-free NA: biomarkers of therapeutic resistance in castration-resistant prostate cancer. *Clin Cancer Res*. 2015;21:2315-2324.

70. Wyatt AW, Azad AA, Volik SV, et al. Genomic alterations in cell-free DNA and enzalutamide resistance in castration-resistant prostate cancer. *JAMA Oncol*. 2016;2:1598-1606.

71. Guardant Health. Guardant 360. http://www.guardanthealth.com. Accessed December 31, 2017.

72. Foundation Medicine. Genomic Testing. FoundationACT. https://www.foundationmedicine.com/genomic-testing/foundation-act. Accessed December 31, 2017.

73. Lanman RB, Mortimer SA, Zill OA, et al. Analytical and clinical validation of a digital sequencing panel for quantitative, highly accurate evaluation of cell-free circulating tumor DNA. *PLoS One*. 2015;10:e0140712.

74. Kim ST, Lee WS, Lanman RB, et al. Prospective blinded study of somatic mutation detection in cell-free DNA utilizing a targeted 54-gene next generation sequencing panel in metastatic solid tumor patients. *Oncotarget*. 2015;6:40360-40369.

75. Kato S, Kurasaki K, Ikeda S, Kurzrock R. Rare tumor clinic: the University of California San Diego Moores Cancer Center experience with a precision therapy approach. *Oncologist*. 2018;23:171-178.

76. Kato S, Krishnamurthy N, Banks KC, et al. Utility of genomic analysis in circulating tumor DNA from patients with carcinoma of unknown primary. *Cancer Res*. 2017;77:4238-4246.

77. Hahn AW, Gill DM, Maughan B, et al. Correlation of genomic alterations assessed by next-generation sequencing (NGS) of tumor tissue DNA and circulating tumor DNA (ctDNA) in metastatic renal cell carcinoma (mRCC): potential clinical implications. *Oncotarget*. 2017;8:33614-33620.

78. Riviere P, Fanta PT, Ikeda S, et al. The mutational landscape of gastrointestinal malignancies as reflected by circulating tumor DNA. *Mol Cancer Ther*. 2018;17:297-305.

79. Blakely CM, Watkins TBK, Wu W, et al. Evolution and clinical impact of co-occurring genetic alterations in advanced-stage EGFR-mutant lung cancers. *Nat Genet*. 2017;49:1693-1704.

80. Kuderer NM, Burton KA, Blau S, et al. Comparison of 2 commercially available next-generation sequencing platforms in oncology. *JAMA Oncol*. 2017;3:996-998.

81. Schwaederle M, Chattopadhyay R, Kato S, et al. Genomic alterations in circulating tumor DNA from diverse cancer patients identified by next-generation sequencing. *Cancer Res*. 2017;77:5419-5427.

82. Torga G, Pienta KJ. Patient-paired sample congruence between 2 commercial liquid biopsy tests. *JAMA Oncol*. Epub 2017 Dec 14.

83. Maia MC, Bergerot PG, Dizman N, et al. Association of circulating tumor DNA (ctDNA) detection in metastatic renal cell carcinoma (mRCC) with tumor burden. *Kidney Cancer*. 2017;1:65-70.

84. Ou SI, Lee TK, Young L, et al. Dual occurrence of ALK G1202R solvent front mutation and small cell lung cancer transformation as resistance mechanisms to second generation ALK inhibitors without prior exposure to crizotinib. Pitfall of solely relying on liquid re-biopsy? *Lung Cancer*. 2017;106:110-114.

85. Ou SI, Cui J, Schrock AB, et al. Emergence of novel and dominant acquired EGFR solvent-front mutations at Gly796 (G796S/R) together with C797S/R and L792F/H mutations in one EGFR (L858R/T790M) NSCLC patient who progressed on osimertinib. *Lung Cancer*. 2017;108:228-231.

86. Ou SI, Horn L, Cruz M, et al. Emergence of FGFR3-TACC3 fusions as a potential by-pass resistance mechanism to EGFR tyrosine kinase inhibitors in EGFR mutated NSCLC patients. *Lung Cancer*. 2017;111:61-64.

87. Klempner SJ, Mehta P, Schrock AB, et al. Cis-oriented solvent-front EGFR G796S mutation in tissue and ctDNA in a patient progressing on osimertinib: a case report and review of the literature. *Lung Cancer (Auckl)*. 2017;8:241-247.

88. ThermoFisher Scientific. Oncomine Lung cfDNA assay. https://www.thermofisher.com/order/catalog/product/A31149. Accessed December 31, 2017.

89. Archer Reveal. ctDNA 28 Liquid biopsy mutation profiling with Archer Reveal ctDNA 28. http://archerdx.com/reveal-ctdna. Accessed December 31, 2017.

90. Genomic Health. Oncotype SEQ Liquid Select. http://www.genomichealth.com/en-US/oncotype_iq_products/oncotype_SEQ. Accessed December 31, 2017.

91. MolecularMD. Platforms, Applications & Methods. http://molecularmd.com/platforms-applications-methods/overview/. Accessed December 31, 2017.

92. Pathway Genomics. Pathway Genomics launches first liquid biopsy test to detect cancer-associated mutations in high-risk patients. https://www.pathway.com/pathway-genomics-launches-first-liquid-biopsy-test-to-detect-cancer-associated-mutations-in-high-risk-patients/. Accessed December 31, 2017.

93. DiCarta. OptiSeq NGS Pan-cancer Panel (65 Genes). http://www.diacarta.com/services/ngs/. Accessed December 12, 2018.

94. Personal Genome Diagnostics. MetDETECT R. http://www.personalgenome.com/pharma-biotech-companies/liquid-biopsy/. Accessed December 31, 2017.

95. Chalmers ZR, Connelly CF, Fabrizio D, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med*. 2017;9:34.

96. Carbone DP, Reck M, Paz-Ares L, et al. First-line nivolumab in stage IV or recurrent non-small-cell lung cancer. *N Engl J Med*. 2017;376:2415-2426.

97. Koeppel F, Blanchard S, Jovelet C, et al. Whole exome sequencing for determination of tumor mutation load in liquid biopsy from advanced cancer patients. *PLoS One*. 2017;12:e0188174.

98. Fabrizio D, Malboeuf C, Lieber D, et al. Analytic validation of a next generation sequencing assay to identify tumor mutational burden from blood (bTMB) to support investigation of an anti-PD-L1 agent, atezolizumab, in a first line non-small cell lung cancer trial (BFAST). *Ann Oncol*. 2017;28(suppl_5):v22-v42.

99. Gandara DR, Kowanetz M, Mok TS, et al. Blood-based biomarkers for cancer immunotherapy: tumor mutational burden in blood (bTMB) is associated with improved atezolizumab (atezo) efficacy in 2L+ NSCLC (POPLAR and OAK). *Ann Oncol*. 2017; 28(suppl_5):mdx380.

100. Mok TS, Gadgeel S, Kim ES, et al. 1383TiP. Blood first line ready screening trial (B-F1RST) and blood first assay screening trial (BFAST) enable clinical development of novel blood-based biomarker assays for

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.19

tumor mutational burden (TMB) and somatic mutations in 1L advanced or metastatic NSCLC. *Ann Oncol*. 2017; 28(suppl_5):mdx380.084.

101. Tie J, Wang Y, Tomasetti C, et al. Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. *Sci Transl Med*. 2016;8: 346ra92.

102. Abbosh C, Birkbak NJ, Wilson GA, et al. Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. *Nature*. 2017;545:446-451.

103. Jamal-Hanjani M, Wilson GA, McGranahan N, et al. Tracking the evolution of non-small-cell lung cancer. *N Engl J Med*. 2017;376:2109-2121.

104. Murphy DJ, Blyth KG. Predicting lung cancer recurrence from circulating tumour DNA. Commentary on 'Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution'. *Cell Death Differ*. 2017;24:1473-1474.

105. Ng TL, Camidge DR. Lung cancer's real adjuvant EGFR targeted therapy questions. *Lancet Oncol*. 2018;19:15-17.

106. Chan AT, Lo YM, Zee B, et al. Plasma Epstein-Barr virus DNA and residual disease after radiotherapy for undifferentiated nasopharyngeal carcinoma. *J Natl Cancer Inst*. 2002;94:1614-1619.

107. Ocana A, Diez-Gonzalez L, Garcia-Olmo DC, et al. Circulating DNA and survival in solid tumors. *Cancer Epidemiol Biomarkers Prev*. 2016;25:399-406.

Downloaded from ascopubs.org by 99.88.42.8 on March 10, 2020 from 099.088.042.008
Copyright © 2020 American Society of Clinical Oncology. All rights reserved.

TEMPUS EX1002.20

BIOINFORMATICS ORIGINAL PAPER Vol. 27 no. 19 2011, pages 2648–2654
doi:10.1093/bioinformatics/btr462

Sequence analysis Advance Access publication August 9, 2011

Exome sequencing-based copy-number variation and loss of heterozygosity detection: ExomeCNV

Jarupon Fah Sathirapongsasuti[1,2,3,*], Hane Lee[3,4], Basil A. J. Horst[4,5,6], Georg Brunner[7], Alistair J. Cochran[4], Scott Binder[4], John Quackenbush[1,2] and Stanley F. Nelson[3,4,*]

[1]Department of Biostatistics, Harvard School of Public Health, [2]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Boston, MA 02115, [3]Department of Human Genetics and [4]Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, [5]Department of Dermatology, [6]Department of Pathology and Cell Biology, Columbia University Medical Center, New York, NY 10032, USA and [7]Department of Cancer Research, Skin Cancer Center Hornheide, Münster, Germany

Associate Editor: Alfonso Valencia

ABSTRACT

Motivation: The ability to detect copy-number variation (CNV) and loss of heterozygosity (LOH) from exome sequencing data extends the utility of this powerful approach that has mainly been used for point or small insertion/deletion detection.

Results: We present ExomeCNV, a statistical method to detect CNV and LOH using depth-of-coverage and B-allele frequencies, from mapped short sequence reads, and we assess both the method's power and the effects of confounding variables. We apply our method to a cancer exome resequencing dataset. As expected, accuracy and resolution are dependent on depth-of-coverage and capture probe design.

Availability: CRAN package 'ExomeCNV'.

Contact: fsathira@fas.harvard.edu; snelson@ucla.edu

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on January 22, 2011; revised on July 8, 2011; accepted on August 2, 2011

1 INTRODUCTION

The development of next-generation sequencing has enabled routine large-scale resequencing projects, permitting us to perform increasingly more comprehensive DNA variant analysis. However, the cost and analytical complexity of sequencing still limit the number of whole genomes that can be sequenced in any single project (Teer and Mullikin, 2010). In fact, the analysis of complete human genome sequence often interprets DNA alterations in protein coding regions primarily. This is in practice a reasonable strategy since ~85% of the disease-causing mutations are found in the coding regions or canonical splice sites (Choi *et al.*, 2009). Thus, whole-exome sequencing presents an effective alternative to whole-genome sequencing and provides an unbiased, cost-effective and time-efficient tool for the study of the genetic basis for disease. Following the first successful application of whole-exome sequencing in re-discovering the cause of a dominantly inherited rare Mendelian disorder Freeman–Sheldon syndrome (Ng *et al.*, 2009), a number of studies have reported similar successes (Chou *et al.*, 2010; Hoischen *et al.*, 2010; Johnston *et al.*, 2010; Raca *et al.*, 2010; Rehman *et al.*, 2010; Volpi *et al.*, 2010). Although the cost of both genome and exome sequencing continues to fall at a rapid pace, whole-exome sequencing has a number of advantages, including a lower cost, more straightforward data analysis and interpretation and significantly greater depth of coverage with a corresponding overall improvement in data quality. Exome sequencing is rapidly becoming a fundamental tool for genetic and functional genomic research laboratories and a diagnostic tool in clinics. At present the main applications of targeted exonic sequencing is for the determination of single nucleotide variants (SNVs) or small indel variants but not structural variation.

Structural variation, especially copy-number variation (CNV) and loss of heterozygosity (LOH), is an important class of genetic variability in Mendelian, common inherited diseases and cancer (Choi *et al.*, 2007; Deng *et al.*, 2010; Fong *et al.*, 1989; Jankowska *et al.*, 2009; Sha *et al.*, 2009; Shuin *et al.*, 1994; Stankiewicz and Lupski, 2010; Wain *et al.*, 2009). As is true of SNVs, there are population-specific, common CNVs and rare, disease-causing CNVs (Kato *et al.*, 2010; Sudmant *et al.*, 2010). Many large-scale projects (Conrad *et al.*, 2010a,b; McCarroll, 2010) and technological platforms (Pinkel *et al.*, 1998; Urban *et al.*, 2006) have been devised to estimate the prevalence and impact of CNV. Array Comparative Genomic Hybridization (CGH) (Pinkel *et al.*, 1998) and SNP genotyping arrays have been widely used as standard methods to detect CNV and LOH. However, with the rapid increase in genomic and exomic sequence, there is growing interest in the use of these data to detect CNVs.

While methods have been developed for CNV estimation in whole-genome sequencing (Chiang *et al.*, 2009; Yoon *et al.*, 2009), these methods make key assumptions that fail to hold in the exome sequencing setting. For example, Yoon *et al.* (2009) assumes random, unbiased distribution of sequence reads, such that read depth can be modeled as a normal distribution across the genome, and deviation from the background indicates the presence of CNV. This random read distribution assumption breaks down in the context of exome capture as the probes have variable specificity and efficiency for the targeted exonic regions. The discrete nature of exome sequences also presents problems to

*To whom correspondence should be addressed.

 © The Author 2011. Published by Oxford University Press. All rights reserved. For Permissions, please email: journals.permissions@oup.com

Downloaded from https://academic.oup.com/bioinformatics/article-abstract/27/19/2648/231564
by guest
on 27 March 2018

TEMPUS EX1002.21

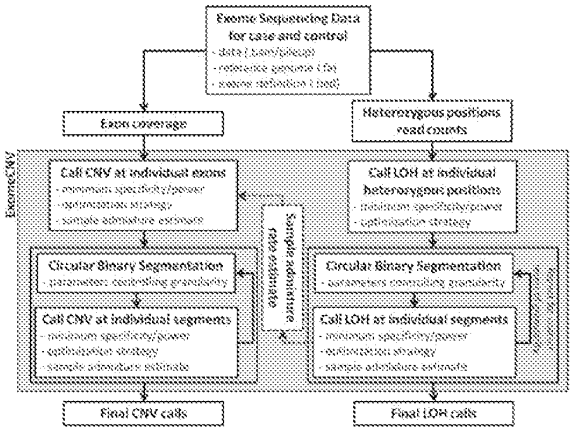


Fig. 1. Overview of ExomeCNV analysis workflows. Two workflows are present: CNV detection and LOH detection. Each involves similar steps of exon/position/segment-wise CNV/LOH calling, Circular Binary Segmentation, and interval merging. User inputs and parameters are listed at each step.

existing methods. Many whole-genome CNV detection tools use segmentation algorithms that assume continuity of search space and do not function properly when given discontinuous and variable length exome sequencing data. SegSeq (Chiang *et al.*, 2009), for example, merges windows of a fixed length based on a log-ratio difference statistic. Lastly, because exons are generally smaller than insert sizes for paired-end sequencing (200–500 bp), paired-end based CNV detection methods are not generally applicable to exome data.

Here we present ExomeCNV, which uses depth-of-coverage and B-allele frequencies from mapped short sequence reads to estimate CNV (Fig. 1, left side) and LOH (Fig. 1, right side). We describe an assessment of its validity, sensitivity, specificity and limitations through an analysis of a melanoma tumor and a matched normal sample. Important model assumption and the effect of important confounding factor such as sample admixture rate are also considered.

2 METHODS

2.1 Correlation of depth-of-coverage

To plot the correlation of depth-of-coverage (Fig. 2), we used the internal exome data that were captured by the same Agilent SureSelect Human All Exon Kit and sequenced on the Illumina GAIIx. Samples 1 through 4 were generated by two lanes of 76 bp single-end sequencing, Sample 5 was generated by three lanes of 76 bp single-end sequencing and Sample 6 was generated by one lane of 76 + 76 bp paired-end sequencing. For each sample, the average depth-of-coverage per exon was normalized by dividing the average coverage by the overall exome average coverage and then, the normalized depth-of-coverage were compared between 15 pairs of samples.

2.2 The CNV detection algorithm

2.2.1 Power analysis of CNV Detection Consider an exon of length L, let X and Y denote the numbers of reads, each of length w, mapped within the exon in question in case (e.g. tumor) and control (e.g. matched normal), respectively. The depth-of-coverage is then Xw/L and Yw/L for case and control, respectively. Although we discuss our method in terms of depth-of-coverage, our method is developed in terms of the count statistics X and Y. Let N_X and N_Y be the total numbers of aligned reads in case and control, respectively. Define the read count ratio:

$$R=\frac{X/N_X}{Y/N_Y}. \tag{1}$$

We divide the raw counts X and Y by the total number of reads N_X and N_Y to mitigate the effect of overall increase in local counts due to the increase in total depth-of-coverage. Finally, we adjust the ratio so that the exome-wide median is 1. Without lost of generality, we assume $N_X = N_Y$ and reduce $R = X/Y$. Because X and Y follow Poisson distributions with parameters λ_X and λ_Y, respectively, with sufficient depth-of-coverage the Poisson distributions converge to normals with equal means and variances: $N(\lambda_X, \lambda_X)$ and $N(\lambda_Y, \lambda_Y)$. Under the null hypothesis of no CNV, $\lambda_X = \lambda_Y$, and under the alternative hypothesis, $\lambda_X = \rho\lambda_Y = \rho\lambda$. ρ indicates the copy-number ratio; for example, $\rho = 0.5$ for deletion and $\rho = 1.5$ for duplication. By Geary–Hinkley transformation (Geary, 1930,1944; Hinkley, 1969), let

$$t(\rho)=\frac{\mu_Y R-\mu_X}{\sqrt{\sigma_Y^2 R^2+\sigma_X^2}}=\frac{\lambda_Y R-\lambda_X}{\sqrt{\lambda_Y R^2+\lambda_X}}=\frac{\lambda R-\rho\lambda}{\sqrt{\lambda R^2+\rho\lambda}}=\frac{(R-\rho)\sqrt{\lambda}}{\sqrt{R^2+\rho}}, \tag{2}$$

and $t(\rho)$ follows the standard normal distribution. Thus, the specificity and sensitivity are $1-\alpha$ and $1-\beta$ where

$$\alpha=\begin{cases}\phi(t(1)) & \text{if } \rho<1\\ 1-\phi(t(1)) & \text{if } \rho\geq 1\end{cases} \tag{3a}$$

$$\beta=\begin{cases}1-\phi(t(\rho)) & \text{if } \rho<1,\\ \phi(t(\rho)) & \text{if } \rho\geq 1\end{cases} \tag{3b}$$

The formulas above describe the achievable specificity and sensitivity of a given cutoff ratio R. Considering values of R from zero to infinity, we can plot the receiver operating characteristic (ROC) curve (Supplementary Materials). In practice, we may wish to identify a cutoff $r(\rho)$ which yields the desired minimum specificity and/or sensitivity for testing a particular copy-number ratio ρ at an exon with certain depth-of-coverage and length. This can be achieved by solving above equations, and an appropriate cutoff $r(\rho)$ can be chosen from a set of solutions to maximize user-selected quantity metrics such as specificity, sensitivity or area under curve.

In the presence of sample admixture, the 'true' copy-number ratio will tend to 1. In particular, if a fraction c of the tumor sample has a normal copy-number (either by contamination of normal tissue or heterogeneity within the tumor), the copy-number ratio of this admixed sample will be $\rho' = c + \rho(1-c)$. Thus, in heterogeneous samples, the only change to the method described above is the replacement of ρ by ρ'. The admixture rate c can be estimated from data by back-calculating c from empirical ρ' in LOH regions (see Supplementary Materials).

2.3 Segmentation and sequential merging

We used the circular binary segmentation (CBS) algorithm (Olshen *et al.*, 2004), as implemented in the R package DNAcopy (Venkatraman and Olshen, 2007), to subdivide the genome (exome). For each segment we combined the coverage by direct sum, and used mean coverage log ratio as the segment's log ratio $\log(R)$. Log ratio is used here to satisfy the input requirement of CBS algorithm. CNV call proceeds on each segment in the same manner as described above.

In order to achieve the most sensitive segmentation, we chose to start CBS with parameters that produce a large number of small segments, call CNV on the segments, and repeat the process with a smaller number of larger segments. We then merge the CNV segments sequentially, from finest segmentation to coarsest. By nature of CBS, finer segments are contained within coarser segments, and in merging step, we need to resolve conflicting CNV calls between finer segments within a larger segment. If a finer (smaller) segment has sufficient coverage to call a CNV event, the call persists.

Downloaded from https://academic.oup.com/bioinformatics/article-abstract/27/19/2648/231564
by guest
on 27 March 2018

TEMPUS EX1002.22

However, if it does not have sufficient coverage to reject the null hypothesis (i.e. not being called, or being called as copy-number neutral), a positive CNV call in a larger segment overrides the negative calls. An illustration of this sequential merging algorithm is given in the Supplementary Materials.

2.4 LOH detection

First, we consider all polymorphic positions in the exome of the control sample, and for each of the positions, the B-allele count B is the number of reads with non-reference- or B-allele at that position. For a polymorphic position i, let N_i be the total number of reads mapped to that position (i.e. depth-of-coverage); thus the B-allele count B_i follows a binomial distribution Binomial (p_i, N_i). A binomial that rejects the null hypothesis: $p_i = 0.5$ can be used to detect LOH at each polymorphic position.

Segmentation is done using CBS algorithm based on the absolute difference in B-allele frequencies (BAFs) $|\mathrm{BAF}_{i,\mathrm{case}} - \mathrm{BAF}_{i,\mathrm{control}}|$, where $\mathrm{BAF}_i = B_i/N_i$. Within each segment, based on the realization that B-allele frequencies deviate from the null value of 0.5 under LOH, an F-test for equality of variance is used to detect significance increase in variance of BAF_{case} from that of BAF_{control} (other statistics were also considered, see Supplementary Materials). Finally, the LOH calls are merged sequentially as described above.

2.5 Exome sequencing and data analysis

2.5.1 Exome sequencing High molecular weight whole genomic DNA from matched skin and tumor pair samples of a melanoma patient were sequenced on the Illumina Genome Analyzer (GAIIx). The UCLA Institutional Review Board (IRB) approved the collections of the DNA samples. The libraries were generated following the Agilent SureSelect Human All Exon Kit version 1.0.1 protocol, and the Illumina Genome Analyzer (GAIIx) flowcell was prepared according to the manufacturer's protocol. We performed three and four lanes of single-end sequencing for each of the skin and tumor samples, respectively, within the UCLA Center of High-throughput Biology (CHTB). The base-calling was performed by the real time analysis (RTA) software (version 1.6) provided by Illumina.

2.5.2 Sequence data analysis Novoalign from Novocraft Short Read Alignment Package (http://www.novocraft.com/index.html) was used to align each lane's QSEQ file to the reference genome. Human Genome reference sequence (hg18, March 2006, build 36.1), downloaded from the UCSC genome database located at http://genome.ucsc.edu and mirrored locally, was indexed using novoindex program (-k 14 -s 3). The output format was set to SAM and default settings were used for all options. Using SAMtools (http://samtools.sourceforge.net/), the SAM files of each lane were converted to BAM files, sorted and merged for each sample and potential PCR duplicates were removed using Picard (http://picard.sourceforge.net/) (Li *et al.*, 2009). To retrieve the depth of coverage information of each base, we generated a PILEUP file for each sample using SAMtools and calculated the average coverage per capture interval using a custom script. Here, we used processed BAM files that were used to call the SNVs while reducing the likelihood of using spuriously mismapped reads to call the variants: the last 5 bases were trimmed and only the reads lacking indels were retained (Clark *et al.*, 2009). The detailed description of the mutational landscape of this tumor sample is in preparation.

2.6 Genome-wide SNP genotyping

Both the skin and the tumor samples were submitted to the Southern California Genotyping Consortium (SCGC) at UCLA for genotyping on the Illumina Omni-1 Quad BeadChip, which consists of 1 140 419 SNPs (1 016 423 genotyping probes and 123 996 CNV probes) distributed across the genome. The Illumina GenomeStudio V2010.1 Genotyping Module version 1.6.3 was used to calculate the BAF values and the log R ratio (LRR) for each probe and the copy number aberration (CNA) and LOH of the autosomes were inferred from these values using the genoCN R package (Sun *et al.*, 2009). The genoCNA function was used with the default parameters.

2.7 Copy-number analysis using ERDS

The same PILEUP file we used to generate the average coverage per capture interval was used to run Estimation by Read Depth with SNVs (ERDS) to demonstrate the need to control for the variability of capture of exons observed in exome sequencing (M.Zhu *et al.*, manuscript in preparation). The SNV file that is required by ERDS was generated by using SAMtools varFilter tool default parameters and SVA (Sequence Variant Analyzer) snp_filter.pl script. The result file generated by ERDS was summarized using the SVA software (http://people.genome.duke.edu/~dg48/sva/index.php).

2.8 Comparison between CNV calling methods

In assessing performance of ExomeCNV, we used all mapped exons as the sample space. CNV calls on other platforms were mapped to the exons and compared to calls by ExomeCNV. Thus specificity is the proportion of copy-neutral exons correctly identified by ExomeCNV, while sensitivity is the proportion of amplified (or deleted) exons correctly identified. Similarly, LOH performance is assessed using all polymorphic positions as the sample space.

3 RESULTS

3.1 ExomeCNV for CNV and LOH detection

ExomeCNV uses a normalized depth-of-coverage ratio approach to identify CNV and LOH from exome sequencing information of paired case/control samples (for example, paired tumor/normal) in a way that optimizes sensitivity and specificity. We begin by assuming that although there are potentially exon-specific biases due to laboratory capture methods and sequence-specific biases, these are independent of sample and so are nearly uniform for a particular exon across samples. As a result, simply assessing the ratio of depth-of-coverage of each exon reduces such bias (see Supplementary Materials).

3.1.1 Correlation of depth-of-coverage across exome sequencing samples To establish validity of this fundamental assumption, we compared depth-of-coverage of exons across five independent samples from five different subjects (Samples 1–5 in Fig 2). All samples were captured using the same probe set (Agilent SureSelect Human All Exon G3362) and sequenced at mean base coverages of 36–39× as a result of two (Samples 1–4) or three (Sample 5) lanes of GAIIx single-end sequencing per sample (see Section 2). As shown in Figure 2, a high correlation was observed among the five samples (Pearson correlation 0.908–0.975, mean = 0.947, SD = 0.027), arguing for the validity of our assumption.

The same level of consistency was not observed when single-end data were compared with paired-end data (Sample 6; Pearson correlation 0.855–0.877, mean = 0.871, SD = 0.009) due to the lack of independence between pairs of reads in paired-end data. Thus, care must be taken to ensure consistency of library preparation and sequencing method between samples used in analysis. Here, all of our analyses used exome sequencing data from a melanoma (Sample 5) and a matched normal skin, both processed and sequenced in the same manner (see Section 2).

3.1.2 Analytic power calculation of exonic CNV and LOH detection For each exon, the number of sequencing reads aligning

Downloaded from https://academic.oup.com/bioinformatics/article-abstract/27/19/2648/231564
by guest
on 27 March 2018

TEMPUS EX1002.23

within it appears to follow a Poisson with mean directly proportional to the size of the exon and the copy-number (see Supplementary Materials), but assuming that we have sufficiently deep coverage, we can approximate this by a normal distribution with mean equal to variance. We apply the Geary–Hinkley transformation (Geary, 1930,1944; Hinkley, 1969), which converts a ratio of two normally distributed variables to a standard normal distribution, and a CNV is identified by a significant deviation of the transformed ratio from the null, standard normal distribution (see Section 2). Allowing only one false positive per genome, we analytically determine the statistical power of this CNV detection approach for different depth-of-coverage, and the results are shown in Figure 3a–b. For detecting deletions, 95% power is achieved for segments of size 500 bp or more (Fig. 3a), while detection of a single copy duplication is achieved with 95% power for segments of size 1,000 bp or more (Fig. 3b) with a mean segmental base coverage of 35×. We note that the power of the method improves substantially with higher depth-of-coverage, and an individual exons deletion/duplication status would be more powerfully observed by including additional flanking intronic sequence in the capture probe design. Genomic DNA admixture, as expected, diminishes power, but even with 35× coverage of a given exon, a length of greater than 1000 bp is observed over 95% of the time. Exons or segments captured with 500 bp at target sequence are observed at 95% power only with greater than 55× base coverage. A more thorough consideration of specificity and ROC curves are produced in the Supplementary Materials.

To estimate LOH, we focused on the non-reference-allele or BAF of polymorphic positions in the sequenced regions. The observed B-allele count at a polymorphic position can be modeled by a binomial distribution with depth-of-coverage as sample size and the probability of observing a B-allele proportional to the B-allele copy-number, which is equivalent to the LOH state. Because the

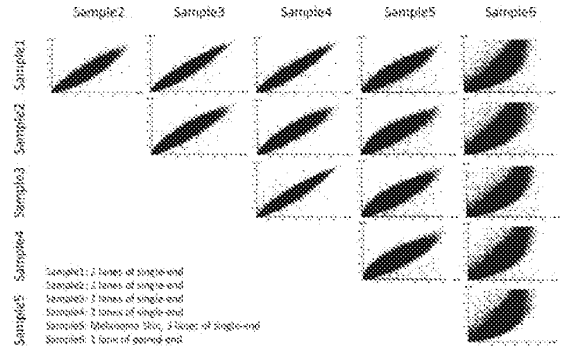


Fig. 2. Correlation of depth-of-coverage across exome sequencing samples. To demonstrate the consistency of capture and sequencing efficiency of individual exons represented by the depth-of-coverage per exon, the normalized individual exon coverage in all pairs of six-independent exomes were plotted. All 6 samples were captured using the Agilent SureSelect Human All Exon G3362. Samples 1–5 had mean base coverages of 36∼39× as a result of 2 (Samples 1–4) or 3 (Sample 5) lanes of GAIIx single-end sequencing per sample. Sample 6 had mean base coverage of 60× as a result of 1 lane of GAIIx paired-end sequencing and demonstrates substantially different biases in individual exonic depth-of-coverage.

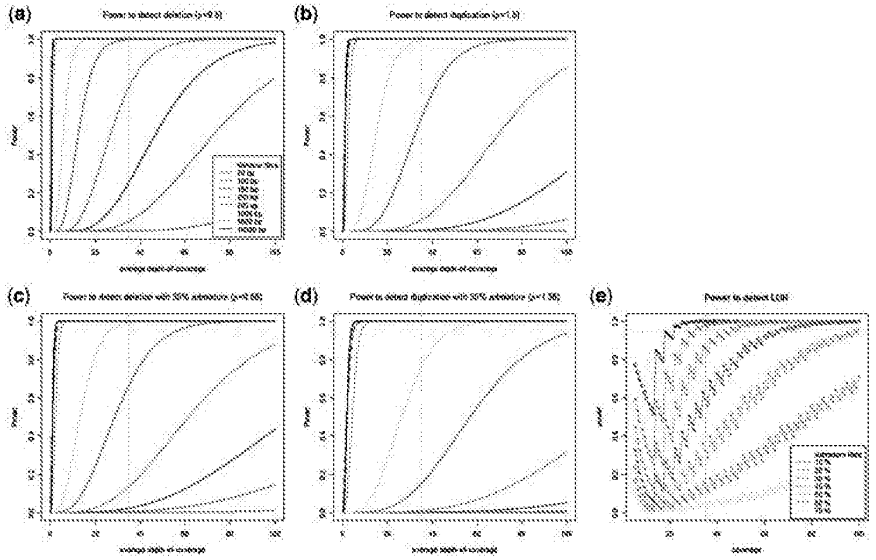


Fig. 3. Examples of the power of ExomeCNV to detect segmental duplication, deletion and LOH based on an analytical calculation. Power is plotted relative to mean depth-of-coverage in the genomic segment, setting false positive to 1 per genome based on an analytical model of genome-wide power of detection at different window sizes (inset, **a**–**d**). Windows are the total length of a given sequence at a given exon or the sum of length of exons adjacent to each other in the genome. The effect of admixture (rate of 30%) on the power to detect deletions and single copy duplications are shown in (c) and (d), respectively. (**e**) plots the power of LOH detection versus depth-of-coverage of individual polymorphic position (single base pair) with variable admixture rates (inset). The periodicity of the power curve is due to discrete nature of the binomial test. The 35× depth of coverage is chosen because it is a typical minimal average depth of coverage for exome sequencing and is thus a conservative view of power within typical exome sequencing datasets.

Downloaded from https://academic.oup.com/bioinformatics/article-abstract/27/19/2648/231564
by guest
on 27 March 2018

TEMPUS EX1002.24

expected value of BAF at a normal (non-LOH) polymorphic position is 0.5, a significant deviation of BAF from 0.5 identifies LOH. With sufficient depth-of-coverage, LOH can be detected at a single polymorphic position.

3.1.3 The effect of sample admixture The specificity and sensitivity of this CNV detection method depends not only on the depth-of-coverage but also the rate of admixture, whereby non-mutated genomes contaminate mutated genomes in the sampled tissue/cells. In the absence of admixture, the average depth-of-coverage ratio is 0.5 for deletion, 1.5 for one-copy duplication and the BAF at an LOH site is either 0 or 1; however, in cancer biopsy sequencing, this is rarely observed in practice due to admixture with normal or non-mutated tumor cells. Thus, we assess the effect of admixture ranging from 10% to 70%, which is frequently observed in tumor samples (Fig. 3e). With admixture, the depth-of-coverage ratio and the BAF will tend to the null values of 1 and 0.5, respectively, making the CNV and LOH detection harder. Figure 3c and d show a reduction of power in detecting deletion (Fig. 3c) and one-copy duplication (Fig. 3d) as a result of 30% admixture. There is an approximate two-fold increase in the size of the exonic sequence detectable in the presence of 30% non-mutated genomic DNA. Capping the false positive rate at 0.001 and assuming 35× mean depth-of-coverage, a power curve (Fig. 3e) shows 0.95 sensitivity of detecting LOH at a single polymorphic position with admixture up to 30%.

3.1.4 Using circular binary segmentation to merge exonic CNV/LOH Because CNV and LOH can, and usually do, span multiple exons, we extended our method above to call CNV/LOH on larger segments derived from summing data of sequentially spaced exons in the human genome. We apply circular binary segmentation (CBS) (Olshen *et al.*, 2004; Venkatraman and Olshen, 2007) to subdivide the genome and then combine depth-of-coverage of exons and BAF of polymorphic positions within each segment, composed of arbitrary number of individual exons, to search for larger CNV and LOH. In the case of CNV, since reads are independent of each other, the sum of depth-of-coverage of all exons in a segment constitutes the segment's depth-of-coverage, and the CNV test can be performed as described above. In the case of LOH, since B-alleles are not always on the same chromosome, BAF cannot be combined by direct summation. Instead, since BAF deviates from the null value 0.5 under LOH, a significance increase in variance of BAF from control (F-test for equality of variances) indicates LOH (several other statistics were also considered, see Supplementary Materials). Finally, we repeated the process of CBS and CNV/LOH-calling, ranging granularity of segmentation from finest to coarsest, and merged the CNV/LOH calls by prioritizing positive calls of finer segments over coarser ones (see Section 2 for details). In the case of our melanoma sample, we performed CBS/sequential merging at five levels of granularity and observed 165 130 merging events in the first iteration followed by 121, 79, 105 and 66 in the subsequent iterations for a total of 165 501 merging events.

3.2 Validation

To test the performance of ExomeCNV we analyzed exome sequencing data from a melanoma and a matched normal skin (Supplementary Materials); the average depth-of-coverage of the data is 42.8× for the tumor and 37.5× for the normal sample, which are sufficient to achieve at least 90% sensitivity and specificity based on the power calculation above.

3.2.1 Validating false positive and false negative rates We first estimated the false positive rate of the algorithm by calling CNVs on two sequencing lanes of the same normal tissue library, treating one as case and the other as control; any CNV call from this would be false positive. Our method correctly called most exons as non-CNV. In particular, setting P-value thresholds to ensure minimum specificity of 0.9, 0.99 and 0.999, we observed specificity of 0.916, 0.995 and 1.0, respectively (Supplementary Materials). Furthermore, we tested the sensitivity of ExomeCNV by analyzing copy-number of sex chromosomes in a pair of male and female exome data that were available internally (see Section 2). Using the male exome as control, ExomeCNV correctly identified female chromosome X as being 'duplicated' and chromosome Y as being 'deleted' (Supplementary Materials) with no false negatives.

3.2.2 Comparison with SNP genotyping array We then used ExomeCNV to predict CNV and LOH in the melanoma samples and compared our results to those obtained from Illumina Omni-1 Quad Beadchip genotyping array assessment of the same samples (Fig. 4 and Supplementary Materials). The sizes of the CNV segments from ExomeCNV range from single exon (120 bp) to whole chromosome (chr 10 and 18) (size distribution of CNV calls is presented in the Supplementary Materials). Treating calls from the genotyping array experiment as a standard, ExomeCNV had 97% specificity and 86% sensitivity for detecting deletions, 92% specificity and 88% sensitivity for detecting amplifications, and 88% specificity and 68% sensitivity for detecting LOH even though there is substantial variability across the genome. Higher depth-of-coverage from the sequence data for each exome would likely further improve concordance. We note that this is a dramatic improvement of ExomeCNV relative to the inappropriate application of ERDS (M.Zhu *et al.*, manuscript in preparation) CNV caller which, when applied to these data, achieves only 16% sensitivity and 83% specificity for deletion and 50% sensitivity and 56% specificity for amplification (see Circo plot showing results from the three methods in the Supplementary Materials). This is due to the fact that there is substantial variability of the efficiency of capture of exons in exome sequencing not accounted for in ERDS. For CNV segments called by ExomeCNV but not by the genotyping arrays, we found that most lie within regions in which there is a low density of genotyping markers; thus the false positive rates (and the associated specificity) for ExomeCNV here may in fact be lower.

4 DISCUSSION AND CONCLUSION

The resolution of CNV detection with ExomeCNV is limited largely by the probe design. The CNV segments identified by our method range from 120 bp (single exons with higher than average coverage) to 240 Mb in size (whole chromosomes); however, the true breakpoint can be anywhere in the space between the terminal exon called within a CNV region and the adjacent exon in a non-CNV region. Hence, although a given CNV event can be detected at a single exon in some instances, the absolute resolution of our method is in fact limited to the inter-exon distance around an exon, which can be as small as 125 bp or as large as 22.8 Mb with the

Downloaded from https://academic.oup.com/bioinformatics/article-abstract/27/19/2648/231564
by guest
on 27 March 2018

TEMPUS EX1002.25

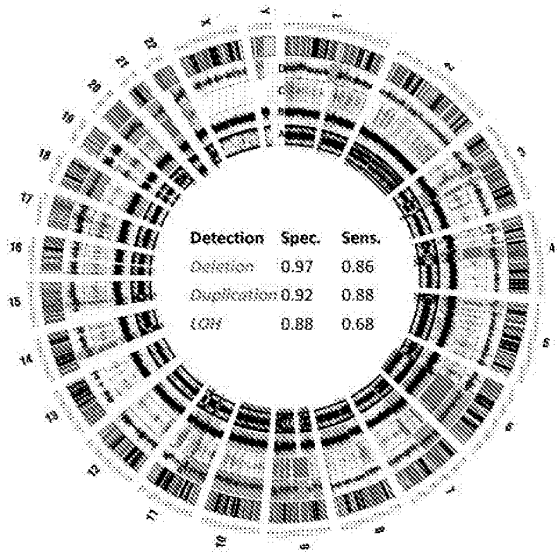


Fig. 4. Analysis of melanoma and paired normal samples. Interpretation of deletion, duplication and LOH from exonic sequence data using ExomeCNV and plotted with Circos. The most outer ring shows the chromosome ideograms in a pter–qter orientation, clockwise with the centromeres in red. From inside to outside, each data track represents (A) B-allele-frequency (BAF) from Omni-1 genotyping array with the region of LOH highlighted in blue underneath the track; (B) Log R Ratio (LRR) from genotyping array with the region of gain highlighted in red and the region of loss highlighted in green; (C) BAF from ExomeCNV output from ~40× depth-of-coverage exome sequencing with the region of LOH highlighted in blue; (D) log ratio of tumor and normal depth-of-coverage with the segment mean in red line, the region of gain highlighted in red and the region of loss highlighted in green. The LOH and CNV for the chromosomes X and Y were not called for the genotyping data as genoCN (the algorithm used to call CNV from Omni-1) is not designed to analyze chromosomes X and Y. The table in the middle summarizes best achievable specificity and sensitivity of ExomeCNV in detecting CNV and LOH relative to CNV/LOH calls from Omni-1 array assessment.

median of 5 kb (statistics based on SureSelect Human All Exon Kit G3362).

Although ExomeCNV relies on the availability of matched control samples, we can also derive a matched control sample from a pool of other samples, which then serves as an effective control. This is useful for the identification of germline inherited or de novo CNVs in an individual. Because the expected copy number in the reference population is constant (usually two), by the law of large numbers, averaging depth-of-coverage from sufficiently many samples yields a good control set, assuming that they are all captured using the same probe set and capture method and sequenced in the same manner. This may limit the application of ExomeCNV to data generated at a given site with a given protocol. Calling CNVs using this pooled sample as background will generate CNV calls that are present in the case sample but not the control population. Also, by the central limit theorem, pooling independent samples helps reduce variance in depth-of-coverage and increases precision of our method. We have pooled as few as eight samples and have observed that this is indeed the case (Supplementary Materials). However, it is important to note that using pooled sample as control imposes a strong assumption that the samples do not share common CNV regions and that the population has an average genomic copy number of two. Other potential challenges of using the pooled sample as control are discussed in the Supplementary Materials.

Because ExomeCNV depends on an estimate of the admixture rate c, misspecification of c would affect its performance. We performed sensitivity analysis and found that misestimating c would have a strong effect on sensitivity and specificity of CNV detection. Fortunately, LOH detection provides some data to directly estimate c, as LOH detection does not depend on a prior knowledge of c (Supplementary Materials). For the melanoma sample, our estimate of 30% admixture rate matches that from genotyping arrays, confirming the validity of this approach. However, there are advantages to slightly overestimating c as it makes the method more conservative and reduces false positives.

As we have shown, CNV and LOH detection is readily possible from exome sequencing data, extending the utility of this powerful approach. The fundamental basis that makes this approach possible is the consistency of depth-of-coverage of each exon (and BAF by extension) across multiple samples for each individual exon, as demonstrated in five samples performed in our laboratory (Supplementary Materials, Fig. 2). This consistency permits reliable parametric modeling of the shift in depth-of-coverage and BAF distributions, hence accurate identification of CNV and LOH. However, we do not observe the same level of consistency when comparing depth-of-coverage across different library types. For instance, a sixth sample was performed using a paired-end approach that results in very different coverage of each exon (Fig. 2), and as a result, ExomeCNV does not perform well when the control sample library is of one type and the case is of another, or when the case and control have significantly different coverage levels. Resolving these issues is a work-in-progress.

From the analytical power calculations, assuming 35× coverage (which is the lower end of a reasonable amount of sequence for variant calling and easy to generate with a variety of technologies), CNV detection has a limit of about 500 bp (in transcript coordinates), which is typically equivalent to 2–3 exons and spans about 10 kb of genomic space on average. Increased depth-of-coverage, which is likely to become the norm as sequencing costs decrease, reduces the interval size that is reliably detectable and should push the method to single exonic deletion resolution. Currently, CNV and LOH information should be detectable in whole-exome sequencing data at a resolution that is almost equivalent to what one can obtain from a dense SNP genotyping array.

ExomeCNV is available as a CRAN package 'ExomeCNV'.

ACKNOWLEDGEMENTS

We would like to thank Kevin Squire and the members of Nelson Lab for thoughtful discussion and suggestion, Markus Schröder for help with R package and Bret Harry, Zugen Chen and Traci Toy for technical support.

Funding: National Heart, Lung, Blood Institute of the US National Institutes of Health grant (1RC2 HL101715, to J.Q.); National Institute for Mental Health (R01 MH071852); National Institute of Arthritis, Musculoskeletal and Skin Disorders (P30 AR057230); National Cancer Institute (P30 CA16042).

Conflict of Interest: none declared.

Downloaded from https://academic.oup.com/bioinformatics/article-abstract/27/19/2648/231564
by guest
on 27 March 2018

TEMPUS EX1002.26

REFERENCES

Chiang,D.Y. *et al.* (2009) High-resolution mapping of copy-number alterations with massively parallel sequencing. *Nat. Methods*, **6**, 99–103.

Choi,C.H. *et al.* (2007) Hypermethylation and loss of heterozygosity of tumor suppressor genes on chromosome 3p in cervical cancer. *Cancer Lett.*, **255**, 26–33.

Choi,M. *et al.* (2009) Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc. Natl Acad. Sci. USA*, **106**, 19096–19101.

Chou,L.S. *et al.* (2010) DNA sequence capture and enrichment by microarray followed by next-generation sequencing for targeted resequencing: neurofibromatosis type 1 gene as a model. *Clin. Chem.*, **56**, 62–72.

Clark,M.J. *et al.* (2009) U87MG Decoded: The Genomic Sequence of a Cytogenetically Aberrant Human Cancer Cell Line. *PloS Genet.*, **6**.

Conrad,D.F. *et al.* (2010a) Mutation spectrum revealed by breakpoint sequencing of human germline CNVs. *Nat. Genet.*, **42**, 385–391.

Conrad,D.F. *et al.* (2010b) Origins and functional impact of copy number variation in the human genome. *Nature*, **464**, 704–712.

Deng,F.Y. *et al.* (2010) Genome-wide copy number variation association study suggested VPS13B gene for osteoporosis in Caucasians. *Osteoporos. Int.*, **21**, 579–587.

Fong,C.T. *et al.* (1989) Loss of heterozygosity for the short arm of chromosome 1 in human neuroblastomas: correlation with N-myc amplification. *Proc. Natl Acad. Sci. USA*, **86**, 3753–3757.

Geary,R.C. (1930) The frequency distribution of the quotient of two normal variates. *J. R. Stat. Soc.*, **93**, 442–446.

Geary,R.C. (1944) Extension of a theorem by Harald Cramer on the frequency distribution of the quotient of two variables. *J. R. Stat. Soc.*, **107**, 56–57.

Hinkley,D.V. (1969) On ratio of 2 correlated normal random variables. *Biometrika*, **56**, 635–639.

Hoischen,A. *et al.* (2010) Massively parallel sequencing of ataxia genes after array-based enrichment. *Hum. Mutat.*, **31**, 494–499.

Jankowska,A.M. *et al.* (2009) Loss of heterozygosity 4q24 and TET2 mutations associated with myelodysplastic/myeloproliferative neoplasms. *Blood*, **113**, 6403–6410.

Johnston,J.J. *et al.* (2010) Massively parallel sequencing of exons on the X chromosome identifies RBM10 as the gene that causes a syndromic form of cleft palate. *Am. J. Hum. Genet.*, **86**, 743–748.

Kato,M. *et al.* (2010) Population-genetic nature of copy number variations in the human genome. *Hum. Mol. Genet.*, **19**, 761–773.

Li,H. *et al.* (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

McCarroll,S.A. (2010) Copy number variation and human genome maps. *Nat. Genet.*, **42**, 365–366.

Ng,S.B. *et al.* (2009) Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*, **461**, 272–276.

Olshen,A.B. *et al.* (2004) Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics*, **5**, 557–572.

Pinkel,D. *et al.* (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat. Genet.*, **20**, 207–211.

Raca,G. *et al.* (2010) Next generation sequencing in research and diagnostics of ocular birth defects. *Mol. Genet. Metab.*, **100**, 184–192.

Rehman,A.U. *et al.* (2010) Targeted capture and next-generation sequencing identifies C9orf75, encoding taperin, as the mutated gene in nonsyndromic deafness DFNB79. *Am. J. Hum. Genet.*, **86**, 378–388.

Sha,B.Y. *et al.* (2009) Genome-wide association study suggested copy number variation may be associated with body mass index in the Chinese population. *J. Hum. Genet.*, **54**, 199–202.

Shuin,T. *et al.* (1994) Frequent somatic mutations and loss of heterozygosity of the von Hippel-Lindau tumor suppressor gene in primary human renal cell carcinomas. *Cancer Res.*, **54**, 2852–2855.

Stankiewicz,P. and Lupski,J.R. (2010) Structural variation in the human genome and its role in disease. *Annu. Rev. Med.*, **61**, 437–455.

Sudmant,P.H. *et al.* (2010) Diversity of human copy number variation and multicopy genes. *Science*, **330**, 641–646.

Sun,W. *et al.* (2009) Integrated study of copy number states and genotype calls using high-density SNP arrays. *Nucleic Acids Res.*, **37**, 5365–5377.

Teer,J.K. and Mullikin,J.C. (2010) Exome sequencing: the sweet spot before whole genomes. *Hum. Mol. Genet.*, **19**, R145–R151.

Urban,A.E. *et al.* (2006) High-resolution mapping of DNA copy alterations in human chromosome 22 using high-density tiling oligonucleotide arrays. *Proc. Natl Acad. Sci. USA*, **103**, 4534–4539.

Venkatraman,E.S. and Olshen,A.B. (2007) A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics*, **23**, 657–663.

Volpi,L. *et al.* (2010) Targeted next-generation sequencing appoints c16orf57 as clericuzio-type poikiloderma with neutropenia gene. *Am. J. Hum. Genet.*, **86**, 72–76.

Wain,L.V. *et al.* (2009) The role of copy number variation in susceptibility to amyotrophic lateral sclerosis: genome-wide association study and comparison with published loci. *PLoS One*, **4**, e8175.

Yoon,S. *et al.* (2009) Sensitive and accurate detection of copy number variants using read depth of coverage. *Genome Res.*, **19**, 1586–1592.

Downloaded from https://academic.oup.com/bioinformatics/article-abstract/27/19/2648/231564
by guest
on 27 March 2018

TEMPUS EX1002.27

How to map billions of short reads onto genomes

Cole Trapnell & Steven L Salzberg

Mapping the vast quantities of short sequence fragments produced by next-generation sequencing platforms is a challenge. What programs are available and how do they work?

© 2009 Nature America, Inc. All rights reserved.

A new generation of DNA sequencers that can rapidly and inexpensively sequence billions of bases is transforming genomic science. These new machines are quickly becoming the technology of choice for whole-genome sequencing and for a variety of sequencing-based assays, including gene expression, DNA-protein interaction, human resequencing and RNA splicing studies[1–3]. For example, the RNA-Seq protocol, in which processed mRNA is converted to cDNA and sequenced, is enabling the identification of previously unknown genes and alternative splice variants; the ChIP-Seq approach, which sequences immunoprecipitated DNA fragments bound to proteins, is revealing networks of interactions between transcription factors and DNA regulatory elements[4]; and the whole-genome sequencing of tumor cells is uncovering previously unidentified cancer-initiating mutations[5].

One of the challenges presented by the new sequencing technology is the so-called 'read mapping' problem. Sequencing machines made by Illumina of San Diego, Applied Biosystems (ABI) of Carlsbad, California, and Helicos of Cambridge, Massachusetts, produce short sequences of 25–100 base pairs (bp), called 'reads', which are sequence fragments read from a longer DNA molecule present in the sample that is fed into the machine. In contrast to whole-genome assembly, in which these reads are assembled together to reconstruct a previously unknown genome, many of the next-generation sequencing projects begin with a known, or so-called 'reference', genome.

Cole Trapnell and Steven L. Salzberg are at the Center for Bioinformatics and Computational Biology, University of Maryland, College Park, Maryland, USA.
e-mail: cole@cs.umd.edu or salzberg@umd.edu

Table 1 A selection of short-read analysis software

Program	Website	Open source?	Handles ABI color space?	Maximum read length
Bowtie	http://bowtie.cbcb.umd.edu	Yes	No	None
BWA	http://maq.sourceforge.net/bwa-man.shtml	Yes	Yes	None
Maq	http://maq.sourceforge.net	Yes	Yes	127
Mosaik	http://bioinformatics.bc.edu/marthlab/Mosaik	No	Yes	None
Novoalign	http://www.novocraft.com	No	No	None
SOAP2	http://soap.genomics.org.cn	No	No	60
ZOOM	http://www.bioinfor.com	No	Yes	240

In this case, to make sense of the reads, their positions within the reference sequence must be determined. This process is known as aligning or 'mapping' the read to the reference. In one version of the mapping problem, reads must be aligned without allowing large gaps in the alignment (we describe this in more detail in the "Short-read mappers" section below). A more difficult version of the problem arises primarily in RNA-Seq, in which alignments are allowed to have large gaps corresponding to introns (discussed below in the "Spliced-read mappers" section).

These read mapping problems are certainly not new, and there are many programs that perform both spliced and unspliced alignment for the older Sanger-style capillary reads. Even so, these programs neither scale up to the much greater volumes of data produced by short-read sequencers nor scale down to the short read lengths. Aligning the reads from ChIP-Seq or RNA-Seq experiments can take hundreds or thousands of central processing unit (CPU) hours using conventional software tools such as BLAST or BLAT. Fortunately, new software packages designed to meet the computational challenges of short-read sequencing are quickly appearing. Before choosing one, it is essential to understand why the mapping problems are computationally difficult, which difficulties have been overcome and what challenges and opportunities remain.

Challenges of mapping short reads

The first challenge is a practical one: if the reference genome is very large, and if we have billions of reads, how quickly can we align the reads to the genome? Sequence alignment is a classic problem in bioinformatics, supported by a large body of literature describing different variants for both exact and inexact alignment. As a practical matter, the task of mapping billions of sequences to a mammalian-sized genome calls for extraordinarily efficient algorithms, in which every bit of memory is used optimally or near optimally.

The second challenge is strategic: if a read comes from a repetitive element in the reference, a program must pick which copy of the repeat the read belongs to. Because this may be impossible to decide with confidence, the program may choose to report multiple possible locations or to pick a location heuristically. Sequencing errors or variations between the sequenced chromosomes and the reference genome exacerbate this problem, because the

TEMPUS EX1002.28

alignment between the read and its true source in the genome may actually have more differences than the alignment between the read and some other copy of the repeat. The spliced mapping problem faces this same challenge but is further complicated by the possible presence of intron-sized gaps.

DNA sequencers from Illumina, ABI, Roche (of Basel, Switzerland), Helicos and other companies produce millions of reads per run. Complete assays may involve many runs, so an investigator may need to map millions or billions of reads to a genome. For example, the recent cancer genome sequencing project by Ley *et al.*[5] generated nearly 8 billion reads from 132 sequencing runs. A large, expensive computer grid might map the reads from this experiment in a few days using traditional alignment algorithms such as BLAST or BLAT, but such grids are not accessible to everyone. To reduce the computing cost of analysis for sequencing-based assays and to make them available to all investigators, we and others have created a new generation of alignment programs capable of mapping hundreds of millions of short reads on a single desktop computer. Vendors of sequencing machines provide specialized mapping software, such as the ELAND program from Illumina, but in this article we focus on third-party packages, some of which are free and open source. These programs are built on algorithms that exploit features of short DNA sequencing reads to map millions of reads per hour while minimizing both processing time and memory requirements.

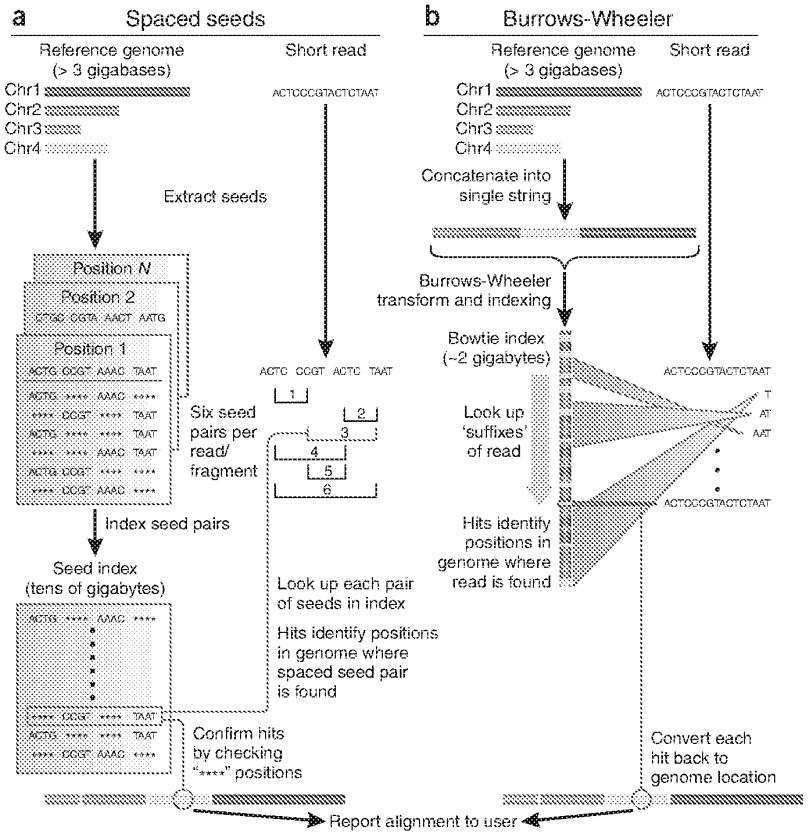


Figure 1 Two recent algorithmic approaches for aligning short (20–200-bp) sequencing reads. (**a**) Algorithms based on spaced-seed indexing, such as Maq, index the reads as follows: each position in the reference is cut into equal-sized pieces, called 'seeds' and these seeds are paired and stored in a lookup table. Each read is also cut up according to this scheme, and pairs of seeds are used as keys to look up matching positions in the reference. Because seed indices can be very large, some algorithms (including Maq) index the reads in batches and treat substrings of the reference as queries. (**b**) Algorithms based on the Burrows-Wheeler transform, such as Bowtie, store a memory-efficient representation of the reference genome. Reads are aligned character by character from right to left against the transformed string. With each new character, the algorithm updates an interval (indicated by blue 'beams') in the transformed string. When all characters in the read have been processed, alignments are represented by any positions within the interval. Burrows-Wheeler–based algorithms can run substantially faster than spaced seed approaches, primarily owing to the memory efficiency of the Burrows-Wheeler search. Chr., chromosome.

Short-read mappers

Such programs as Maq and Bowtie (**Table 1**) use a computational strategy known as 'indexing' to speed up their mapping algorithms. Like the index at the end of a book, an index of a large DNA sequence allows one to rapidly find shorter sequences embedded within it. Maq is based on a straightforward but effective strategy called spaced seed indexing[6] (**Fig. 1a**). In this strategy, a read is divided into four segments of equal length, called the 'seeds'. If the entire read aligns perfectly to the reference genome, then clearly all of the seeds will also align perfectly. If there is one mismatch, however, perhaps due to a single-nucleotide polymorphism (SNP), then it must fall within one of the four seeds, but the other three will still match perfectly. Using similar reasoning, two mismatches will fall in at most two seeds, leaving the other two to match perfectly. Thus, by aligning all possible pairs of seeds (six possible pairs) against the reference, it is possible to winnow the list of candidate locations within the reference where the full read may map, allowing at most two mismatches. Maq's spaced seed index enables it to perform this winnowing operation very efficiently. The resulting set of candidate reads is typically small enough that the rest of the read—that is, the other two seeds that might contain the mismatches—may be individually checked against the reference.

Bowtie takes an entirely different approach, borrowing a technique originally developed for compressing large files called the Burrows-Wheeler transform. Using this transform, the index for the entire human genome fits into less than two gigabytes of memory (an amount that is commonly available on today's desktop and even laptop computers)—in contrast to a spaced seed index, which may require over 50 gigabytes—and yet reads can still be aligned efficiently. Bowtie aligns a read one character at a time to the Burrows-Wheeler–transformed genome (**Fig. 1b**). Each successively aligned new character allows Bowtie to winnow the list of positions to which the read might map. If Bowtie cannot find a location where a read aligns perfectly, the algorithm backtracks to a previous character of the read, makes a substitution and resumes the search. In effect, the Burrows-Wheeler transform enables Bowtie to conquer the mapping problem by first solving a simple subproblem—align one character—and then building on that solution to solve a slightly harder problem—align two characters—and then continuing on to three characters, and so on, until the entire read has been aligned. Bowtie's alignment algorithm is substantially more complicated than Maq's, but Bowtie's alignment speed is more than 30-fold faster[7].

© 2009 Nature America, Inc. All rights reserved.

TEMPUS EX1002.29

Maq and Bowtie both report alignments with up to two mismatches when run in their default modes. In some alignment scenarios, a user may need to allow more mismatches. These two programs were originally designed for reads between 20 and 40 bp long, and both were optimized for human resequencing projects. Even so, Illumina sequencers can now produce reads longer than 100 bp. Additionally, some sequencing projects (such as bacterial or fungal genome sequencing) produce sequences that have many nucleotide-level differences with respect to the closest fully sequenced genome. Finally, the overall quality of reads produced by the new technologies is sensitive to factors such as library preparation, sequencing protocol and even the temperature of the room housing the sequencing machine. Thus, it is essential to know how to change the various default options for any short-read mapper and to be able to identify when those defaults are no longer appropriate.

Several of the new short-read mappers (**Table 1**) are open source, are simple to install and have good documentation and active user communities. The installation package for Bowtie includes a prebuilt index for *Escherichia coli* and a set of sample *E. coli* reads. To run the program on the sample data, just enter the following on the command line:

```
bowtie e_coli reads/e_coli_1000.fq
```

This command will produce a tabular report showing each matching read's identifier, the position(s) where it aligns to the reference sequence, and the number and location of mismatches. Maq reports this same information when you run it with the command:

```
maq.pl easyrun -d outdir
reference.fasta reads.fastq
```

For a given experiment, the fraction of reads that align to the genome depends on many factors. Assuming the sequenced DNA does not contain many mismatched nucleotides compared to the reference, and assuming the reads have passed rudimentary quality filters, most mapping software will find an alignment for 70–75% of the reads. This might seem surprisingly low, but the sequencing technology is still immature—and it's worth noting that Sanger sequencing had success rates of less than 80% until the late 1990s. Note that many reads will align to multiple positions in the genome. Most read mappers can be directed to report alignments only for reads that map to a unique location in the genome.

After aligning the reads, next one might want to call SNPs or view the alignments against the reference sequence. One package for this task is the SAM tools (http://samtools.sourceforge.net). SAM includes a consensus base caller and viewer that can be used either with Maq or with Bowtie.

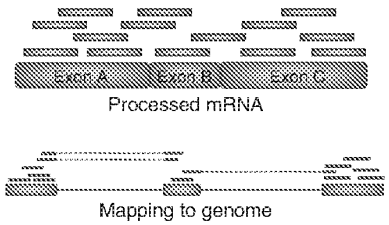


Figure 2 RNA-Seq assays produce short reads sequenced from processed mRNAs. Aligning these reads to the genome with Bowtie or Maq will produce the alignments shown in black but will fail to align the blue reads. A spliced-read mapper such as TopHat or ERANGE will also report the (blue) alignments spanning intron boundaries.

Most read mapping software is designed with whole-genome resequencing in mind, but the programs can be configured for other assays. The manuals for Bowtie and Maq are quite detailed, and the array of choices a user can make can be daunting. Moreover, the list of programs capable of short-read mapping is rapidly growing (**Table 1**), and not every program is ideal or appropriate for every experiment. Fortunately, there are ways to get help. The SeqAnswers message board (http://www.seqanswers.com) is an excellent resource for novice and expert users, frequented by the developers of many short-read mapping programs. One of the most popular SeqAnswers threads contains a catalog of current software for primary analysis and visualization of short-read data.

Spliced-read mappers

The spliced alignment problem, in which cDNA (from processed mRNA) sequences are aligned back to genomic DNA, requires more specialized algorithms. Reads sampled from exon-exon junctions need to be mapped differently from reads that are contained entirely within exons (**Fig. 2**).

To align cDNA reads from RNA-Seq[1–3] experiments, packages such as ERANGE (http://woldlab.caltech.edu/rnaseq) use the positions of exons and introns within known genes as a guide. This allows ERANGE to construct the sequences spanning exon-exon junctions and use them as reference sequences, and then to invoke a standard read mapper such as Maq or Bowtie to align the spliced reads[2]. Because this approach will not discover entirely new splice junctions, some studies have used machine learning methods to predict possible junctions by training statistical models using available reference annotations[8]. In contrast, the TopHat spliced-read mapper (http://tophat.cbcb.umd.edu) does not rely on annotations. Instead, it uses Bowtie (in an initial alignment pass) to identify exons that fully contain some of the reads, and then aligns the remaining reads to junctions between those exons[9]. Another program, G-Mo.R-Se (http://www.genoscope.cns.fr/externe/gmorse), performs a similar spliced alignment while constructing gene models from RNA-Seq data[10].

Limitations and open problems

The current solutions for short-read mapping all have limitations. Mapping programs such as Maq and Bowtie offer very limited support for aligning reads with insertions or deletions (indels). Some read mappers, such as SHRiMP (http://compbio.cs.toronto.edu/shrimp), support ABI's 'color space' sequence representation, but most do not. The spliced alignment programs suffer from these same problems and add a few of their own. Annotation-based methods are of course only as good as the annotations, and many organisms have annotations supported only by homology or computational predictions. Machine learning methods will perform poorly if they are trained on incorrect annotations, and they are prone to overtraining.

Many challenges and questions remain for developers of read mapping software. As all the sequencing machine vendors are trying to produce longer reads, will the short-read mapping programs scale well as the reads get longer? Maq, Bowtie and several other short-read packages support reads longer than 100 bp, but at some point, software designed for longer reads, such as BLAT, may be a better fit for downstream analysis. Furthermore, when mapping reads from an organism that has diverged significantly from its reference genome, how should a program's parameters be adjusted, and can that adjustment happen automatically? How useful is mapping quality in downstream analysis, and should it be computed while aligning reads, as Maq does, or later? The answers to each of these questions will depend on the type of assay and the scale of the analysis, and as long as the technology continues to change, the programs will have to change rapidly to keep up.

1. Nagalakshmi, U. *et al. Science* **320**, 1344–1349 (2008).
2. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. & Wold, B. *Nat. Methods* **5**, 621–628 (2008).
3. Wang, E.T. *et al. Nature* **456**, 470–476 (2008).
4. Johnson, D.S., Mortazavi, A., Myers, R.M. & Wold, B. *Science* **316**, 1497–1502 (2007).
5. Ley, T.J. *et al. Nature* **456**, 66–72 (2008).
6. Li, H., Ruan, J. & Durbin, R. *Genome Res.* **18**, 1851–1858 (2008).
7. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S.L. *Genome Biol.* **10**, R25 (2009).
8. Pan, Q., Shai, O., Lee, L.J., Frey, B.J. & Blencowe, B.J. *Nat. Genet.* **40**, 1413–1415 (2008).
9. Trapnell, C., Pachter, L. & Salzberg, S.L. *Bioinformatics* published online, doi:10.1093/bioinformatics/btp120 (March 16, 2009).
10. Denoeud, F. *et al. Genome Biol.* **9**, R175 (2008).

© 2009 Nature America, Inc. All rights reserved.

TEMPUS EX1002.30

Allele-specific copy number analysis of tumors

Peter Van Loo[a,b,c,1], Silje H. Nordgard[a,d,1], Ole Christian Lingjærde[e], Hege G. Russnes[a,f,g], Inga H. Rye[f], Wei Sun[d,h], Victor J. Weigman[d], Peter Marynen[c], Anders Zetterberg[i], Bjørn Naume[j], Charles M. Perou[d], Anne-Lise Børresen-Dale[a,g,2], and Vessela N. Kristensen[a,g,k,2,3]

[a]Department of Genetics, Institute for Cancer Research, Clinic for Cancer and Surgery, Oslo University Hospital, Montebello, N-0310 Oslo, Norway; [b]Department of Molecular and Developmental Genetics, VIB, B-3000 Leuven, Belgium; [c]Department of Human Genetics, University of Leuven, B-3000 Leuven, Belgium; [d]Department of Genetics, Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599-7264; [e]Department of Informatics, University of Oslo, Blindern, N-0316 Oslo, Norway; [f]Department of Pathology, Clinic for Cancer and Surgery, Oslo University Hospital, Montebello, N-0310 Oslo, Norway; [g]Institute for Clinical Medicine, University of Oslo, Montebello, N-0310 Oslo, Norway; [h]Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27599-7420; [i]Department of Oncology–Pathology, Karolinska Institute, Cancer Center Karolinska, SE-171 77 Stockholm, Sweden; [j]The Cancer Clinic, Clinic for Cancer and Surgery, Oslo University Hospital, Montebello, N-0310 Oslo, Norway; and [k]Institute for Clinical Epidemiology and Molecular Biology, Akershus University Hospital, Faculty of Medicine, University of Oslo, N-1474 Nordbyhagen, Norway

Edited* by David Botstein, Lewis-Sigler Institute, Princeton, NJ, and approved August 23, 2010 (received for review July 9, 2010)

We present an allele-specific copy number analysis of the in vivo breast cancer genome. We describe a unique bioinformatics approach, ASCAT (allele-specific copy number analysis of tumors), to accurately dissect the allele-specific copy number of solid tumors, simultaneously estimating and adjusting for both tumor ploidy and nonaberrant cell admixture. This allows calculation of "ASCAT profiles" (genome-wide allele-specific copy-number profiles) from which gains, losses, copy number-neutral events, and loss of heterozygosity (LOH) can accurately be determined. In an early-stage breast carcinoma series, we observe aneuploidy (>2.7n) in 45% of the cases and an average nonaberrant cell admixture of 49%. By aggregation of ASCAT profiles across our series, we obtain genomic frequency distributions of gains and losses, as well as genome-wide views of LOH and copy number-neutral events in breast cancer. In addition, the ASCAT profiles reveal differences in aberrant tumor cell fraction, ploidy, gains, losses, LOH, and copy number-neutral events between the five previously identified molecular breast cancer subtypes. Basal-like breast carcinomas have a significantly higher frequency of LOH compared with other subtypes, and their ASCAT profiles show large-scale loss of genomic material during tumor development, followed by a whole-genome duplication, resulting in near-triploid genomes. Finally, from the ASCAT profiles, we construct a genome-wide map of allelic skewness in breast cancer, indicating loci where one allele is preferentially lost, whereas the other allele is preferentially gained. We hypothesize that these alternative alleles have a different influence on breast carcinoma development.

breast carcinoma | single-nucleotide polymorphism arrays | bioinformatics | cancer

Genomic changes are key causative events of cancer. Cancer genomes are characterized by numerous sequence changes compared with their normal host counterparts, ranging in size from single base changes (point mutations) to insertions or deletions of large chromosomal fragments and even whole-genome duplications (1, 2). These cancer genomes have been extensively charted by array-comparative genomic hybridization (CGH), SNP arrays (3, 4), and more recently by whole-genome sequencing (5–8). However, correct assembly and interpretation of the data have proven difficult, because tumors often deviate from a diploid state (9, 10), and many contain multiple populations of both tumor and nontumor cells (11, 12). For these reasons, most studies have been limited to reporting gains and losses (array CGH), possibly supplemented by allelic imbalances (SNP arrays), and are unable to assign correct (allele-specific) copy numbers to all loci in the reference genome. Similarly, for correct assembly of complete cancer genomes from sequencing data, the calculation of accurate copy numbers for all loci is a necessary first step for correct interpretation of changes ranging from point mutations to large-scale genomic rearrangements.

We present here an allele-specific copy number analysis of the in vivo breast cancer genome, in which both aneuploidy of the tumor cells and nonaberrant cell infiltration are taken into account. We obtain accurate genome-wide allele-specific copy-number profiles [called "ASCAT (allele-specific copy number analysis of tumors) profiles"] for 91 of 112 breast carcinomas assayed. On the basis of these ASCAT profiles, differences in aberrant tumor cell fraction, ploidy, gains, losses, loss of heterozygosity (LOH), and copy number-neutral events are revealed among the five previously identified molecular breast cancer subtypes. Finally, by evaluating the relative frequency of deletions and duplications of the two possible alleles at each SNP locus, we construct a genome-wide map of allelic skewness, pointing to candidate genes/loci that may drive breast cancer development.

Results

Allele-Specific Copy Number Analysis of Breast Carcinomas. We performed genotyping of 112 breast carcinoma samples using Illumina 109K SNP arrays and constructed an algorithm (ASCAT) to estimate the fraction of aberrant cells and the tumor ploidy, as well as whole-genome allele-specific copy number profiles taking both properties into account (Fig. 1 and Figs. S1 and S2). Using ASCAT, we obtained genome-wide allele-specific copy number profiles (hereafter called ASCAT profiles) for 91 (81%) of the breast carcinomas. Most of the 21 cases (19%) for which ASCAT indicated that no acceptable solution could be found were characterized by significantly larger residual variance in the Log R profiles (Fig. S3). Hence, ASCAT is able to calculate the allele-specific copy numbers of all assayed SNP loci, taking into account tumor aneuploidy and the fraction of aberrant tumor cells, for most of our breast carcinoma cases, and indicates when the quality of the input data are questionable.

We validated the ASCAT profile predictions in three ways. First, we checked ASCAT's consistency and sensitivity to a varying percentage of aberrant tumor cells by applying the algorithm to a dilution series of a tumor sample mixed with different proportions of its germline DNA. Overall, ASCAT profiles were very similar for the different dilutions (Fig. S4). Second, we validated ASCAT's ploidy predictions by experimentally determining the amount of DNA in the tumor cells for 79 of the 91 scored breast cancer cases. We obtained good correspondence with ASCAT's predictions (Fig. 2 and Fig. S5). Finally, FISH experiments were

Author contributions: P.V.L., S.H.N., O.C.L., P.M., C.M.P., A.-L.B.-D., and V.N.K. designed research; P.V.L., S.H.N., O.C.L., H.G.R., I.H.R., W.S., A.Z., C.M.P., A.-L.B.-D., and V.N.K. performed research; V.J.W. and B.N. contributed new reagents/analytic tools; P.V.L., S.H.N., O.C.L., H.G.R., and I.H.R. analyzed data; and P.V.L., S.H.N., O.C.L., A.-L.B.-D., and V.N.K. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

[1]P.V.L. and S.H.N. contributed equally to this work.

[2]A.-L.B.-D. and V.N.K. contributed equally to this work.

[3]To whom correspondence should be addressed. E-mail: Vessela.N.Kristensen@rr-research.no.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1009843107/-/DCSupplemental.

Downloaded by guest on March 10, 2020

TEMPUS EX1002.31

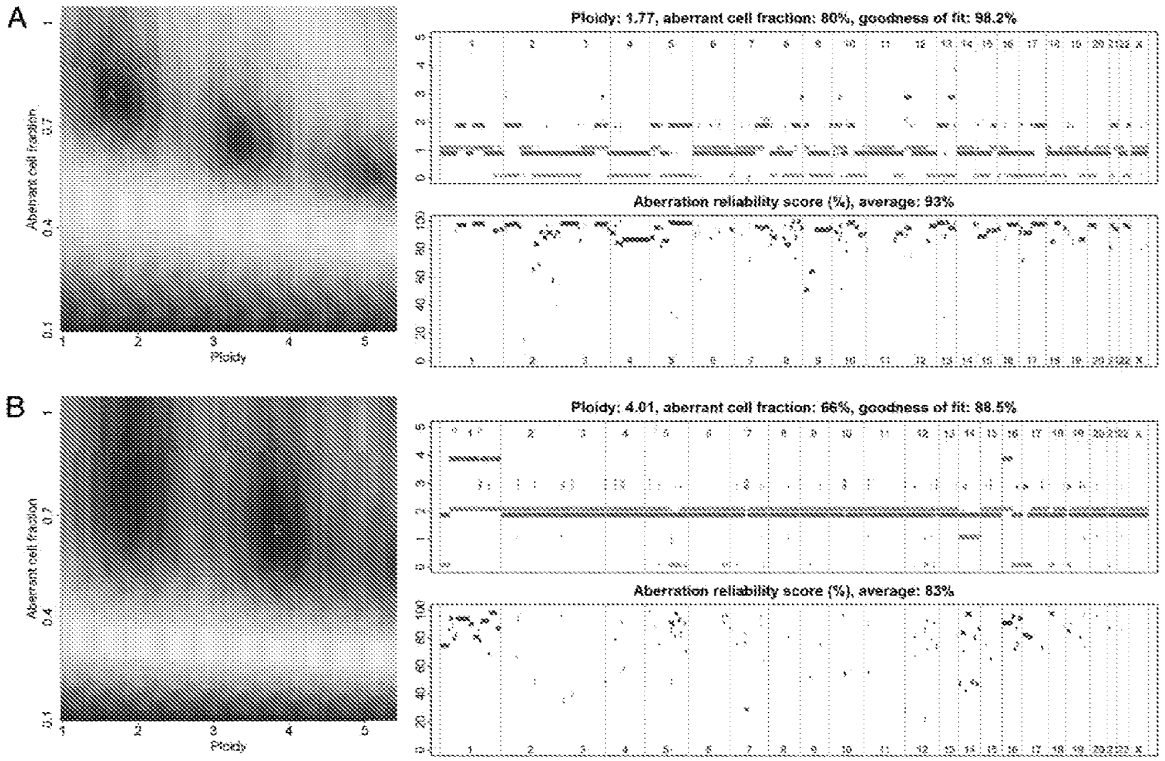


Fig. 1. ASCAT profiles and their calculation. Two examples are given: (*A*) a tumor with ploidy close to 2n and (*B*) a tumor with ploidy close to 4n. (*Left*) ASCAT first determines the ploidy of the tumor cells ψ_t and the fraction of aberrant cells ρ. This procedure evaluates the goodness of fit for a grid of possible values for both parameters (blue, good solution; red, bad solution; detailed in *Materials and Methods*). On the basis of this goodness of fit, the optimal solution is selected (green cross). Using the resulting tumor ploidy and aberrant cell fraction, an ASCAT profile is calculated (*Upper Right*), containing the allele-specific copy number of all assayed loci [copy number on the *y* axis vs. the genomic location on the *x* axis; green, allele with lowest copy number; red, allele with highest copy number; for illustrative purposes only, both lines are slightly shifted (red, down; green, up) such that they do not overlap; only probes heterozygous in the germline are shown]. Finally, for all aberrations found, an aberration reliability score is calculated (*Lower Right*).

performed on 11 of the breast carcinomas for three loci (Table S1), suggesting a good correspondence with the copy numbers estimated by ASCAT, although FISH seems to consistently underestimate the copy number as compared with ASCAT. Together, these validation experiments confirm that ASCAT accurately predicts ASCAT profiles over a broad range of tumor ploidy and fraction of aberrant tumor cells.

Ploidy and Fraction of Aberrant Tumor Cells in Breast Carcinoma. To investigate the relevance of aneuploidy and involvement of nonaberrant cells in breast carcinoma, we examined the ploidy and aberrant cell fraction estimates for our breast carcinoma series. Because the tumor samples were macrodissected to remove as much as possible of the surrounding nontumor tissue by a pathologist, the aberrant cell fraction estimates will reflect intratumoral nonaberrant cells and not normal cells surrounding the tumor. We found that the tumors are on average infiltrated with 49% nonaberrant cells and that 45% of them have a ploidy of 2.7n or higher. These results confirm the importance of taking both nonaberrant cell admixture and tumor aneuploidy into account.

Breast carcinomas can be divided into five distinct subtypes, depending on their expression pattern of 561 transcripts (13). These five subgroups, Luminal A, Luminal B, ERBB2, Basal-like, and Normal-like breast carcinoma, are associated with different clinical outcomes (14). We correlated our ploidy and aberrant cell fraction estimates with these gene-expression–based breast cancer subtypes. Stratifying the estimated percentages of aberrant cells by breast cancer subtype revealed considerable differences (Fig. 3*A*), with the highest fraction of aberrant tumor cells for the Luminal A subtype and the lowest fraction for the ERBB2 and Normal-like subtypes. An evaluation of tumor ploidy stratified by molecular subtype showed lowest ploidy for Luminal A, Basal-like, and Normal-like subtypes and highest ploidy for the ERBB2 subtype (Fig. 3*B*). For the Luminal A subtype, the specific ploidy distribution along with its characteristic paucity of aberrations (and preference for aberrations involving whole chromosome arms) implies a common state of diploidy for these tumors, with a minority of Luminal A carcinomas having undergone polyploidization by endoreduplication (with only few additional aberrations), resulting in a tetraploid state.

ASCAT Profiles Allow Accurate Dissection of Gains, Losses, LOH, and Copy Number-Neutral Events and Grant Insight into Tumor Development. The frequency of gains and losses in a population can be deduced from ASCAT profiles. In our breast carcinoma series, this resulted in similar but slightly more pronounced patterns compared with previous (array-CGH) reports (15–17) (Fig. S6*A*). However, stratification by molecular subtype resulted in considerably higher frequencies of gains and losses specifically for the ERBB2 and the Normal-like subtypes than those directly derived from Log R data (Fig. S6*B*). Hence, contrary to previous reports describing only a limited number of aberrations for these two subtypes (15–17), the use of ASCAT profiles results in clear gains and losses. These aberrations were missed by earlier approaches owing to the high fractions of nonaberrant cells in the ERBB2 and Normal-like subtypes (Fig. 3*A*), a feature taken into account when using ASCAT profiles.

The ASCAT profiles also allow us to investigate LOH and copy number-neutral events. This is not possible using direct evaluation of SNP array data from tumor samples, owing to the admixture

Downloaded by guest on March 10, 2020

TEMPUS EX1002.32

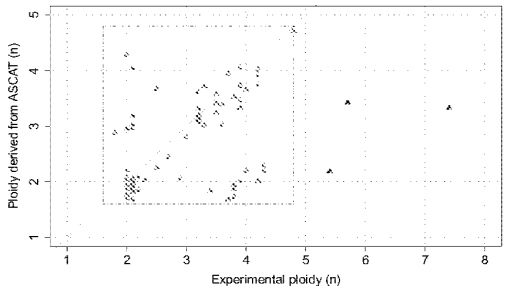


Fig. 2. Validation of tumor ploidy predicted by ASCAT. ASCAT's ploidy estimates are plotted relative to experimentally measured ploidy. Here we define ploidy as the amount of DNA relative to a haploid genome. For 58 of the 79 assayed samples (73.4%), ASCAT's ploidy predictions correspond well with the experimentally determined ploidy (green triangles close to the diagonal). Three samples (3.8%) have an experimentally determined ploidy larger than 5n (blue triangles), outside of the ploidy range used by ASCAT (1.6n–4.8n, depicted by the blue square). Ten breast carcinomas (12.7%) have a predicted ploidy close to 2n, whereas the experimentally determined ploidy was close to 4n (pink triangles). For most of these cases, manual inspection of the copy number profiles could not reveal any indications (missed by ASCAT) that these samples are in fact close to tetraploid (Fig. S5*A*). Indeed, cases that are tetraploid but only show even-numbered allele-specific copy numbers would be recognized as diploid, because the SNP array data do not provide any information to distinguish such tetraploid samples from diploid samples. Alternatively, the experimental method for ploidy determination, applied to a different part of the tumor as the SNP arrays, could be measuring tumor cells in the S phase of the cell cycle, or a different subclone of the tumor. Finally, eight samples (10.1%) show clearly higher ploidy by ASCAT prediction compared with the experimentally determined ploidy (red triangles). A possible explanation for this is the presence of multiple populations of aberrant tumor cells with (slightly) different aberrations (Fig. S5*B*).

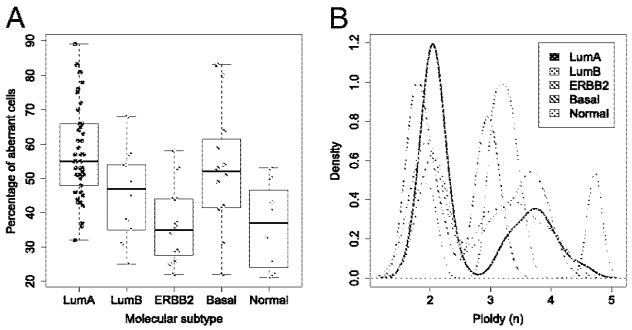


Fig. 3. Percentage of aberrant tumor cells and ploidy across the five breast cancer subtypes. Molecular subtypes used: LumA, Luminal A (n = 45); LumB, Luminal B (n = 10); ERBB2 (n = 12); Basal, Basal-like (n = 12); Normal, Normal-like (n = 8). (*A*) Distribution of percentage of aberrant tumor cells across the five subtypes. The box plots show the median (thick lines) and the lower and upper quartile (boxes). The whiskers reach up to the most extreme value within 1.5 times the interquartile range from the box. Whereas Luminal A carcinomas harbored the highest levels of aberrant tumor cells ($P = 6.9 \times 10^{-6}$, unpaired t test with unequal variance, testing for differences between the Luminal A subtype and all other carcinomas), tumors of the ERBB2 and Normal-like subtype displayed the lowest fraction of aberrant cells ($P = 3.7 \times 10^{-4}$ and $P = 8.4 \times 10^{-3}$, respectively). (*B*) Distribution of ploidy across the five subtypes. The vast majority of Luminal A tumors showed a ploidy close to 2n, with a smaller fraction showing a ploidy close to 4n. Carcinomas of the Luminal B subtype were approximately equally divided among 2n and 4n tumors, with two tumors being 3n. On average, the ERBB2 subgroup displayed the highest level of ploidy but also the broadest range. The Basal-like subgroup showed cases with a ploidy 1.6n–2n and cases of 2.8n–3.2n. Normal-like tumors showed a group of cases with ploidy close to 2n and a group of cases with ploidy above 3n.

with cells not showing these events. A distinct pattern of LOH across the genome emerges (Fig. 4*A*), with LOH being most frequent on chromosome arms 8p, 11q, 16q, and 17p. The q arm of chromosome 16 shows the highest fraction of LOH, including among others SNPs residing in multiple members of the cadherin family (e.g., *CDH1*, *CDH3*, *CDH15*, *CDH13*, and *CDH8*). A calculation of the frequency of copy number-neutral events reveals many regions with a frequency above 20%, with some peaks up to 50% (Fig. 4*B*). We define here a copy number-neutral event as an allelic bias for an SNP heterozygous in the germline such that the total copy number does not differ from the tumor ploidy. Following this definition, copy number-neutral events are all genomic aberrations that cannot be detected by array-CGH. We observe that many genomic regions with higher frequency of loss were also more likely to harbor copy number-neutral events. This is particularly apparent for chromosome/chromosome arm 1p, 2, 3, 4q, 9q, 15, and 19p. This suggests that the frequency of actual loss (loss of one allele, possibly combined with gain of the other allele) is considerably higher than what was previously reported (considering only the total amount of DNA, not distinguishing both alleles). Also notable is chromosome 17q, showing both a high frequency of gains and elevated levels of copy number-neutral events. This chromosome arm harbors among others the *ERBB2* gene, a gene renowned for its relevance in breast cancer. The overall highest frequencies of copy number-neutral events were seen for chromosome 2, 3, 4, 6, 12, and 15, chromosomes not previously reported as key areas for genomic aberrations in breast cancers, suggesting that copy number-neutral events may represent an as-yet unexplored picture of genomic aberrations in breast cancer.

Stratification of genome-wide LOH and copy number-neutral event profiles by breast cancer subtype reveals hitherto unknown differences (Fig. 4 *C* and *D*). A considerably higher frequency of LOH specifically in the Basal-like subtype is immediately apparent ($P = 1.0 \times 10^{-3}$ by an unpaired t test with unequal variance, testing for differences between Basal-like carcinomas and all other carcinomas). This observation, combined with the particular ploidy range of the Basal-like breast carcinomas (Fig. 3*B*), makes us hypothesize that the genomes of Basal-like tumors initially are reduced from a diploid to a partial haploid state (around 1.5n) and subsequently undergo a whole-genome duplication resulting in a ploidy around 3n (Fig. S7).

Alleles Preferentially Gained or Lost in Breast Carcinoma. The ASCAT profiles of our series of 91 breast carcinomas allow us to create a genome-wide map of alleles preferentially gained or lost (Fig. 5). If for a certain SNP the most commonly lost allele is the B allele, whereas the A allele is preferentially kept (or gained), the A allele likely provides a relative advantage for the breast carcinomas. For example, the SNP rs6575883 within the *PPP2R5C* gene on 14q32.31 is germline heterozygous in 30 of the cases. Fifteen of those have a loss, all losing the B allele. In addition, there are four gains, all of the A allele, and two copy number-neutral events, both showing gain of A and loss of B. All these observations point to a skewness in the loss/gain of the two alternative alleles and suggest a relative advantage of the A allele and disadvantage of the B allele acting during breast carcinoma development. A statistical evaluation (Fig. S8) resulted in $P = 9.5 \times 10^{-7}$. Hence, despite the relatively limited size of our dataset (further complicated by the fact that for each SNP only heterozygous cases could be evaluated for this analysis), we were able to identify probes with highly significant allelic skewness in a genome-wide statistical evaluation. This confirms that at least part of the allelic skewness shown in Fig. 5 is likely due to selection, suggesting that loci subject to allelic skewness are potential unique markers for breast cancer development.

Discussion

Since the initial report in 1992 describing CGH (18), and the later adaptation to array technology (19–21), CGH has established itself as the de facto standard for detection of chromosomal aberrations in tumors. However, more than a decade later, it remains difficult to determine accurate genome-wide copy number profiles of tumors from these high-throughput arrays.

Downloaded by guest on March 10, 2020

TEMPUS EX1002.33

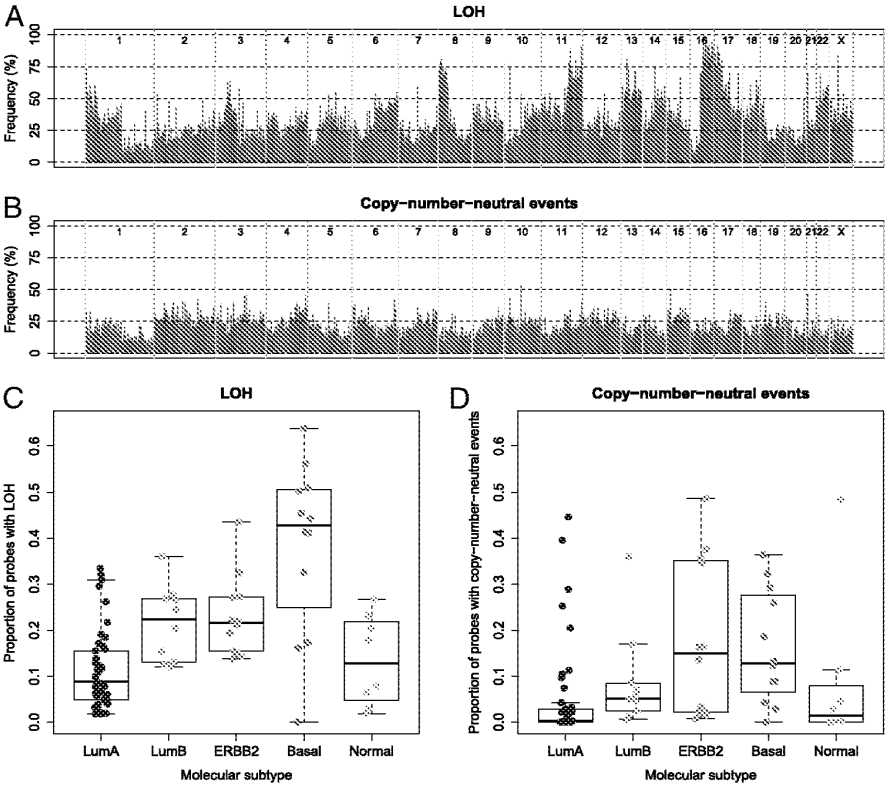


Fig. 4. Frequency of LOH and copy number-neutral events. (*A*) Frequency of LOH across the genome. Probes are shown in genomic order along the *x* axis, from chromosome 1 to chromosome X, where different chromosomes are delimited by gray lines. (*B*) Frequency of copy number-neutral events across the genome. For diploid tumors, copy number-neutral events correspond to a subset of LOH (copy number-neutral LOH), but for, for example, tetraploid tumors, a copy number-neutral event can also be three copies of A and one copy of B. (*C*) Proportion of LOH per case (percentage of probes heterozygous in the germline that have lost this heterozygosity in the tumor), stratified by molecular breast cancer subtypes. Molecular subtypes used and box plot legends are the same as in Fig. 3. The Luminal A subtype shows a significantly lower frequency of LOH compared with the four other subtypes ($P = 2.3 \times 10^{-6}$, unpaired *t* test with unequal variance). Even more striking is the elevated level of LOH for the Basal-like subtype ($P = 1.0 \times 10^{-3}$). Indeed, two thirds of the Basal-like tumors show LOH at more than 40% of the loci heterozygous in the germline. (*D*) Proportion of copy number-neutral events per case, stratified by molecular breast cancer subtypes. The Luminal A ($P = 4.7 \times 10^{-3}$, unpaired *t* test with unequal variance, testing for differences between the Luminal A subtype and all other carcinomas) and Normal-like ($P = 0.95$) subtype display low levels of copy number-neutral events, the Luminal B subgroup shows intermediate levels ($P = 0.99$), and the Basal-like ($P = 0.043$) and ERBB2 subtypes ($P = 0.064$) show the highest frequencies of copy number-neutral events.

Complicating factors are that tumor cells are often aneuploid (9, 10) and that tumor samples contain multiple populations of both tumor and nontumor cells (11, 12). Although some studies aim to bring these effects into the equation (22, 23), these difficulties still remain today. In addition to these limitations, array-CGH provides no information regarding which of the two alternative alleles has been gained or lost, and it overlooks copy number-neutral aberrations.

The introduction of SNP array technology (24, 25) holds the promise to solve these issues, because allele-specific measurements allow estimation of the amount of aberrant and nonaberrant cells in a specimen and clearly show deviations from diploidy. Recently, a number of computational methods have been developed that aim to take into account either tumor aneuploidy or infiltration of nontumoral cells (26–31). However, to calculate correct genome-wide allele-specific copy numbers from SNP array data of nonmicrodissected tumor samples, both these effects need to be modeled simultaneously.

We developed a unique algorithm, ASCAT, to infer ASCAT profiles (accurate genome-wide allele-specific copy number profiles) from SNP array data, estimating and correcting for both tumor cell aneuploidy and nonaberrant cell admixture. We validated ASCAT's copy number predictions by FISH, its sensitivity to increasing nonaberrant cell involvement by application to a dilution series of a tumor sample, and its ploidy predictions by experimental ploidy measurements.

The dissection of cancer genomes is taken to the next step by the recent introduction of cancer genome sequencing (5–8). We believe ASCAT profiles could be useful tools for interpretation of these data, aiding in the assembly of the data and in the identification of changes varying in size from point mutations to complex rearrangements.

The distribution of ploidy and aberrant cell fraction across our early-stage breast carcinoma series (Fig. 3) suggests that analysis methods not taking both ploidy and nonaberrant cell infiltration into account will misinterpret at least 50% of the cases. This may lead to underestimation of the number of aberrations in tumors showing high nonaberrant cell admixture (as observed in ERBB2 and Normal-like breast carcinomas) or to misinterpretation of nearly all aberrations in aneuploid tumors.

ASCAT profiles allow identification of LOH and copy number-neutral events, invisible to array-CGH (and misinterpreted by SNP array methods not correcting for both aneuploidy and nonaberrant cell infiltration). The genomic distribution of LOH (Fig. 4*A*) corresponds very well with that of losses (Fig. S6*A*). This is in large part because many losses also result in loss of heterozygosity. Unexpectedly, however, there were also correspondences between losses and copy number-neutral events (two entirely distinct classes of aberrations), suggesting that the frequency of loss of one allele (possibly combined with gain of the other allele) is considerably higher than previously reported.

We identify characteristic differences in tumor ploidy, nonaberrant cell admixture, and frequency of gains, losses, LOH, and copy number-neutral events among the five molecularly defined breast cancer subtypes, many of which were previously unknown. These findings confirm the added value of our approach and at the same time support the hypothesis that these molecular subtypes are distinct biological entities. For example, we find a high fraction of nonaberrant cells in the ERBB2 and Normal-like subtypes (Fig. 3*A*), two tumor subtypes for which previous reports described only a limited number of aberrations (15–17). The correction for nonaberrant cell involvement shows that these tumors in fact do not harbor fewer aberrations than the other subtypes (Fig. S6*B*) but rather that these aberrations were missed by earlier approaches not adjusting for nonaberrant cell involvement. The *ERBB2* gene is an important tumor antigen for the induction of $CD8^+$-mediated T-cell responses in breast carcinomas, and patients carrying tumors overexpressing

Downloaded by guest on March 10, 2020

TEMPUS EX1002.34

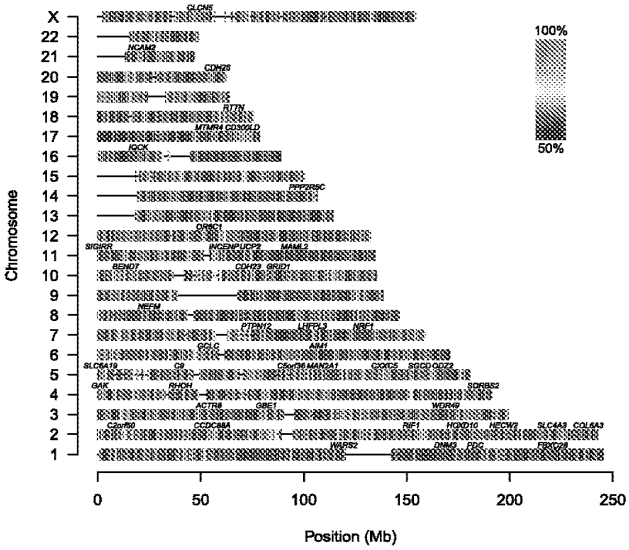


Fig. 5. Genome-wide map of allelic skewness. SNPs that show no allelic skewness (no allele is preferentially gained or lost) should show approximately equal frequencies of loss and gain for both alleles. Here, the frequency of the most frequently gained/lost allele is shown. Alleles without allelic skewness should have a frequency of 50% (blue), whereas alleles that are completely skewed have a frequency of 100% (red). For each SNP, we selected the cases from our series that are germline heterozygous. We count how many cases show gains (of A vs. of B), losses (of A vs. of B), and copy number-neutral events (with gain of A and loss of B vs. with gain of B and loss of A). We combined the counts for gain of A, loss of B, and copy number-neutral events with gain of A and loss of B, and the counts for gain of B, loss of A, and copy number-neutral events with gain of B and loss of A, and display the frequency of the most frequently skewed allele. Only probes with a total of at least 10 gains, losses, and copy number-neutral events are shown. Gene symbols shown contain at least one SNP with a most frequently gained/lost allele frequency of 95% or more.

this growth factor receptor are often mounting immune responses to ERBB2-derived peptides (32). Hence, the attraction of T cells may at least in part explain the higher fraction of nonaberrant cells in tumors of the ERBB2 subtype.

Our findings also allow insight into tumor development. Luminal A breast carcinomas are typically diploid, showing only a limited number of aberrations (mostly affecting entire chromosome arms). A minority of them have become tetraploid by endoreduplication, with very few additional aberrations. Breast carcinomas of the Basal-like subtype in contrast show numerous aberrations, with most chromosomes affected. There are considerably more losses than gains, resulting in a ploidy of 1.6n to 2n. In a later stage, some of these tumors undergo a whole-genome duplication. We hypothesize that these partially haploid genomes become unstable (at around 1.5n), resulting in a selection for more stable triploid genomes. The ASCAT profiles of these triploid basal-like carcinomas, displaying extensive LOH, confirm that this whole-genome duplication occurs late in tumor development, after the tumor genome has acquired a large number of aberrations (Fig. S7).

Finally, we construct a genome-wide map of allelic skewness, indicating loci where one allele is preferentially lost whereas the other allele is preferentially gained. We hypothesize that these alternative alleles have a different effect on breast carcinoma development, with the allele preferentially gained showing a beneficial effect on the breast carcinoma, compared with the allele preferentially lost. Interestingly, the gene containing the SNP showing the most extreme allelic skewness (21 aberrations, all pointing to preferential gain of A and preferential loss of B), *PPP2R5C* (PP2A, B subunit, B56γ isoform), has been shown to mediate DNA-damage induced dephosphorylation of p53 (33), and as part of the heterotrimeric complex PP2A may play an important tumor suppressive role in multiple cancers, including breast cancer (34, 35).

Materials and Methods

Breast Carcinoma Series. The study population of early-stage breast carcinomas has been described previously (16, 36). It consists of 112 blood-tumor pairs. A part of each of the surgically removed tumor specimens was frozen directly at −80 °C and stored. The frozen tumors were then macrodissected by a pathologist. Two frozen sections from each were examined by microscope to secure representative tumor tissue before DNA extraction. The blood DNA was isolated from the lymphocyte fraction of peripheral blood. Both were analyzed using the Human-1 109K BeadChip SNP array platform (Illumina). A more detailed description can be found in our earlier study (16). Because these were all breast cancers from female patients, the SNP array data for chromosome Y was not used, leaving 109,302 SNPs in total.

Although multiple methods have been developed to perform molecular breast cancer subtyping (37, 38), in this study we opted to use the same subtypes as in our previous study (16, 36), for easier comparison.

ASCAT Algorithm. Illumina SNP arrays deliver two output tracks: Log R, a measure of total signal intensity, and B allele frequency (BAF), a measure of allelic contrast (25). The Log R track is similar to the output given by common array-CGH platforms and quantifies the (total) copy number of each genomic locus. The BAF track shows the relative presence of each of the two alternative nucleotides (called "A" and "B") at each SNP locus profiled.

Two recurrent phenomena complicate the analysis of genotype profiles of cancer samples and occurred frequently in our series as well: infiltration of nonaberrant cells and aneuploidy of tumor cells (Fig. S1). We express the Log R and BAF data (*r* and *b*, respectively) as a function of the allele-specific copy number ($n_{A,i}$ and $n_{B,i}$), accounting for nonaberrant cell infiltration and tumor aneuploidy (*SI Materials and Methods* for details):

$$r_i = \gamma \log_2\left(\frac{2(1-\rho) + \rho(n_{A,i} + n_{B,i})}{\psi}\right) \quad [1]$$

$$b_i = \frac{1 - \rho + \rho n_{B,i}}{2 - 2\rho + \rho(n_{A,i} + n_{B,i})} \quad [2]$$

In Eqs. **1** and **2**, *i* represents the genomic location, and γ is a constant depending on the SNP array technology used. The average ploidy of the sample is modeled by $\psi = 2(1 - \rho) + \rho\psi_t$, with ψ_t the tumor ploidy (ranging from 1.6 to 4.8, corresponding with a tumor ploidy range of 1.6n to 4.8n). The aberrant cell fraction of a sample is modeled by ρ, a value between 0 and 1. The parameter γ can be obtained from the literature (25) (it is the drop in Log R in case of a deletion in a 100% pure sample), whereas ρ and ψ_t need to be estimated from the data for each tumor sample separately. On the basis of these equations, we can express the allele-specific copy number estimates as a function of the data and the parameters (*SI Materials and Methods*).

To make our method less sensitive to noise in the input data, both Log R and BAF are preprocessed by a specially designed segmentation and filtering algorithm, Allele-Specific Piecewise Constant Fitting (ASPCF) (*SI Materials and Methods* for details). First, probes for which the germline DNA is homozygous (i.e., probes in the BAF bands at heights 0 and 1) are removed from the BAF track, because they are uninformative for determination of the total copy number. Because our breast carcinoma series consisted of blood and tumor pairs, we used the genotypes generated from the blood samples to eliminate the probes homozygous in the germline (Fig. S2*A*). ASPCF then fits piecewise constant functions simultaneously to the Log R and BAF data, requiring change points to appear at the same genomic locations in the two fitted functions (Fig. S2*B*). As a result, a segmentation of the genome is obtained, each segment corresponding to a genomic region between two adjacent change points (or between a change point and the start/end of a chromosome arm). For Log R, a single fitted value is obtained for each segment, whereas for BAF the output from ASPCF may consist of either one or two values per segment. These values are symmetric around 0.5. If the aberrant cells are found to be balanced (equal number of As and Bs), only one value, 0.5, is returned. If the aberrant cells show an allelic bias, it will be present in both directions (e.g., SNPs with ABB and AAB genotype will both be present), resulting in two values symmetric around 0.5 being output from ASPCF (Fig. S2*B*).

These ASPCF-smoothed data are subsequently used as input of our ASCAT algorithm (implemented in R), to estimate the parameters ρ (aberrant cell fraction) and ψ_t (tumor ploidy), as well as the absolute allele-specific copy

Downloaded by guest on March 10, 2020

TEMPUS EX1002.35

number calls ($\hat{n}_{A,i}$ and $\hat{n}_{B,i}$). Using the fact that true copy numbers are nonnegative whole numbers, we seek values for ρ and ψ_t such that the allele-specific copy number estimates are as close as possible to nonnegative whole numbers for germline heterozygous SNPs. Optimal values for ρ and ψ_t were estimated as follows:

(*i*) Genome-wide allele-specific copy number profiles were calculated for a grid of ρ (0.10, 0.11, ..., 1.05) and ψ_t values (1.00, 1.05, ..., 5.40)

(*ii*) for each parameter value combination, the total distance to a nonnegative whole-number solution for the genome-wide allele-specific copy number profiles was calculated and summed over all SNPs (Eq. 3).

$$d(\rho,\psi_t)=\sum_i w_i\Big(\big(\hat{n}_{A,i}(\rho,\psi_t)-round(\hat{n}_{A,i}(\rho,\psi_t))\big)^2+\big(\hat{n}_{B,i}(\rho,\psi_t)-round(\hat{n}_{B,i}(\rho,\psi_t))\big)^2\Big) \quad [3]$$

Here, the *round()* function rounds to the nearest nonnegative whole number. The weight w_i = 1 for probes in segments with allelic bias (BAF $\neq$ 0.5), and w_i = 0.05 for probes in segments without allelic bias (BAF = 0.5), because the former were deemed more likely aberrant segments.

(*iii*) All local minima were determined and were considered as possible interpretations of the data. For each possible interpretation, a goodness-of-fit score is calculated. This goodness-of-fit g is calculated as a linear rescaling of the total distance to nonnegative whole numbers to a percentage: g = 100% when d = 0 and g = 0 when d = the distance obtained when the allele-specific copy numbers for each SNP differ 0.25 from nonnegative whole numbers ($d=\sum w_i(2{\cdot}0.25^2)$). The value 0.25 was selected as a reasonable maximum distance (averaged over all probes), taking into consideration the fact that this goodness-of-fit is calculated only for local minima.

(*iv*) Local minima corresponding to unlikely interpretations are automatically excluded by ASCAT: (*1*) solutions with ploidy (calculated as the average total copy number) outside a user-defined range (1.6n–4.8n), (*2*) solutions with a too-low percentage of aberrant tumor cells (ρ < 0.20), (*3*) "floating" solutions—solutions that show genomic aberrations but that do not show any SNPs with copy number 0 of either allele (by this criterion, ASCAT avoids interpretations with higher ploidy when there is no evidence of higher ploidy), and (*4*) solutions with a goodness-of-fit below 80%.

(*v*) If one candidate solution remains, then this solution is reported. If multiple solutions remain, these are ranked according to their goodness of fit, and the highest ranking solution is reported. For the reported solution, ASCAT returns the percentage of aberrant tumor cells, the tumor ploidy (calculated as the average total copy number), the goodness of fit, and the whole-genome allele-specific copy number profile of the tumor (ASCAT profile), as well as an aberration reliability score for each aberration found (*SI Materials and Methods* for details).

Software and Data Availability. The ASCAT and ASPCF software, the SNP array data, and the allelic skewness data are available at http://www.ifi.uio.no/bioinf/Projects/ASCAT.

ACKNOWLEDGMENTS. We thank Elmar Bucher and Tuuli Lappalainen for bioinformatical assistance; Therese Sørlie for sharing the correlation values to the molecular breast cancer subtypes; Grethe I. G. Alnaes and Fredrik E. Johansen for performing part of the Illumina genotyping; and Trond Stokke for valuable discussion. Operating costs for genotyping were provided by Norwegian Research Council Grants 155218/V40 and 175240/S10 (to A.-L.B.-D.), Functional Genomics-Norsk Forskningsråd (Norwegian Research Council) (FUGE-NFR) FUGE-NFR 181600/V11 (to V.N.K.), and a Swizz Bridge Award (A.-L.B.-D.). Laboratory assistance was funded by Norwegian Cancer Society Grants D99061 (to A.-L.B.-D.) and D03067 (to V.N.K.). P.V.L. is a postdoctoral researcher of the Research Foundation–Flanders (FWO) and is as a visiting scientist at the Institute for Cancer Research supported by travel grants from the FWO and from the European Association for Cancer Research. S.H.N. is a postdoctoral fellow of the Norwegian Cancer Association (PK01-2007-0356) and is supported by a travel grant from Lillemor Grobstoks Legacy for Cancer Research. C.M.P. was supported by National Cancer Institute Breast Specialized Program of Research Excellence Grant P50-CA58223-09A1, the Breast Cancer Research Foundation, and the V Foundation for Cancer Research.

1. Stratton MR, Campbell PJ, Futreal PA (2009) The cancer genome. *Nature* 458:719–724.
2. Balmain A, Gray J, Ponder B (2003) The genetics and genomics of cancer. *Nat Genet* 33 (Suppl):238–244.
3. Mullighan CG, et al. (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* 446:758–764.
4. Cancer Genome Atlas Research Network (2008) Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 455:1061–1068.
5. Mardis ER, et al. (2009) Recurring mutations found by sequencing an acute myeloid leukemia genome. *N Engl J Med* 361:1058–1066.
6. Stephens PJ, et al. (2009) Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* 462:1005–1010.
7. Pleasance ED, et al. (2010) A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* 463:184–190.
8. Pleasance ED, et al. (2010) A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* 463:191–196.
9. Holland AJ, Cleveland DW (2009) Boveri revisited: Chromosomal instability, aneuploidy and tumorigenesis. *Nat Rev Mol Cell Biol* 10:478–487.
10. Rajagopalan H, Lengauer C (2004) Aneuploidy and cancer. *Nature* 432:338–341.
11. Campbell LL, Polyak K (2007) Breast tumor heterogeneity: Cancer stem cells or clonal evolution? *Cell Cycle* 6:2332–2338.
12. Witz IP, Levy-Nissenbaum O (2006) The tumor microenvironment in the post-PAGET era. *Cancer Lett* 242:1–10.
13. Perou CM, et al. (2000) Molecular portraits of human breast tumours. *Nature* 406: 747–752.
14. Sørlie T, et al. (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 98:10869–10874.
15. Bergamaschi A, et al. (2006) Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. *Genes Chromosomes Cancer* 45:1033–1040.
16. Nordgard SH, et al. (2008) Genome-wide analysis identifies 16q deletion associated with survival, molecular subtypes, mRNA expression, and germline haplotypes in breast cancer patients. *Genes Chromosomes Cancer* 47:680–696.
17. Chin K, et al. (2006) Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell* 10:529–541.
18. Kallioniemi A, et al. (1992) Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* 258:818–821.
19. Solinas-Toldo S, et al. (1997) Matrix-based comparative genomic hybridization: Biochips to screen for genomic imbalances. *Genes Chromosomes Cancer* 20:399–407.
20. Pinkel D, et al. (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet* 20:207–211.
21. Pollack JR, et al. (1999) Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* 23:41–46.
22. Lyng H, et al. (2008) GeneCount: Genome-wide calculation of absolute tumor DNA copy numbers from array comparative genomic hybridization data. *Genome Biol* 9: R86.
23. Wang K, Li J, Li S, Bolund L, Wiuf C (2009) Estimation of tumor heterogeneity using CGH array data. *BMC Bioinformatics* 10:12.
24. McCarroll SA, et al. (2008) Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet* 40:1166–1174.
25. Peiffer DA, et al. (2006) High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. *Genome Res* 16:1136–1148.
26. Attiyeh EF, et al. (2009) Genomic copy number determination in cancer cells from single nucleotide polymorphism microarrays based on quantitative genotyping corrected for aneuploidy. *Genome Res* 19:276–283.
27. Staaf J, et al. (2008) Segmentation-based detection of allelic imbalance and loss-of-heterozygosity in cancer cells using whole genome SNP arrays. *Genome Biol* 9:R136.
28. Gardina PJ, Lo KC, Lee W, Cowell JK, Turpaz Y (2008) Ploidy status and copy number aberrations in primary glioblastomas defined by integrated analysis of allelic ratios, signal ratios and loss of heterozygosity using 500K SNP Mapping Arrays. *BMC Genomics* 9:489.
29. Pounds S, et al. (2009) Reference alignment of SNP microarray signals for copy number analysis of tumors. *Bioinformatics* 25:315–321.
30. Greenman CD, et al. (2010) PICNIC: An algorithm to predict absolute allelic copy number variation with microarray cancer data. *Biostatistics* 11:164–175.
31. Sun W, et al. (2009) Integrated study of copy number states and genotype calls using high-density SNP arrays. *Nucleic Acids Res* 37:5365–5377.
32. Pupa SM, Tagliabue E, Ménard S, Anichini A (2005) HER-2: A biomarker at the crossroads of breast cancer immunotherapy and molecular medicine. *J Cell Physiol* 205:10–18.
33. Li HH, Cai X, Shouse GP, Piluso LG, Liu X (2007) A specific PP2A regulatory subunit, B56gamma, mediates DNA damage-induced dephosphorylation of p53 at Thr55. *EMBO J* 26:402–411.
34. Sablina AA, et al. (2007) The tumor suppressor PP2A Abeta regulates the RalA GTPase. *Cell* 129:969–982.
35. Esplin ED, et al. (2006) The glycine 90 to aspartate alteration in the Abeta subunit of PP2A (PPP2R1B) associates with breast cancer and causes a deficit in protein function. *Genes Chromosomes Cancer* 45:182–190.
36. Naume B, et al. (2007) Presence of bone marrow micrometastasis is associated with different recurrence risk within molecular subtypes of breast cancer. *Mol Oncol* 1:160–171.
37. Sørlie T, et al. (2003) Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* 100:8418–8423.
38. Parker JS, et al. (2009) Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 27:1160–1167.

Downloaded by guest on March 10, 2020

TEMPUS EX1002.36

Published online 25 May 2019

Nucleic Acids Research, 2019, Vol. 47, No. 15 e87
doi: 10.1093/nar/gkz474

High efficiency error suppression for accurate detection of low-frequency variants

Ting Ting Wang[1,2], Sagi Abelson[2,3], Jinfeng Zou[2], Tiantian Li[2], Zhen Zhao[2], John E. Dick[2,4], Liran I. Shlush[2,5], Trevor J. Pugh[1,2,3,*] and Scott V. Bratman[1,2,6,*]

[1]Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada, [2]Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada, [3]Ontario Institute for Cancer Research, Toronto, Ontario, Canada, [4]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada, [5]Department of Immunology, Weizmann Institute of Science, Rehovot, Israel and [6]Department of Radiation Oncology, University of Toronto, Toronto, Ontario, Canada

Received May 23, 2018; Revised April 28, 2019; Editorial Decision May 13, 2019; Accepted May 16, 2019

ABSTRACT

Detection of cancer-associated somatic mutations has broad applications for oncology and precision medicine. However, this becomes challenging when cancer-derived DNA is in low abundance, such as in impure tissue specimens or in circulating cell-free DNA. Next-generation sequencing (NGS) is particularly prone to technical artefacts that can limit the accuracy for calling low-allele-frequency mutations. State-of-the-art methods to improve detection of low-frequency mutations often employ unique molecular identifiers (UMIs) for error suppression; however, these methods are highly inefficient as they depend on redundant sequencing to assemble consensus sequences. Here, we present a novel strategy to enhance the efficiency of UMI-based error suppression by retaining single reads (singletons) that can participate in consensus assembly. This 'Singleton Correction' methodology outperformed other UMI-based strategies in efficiency, leading to greater sensitivity with high specificity in a cell line dilution series. Significant benefits were seen with Singleton Correction at sequencing depths ≤16 000×. We validated the utility and generalizability of this approach in a cohort of >300 individuals whose peripheral blood DNA was subjected to hybrid capture sequencing at ∼5000× depth. Singleton Correction can be incorporated into existing UMI-based error suppression workflows to boost mutation detection accuracy, thus improving the cost-effectiveness and clinical impact of NGS.

INTRODUCTION

High-throughput sequencing technologies have revolutionized genetic and biomedical research by uncovering alterations responsible for the development of disease. Although considerable progress has been made toward germline and somatic variant detection, identification of variants at lower allele frequencies remains hindered by sequencing errors and technical artefacts. This has numerous implications in oncology, particularly in liquid biopsy applications, where tumour DNA fragments may be present at frequencies <0.01% (1,2). Sensitive detection is difficult in these scenarios as sequencer error rates average ∼0.1–1% (3,4).

A promising strategy to suppress errors uses unique molecular identifiers (UMIs) to compare multiple reads derived from the same DNA fragment (Figure 1A) (5–7). Errors that are found in individual reads are removed, and only variants present across all redundant reads are retained to form a single-strand consensus sequence (SSCS). In addition, strand-aware duplex correction is needed to eliminate artefacts from oxidative damage; duplex consensus sequences (DCSs) retain only true variants found on both strands of a fragment by comparing complementary SSCSs (Figure 1A) (8–10). While duplex methods allow for greater error suppression (Supplementary Figure S1), the efficiency of DCS recovery from SSCSs is poor (15–47%, Figure 1B) and reliant on sequencing coverage (Supplementary Figure S2).

A major limitation of current UMI-based error correction methods is the dependence on redundant sequencing

*To whom correspondence should be addressed. Tel: +1 416 946 2132; Fax: +1 416 946 6561; Email: scott.bratman@rmp.uhn.ca
Correspondence may also be addressed to Trevor J. Pugh. Tel: +1 416 581 7689; Fax: +1 416 581 7430; Email: trevor.pugh@utoronto.ca
Present addresses:
Scott V. Bratman, MaRS Centre, 101 College Street, Princess Margaret Cancer Research Tower, Room 13-305. Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 1L7, Canada.
Trevor J. Pugh, MaRS Centre, 101 College Street, Princess Margaret Cancer Research Tower, Room 9-305. Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 1L7, Canada.

© The Author(s) 2019. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Downloaded from https://academic.oup.com/nar/article-abstract/47/15/e87/5486633 by guest on 11 March 2020

(11). This results in poor efficiency with low yield of unique sequences despite high sequencing costs. These inefficiencies are further magnified in duplex UMI methods, where both strands of a molecule must be redundantly sequenced (8–10). This is problematic, as uneven sequencing often arises from amplification biases, stochastic sampling, and inadequate coverage (11–13). These factors limit the applicability of duplex correction to only 0.5–2.5% of sequenced reads (Figure 1B). Furthermore, current UMI-based strategies do not utilize error suppression for single reads (singletons) that have not been redundantly sequenced. This is detrimental as singletons may account for over half of all reads in a moderately deep sequenced sample (defined as ~1000×–10 000× coverage in this study).

To address these limitations, we developed a 'Singleton Correction' methodology that enables error suppression in singletons (Figure 1A). By utilizing the large number of singletons present in hybrid capture deep sequencing data, Singleton Correction allows dramatically more sequences to be corrected. Unlike traditional UMI methods that are restricted to redundant reads, our method also eliminates errors in singletons using reads from the complementary strand. Here, we analyzed a combination of cell line and clinical samples and found that Singleton Correction consistently improved the efficiency of traditional duplex correction methods and increased sensitivity while maintaining high specificity for calling low-allele-frequency variants.

MATERIALS AND METHODS

Targeted panel design

We constructed hybrid capture panels targeting genomic footprints representing two different experimental strategies. A 13 kb panel we named 'SmallDeep' was intended for ultra-deep sequence coverage and encompassed exons of five genes (*KRAS, NRAS, BRAF, EGFR* and *PIK3CA*) important in the mitogen-activated protein kinase (MAPK) pathway. We have previously used this panel for cell-free DNA sequencing analysis in multiple myeloma (14). A 1.2 Mb panel we named 'LargeMid' was intended for moderately deep sequence coverage and encompassed exons from 260 leukemia associated genes (xGen® Acute Myeloid Leukemia Cancer Panel, IDT). We have previously used this panel for the identification of pre-leukemic mutations in peripheral blood leukocytes of individuals who later developed acute myeloid leukemia (15).

Cell line dilution series

To evaluate analytical performance of mutational profiling, we created cell line dilution series using sheared genomic DNA from cancer cell lines with known genetic alterations to emulate varying levels of mutant allele frequencies (Supplementary Table S1). The source of cell line genomic DNA was as follows: MOLM13 was obtained from DSMZ, SW48 was obtained from ATCC, HCT116 was a kind gift of Dr Daniel De Carvalho, and MM1S was obtained from Dr Rodger Tiedemann. For LargeMid, we performed a dilution series at ratios of 1/5 in duplicate from 5% to 0.04% (six dilution points including 100% and 0% levels, $n = 2$ libraries per dilution point, total of 12 libraries). For SmallDeep, we used a dilution series at ratios of 1/10 from 1:1 to $1:10^6$ (eight dilution points including 100% and 0% levels, $n = 1$ library per dilution point, total of eight libraries).

Next-generation sequencing library preparation

Illumina-compatible next-generation sequencing (NGS) libraries were prepared for each dilution point from genomic DNA. Briefly, 60–100 ng DNA was sheared before library construction using a Covaris M220 sonicator (Covaris, Woburn, MA, USA) to attain median fragment sizes of 180–250 bp. The DNA libraries were constructed using the KAPA Hyper Prep kit (#KK8504, Kapa Biosystems, Wilmington, MA, USA) with custom adapters containing 2 bp in-line duplex unique molecular identifiers (UMIs, Supplementary Tables S2 and S3). Following end repair and A-tailing, we performed adapter ligation overnight using 100-fold molar excess of adapters. Agencourt AMPure XP beads (Beckman-Coulter) were used for library clean up and ligated fragments were amplified between 4 and 8 cycles using 0.5 μM Illuminal universal and sample-specific index primers.

Target capture and sequencing

Indexed Illumina libraries were pooled together in a single capture hybridization (Supplementary Table S1). Following the IDT Hybridization capture protocol, each pool of DNA was combined with 5 μl of 1 mg Cot-I DNA (Invitrogen) and 1 nmol each of xGen Universal Blocking Oligo (Integrated DNA Technologies, Coralville, IA, USA) to prevent cross hybridization and minimize off-target capture. Samples were dried and re-suspended in hybridization buffer and enhancer. Target capture with custom xGen Lockdown Probes (Integrated DNA Technologies, Coralville, IA, USA) was performed overnight. Streptavidin-coated magnetic beads were used to isolate hybridized targets according to manufacturer's specifications. Captured DNA fragments were amplified with 10–15 cycles of PCR. Pooled libraries were sequenced using 100–125 bp paired-end runs on Illumina platforms (HiSeq v3 2000, HiSeq 2500) at the Princess Margaret Genomics Centre (www.pmgenomics.ca). NGS libraries for SmallDeep and LargeMid were sequenced to an average of 186 312× and 4223× target coverage, respectively (see QC metrics in Supplementary Table S1).

Data preprocessing

Sequencing reads were de-multiplexed using sample-specific indices followed by removal of the first 3 bp of each read, as these correspond to the 2 bp UMI and single T invariant spacer sequence necessitated for ligation. Reads without the invariant T sequence were discarded as they were not compliant with this design. The extracted UMIs from paired-end reads were grouped and written into the FASTQ sequence identifier header of each read for downstream *in silico* molecular identification. FASTQ files were mapped to the human reference genome hg19 using BWA (v 0.7.12) (16), processed using the Genome Analysis ToolKit (GATK) IndelRealigner (v 3.4-46) (17), and sorted by

Downloaded from https://academic.oup.com/nar/article-abstract/47/15/e87/5486633 by guest on 11 March 2020

genome position and indexed using SAMtools (v 1.3) (18). This process created sorted BAM files containing sequence alignment data.

Barcodes used in UMIs

Short oligonucleotide barcodes have the benefit of reduced cost for barcode synthesis and conservation of nucleotide bases for biological DNA in short read sequencing. To characterize unique molecules, we utilized a 4 bp barcode (comprised of a pair of 2 bp in-line UMIs on the end of each fragment) in combination with four sequence features from paired-end reads: (i) genomic position, (ii) concise idiosyncratic gapped alignment report (CIGAR), (iii) read orientation and (iv) read number. Hybridization capture approaches have the benefit of catching a wide range of molecules with varying mapping positions, whereas amplicon-based methods capture fragments with conserved positions. By utilizing the diverse genome mapping locations of hybrid capture fragments, shorter barcodes can be employed in combination for unique molecular identification (10).

Analysis of single strand UMIs

Using our UMIs, reads derived from the same strand of a molecule were condensed into single strand consensus sequences (SSCS). First, a filter was applied to exclude reads which were unmapped, paired with an unmapped mate, or had multiple alignments. Paired reads were assigned UMIs as described above using barcode, genome mapping, CIGAR string, strand of origin, orientation, and read number information. Reads sharing the same UMIs were grouped into the same read family. Only families with 2 or more members were error suppressed and collapsed to form SSCSs as following:

- For each position across a sequence length, a Phred quality threshold of Q30 was enforced for every read (only bases with an error probability of one in a thousand or less (>Q30) were evaluated for consensus formation).
- The most frequent base at each position across all replicate reads of the same molecule was established as the consensus. The most common base was assigned if the proportion of reads representing that base was greater than or equal to the threshold required to confidently call a consensus (default cutoff 0.7—based on previous literature (9)), otherwise an *N* was assigned.
- As each SSCS represents multiple reads derived from the same strand of a unique fragment, a consensus query name was assigned to each SSCS pair. Similar to our UMIs, the pairing tag consists of a barcode along with four sequence features: (i) genome mapping ordered by coordinate, (ii) strand of origin inferred from read orientation and number, (iii) CIGAR string ordered by strand of origin and read number and (iv) read family size (number of reads supporting SSCS).

Singleton correction

We developed two approaches for Singleton Correction using the duplex nature of DNA molecules for elimination of technical artefacts. Following the formation of SSCS, singletons were grouped with their complementary SSCS for (i) Singleton Correction by SSCS. If a complementary SSCS could not be identified, single reads were paired with their complementary singleton for (ii) Singleton Correction by singletons. Through this step-wise approach, reads corresponding to the dual strands of a template molecule were used to perform Singleton Correction as following:

- UMIs were assigned to singleton and SSCS reads. For each singleton, a duplex identifier was determined by interchanging barcodes and switching the read number. If R_1 and R_2 on a positive strand had AC/GT as barcodes, their duplex barcodes would be GT/AC on the minus strand. R_1 in the forward orientation on the plus strand corresponds to R_2 in the forward orientation on the minus strand.
- Singleton Correction was achieved using either a complementary (i) SSCS or (ii) singleton corresponding to the opposite DNA strand. For each base, a Phred quality filter of Q30 was enforced to remove error prone bases. Consensus sequences were established by taking concordant bases at each position and assigning *N*s for mismatches.
- Error suppressed singleton pairs were assigned a consensus query name as described above for SSCS reads.

Recovered singleton were written to separate BAM files depending on method of correction (i.e. Singleton Correction by SSCS or Singleton Correction by singletons). They were subsequently merged with SSCS reads for downstream duplex formation.

Analysis of duplex barcodes

For optimal error suppression, duplex consensus sequences (DCS) can be established by condensing SSCSs that originated from opposite/complementary strands of a template DNA molecule. This second layer of duplex error suppression eliminates asymmetric strand artefacts. DCSs were established by preserving matched bases between reads from complementary strands. Although DCSs have the lowest rates of error, they only depict a portion of the total molecular population. To portray accurate molecular representation for variant calling, a BAM file containing all unique molecules was created by combining DCS, SSCS (without duplex pair), and uncorrected singletons.

Error analysis

We determined base substitution (error) rates using the integrated digital error suppression (iDES) tool (https://cappseq.stanford.edu/ides/download.php#bgReport) (10). BAM files were first converted to base frequency files for each genomic position using *ides-bam2freq.pl*. With the *ides-bgreport.pl*, background errors were calculated using non-reference bases <5% allele frequency with at least one read support. Error rates were determined as the number of non-reference bases over all sequenced bases within our targeted panel. We evaluated error rates at each step of error correction.

Downloaded from https://academic.oup.com/nar/article-abstract/47/15/e87/5486633 by guest on 11 March 2020

Recovery efficiency

Efficiency of consensus formation reflects the frequency of consensus sequences generated per read. This is determined by the average number of reads needed to construct a consensus sequence. For example, an efficiency rate of 10% indicates each read contributes to 0.1 of a consensus sequence, or 10 reads are needed to form a single consensus sequence.

In order to compare targeted panels of different sizes, efficiency rates were calculated using the mean target coverage (cov). GATK (v 3.6) DepthOfCoverage was used to determine mean fragment coverage per target position. Notably, we performed fragment counting as it considers overlapping reads as a single entity rather than double-counting those reads:

$$Efficiency = \frac{cov_{(DCS\ or\ SSCS)}}{cov_{uncollapsed}}$$

As DCS formation is dependent on the number of SSCS and corrected singletons, DCS recovery rates were estimated by comparing observed over expected rates:

$$Recovery_{DCS} = \frac{observed_{DCS}}{expected_{DCS}} = \frac{cov_{DCS}}{\left(\frac{Cov_{SSCS}}{2}\right)}$$

Comparison of previous UMI methods

Error rates (Supplementary Figure S1) and efficiency rates (Figure 1B) of previous methods were obtained as follows: Schmitt *et al.* error rates were reported in the text and efficiency rates were derived from Supplementary Table S1 (8), Kennedy *et al.* efficiency rates were obtained from Supplementary Table S1 (9), Schmitt *et al.* error rate and DCS efficiency rate (duplex nucleotides/(reads × (101 − 17)) were derived from Supplementary Table S2 (19), Newman *et al.* error rates were reported in text and efficiency rates were approximated from Supplementary Figure S8 (10). Efficiency rates were calculated with the equation described under 'Recovery efficiency', unless otherwise specified.

In silico downsampling

To compare hybrid capture panels of different sizes sequenced to various depths, we performed downsampling of sequencing coverage for each sample. We chose nine intervals ranging between 128 000× and 500× coverage. Each sample was reduced to the set intervals through *in silico* downsampling of paired-end reads with Samtools (v 1.3). This process was repeated 10 times for each sample across all intervals to address sampling biases. Each downsampled file was then processed through our barcoding pipeline to generate individual BAM files for singletons, SSCS, DCS, Singleton Correction by SSCS, and Singleton Correction by singletons. Corrected singletons were merged with SSCS to generate SSCS (SC) for downstream formation of DCS (SC). Efficiency for consensus formation and molecular recovery was assessed for each error suppression strategy across the broad range of coverage intervals.

Cell line dilution mutation analysis

To assess the sensitivity and specificity of UMI-based error suppression utilizing Singleton Correction, we analyzed mixed cancer cell lines diluted in 1/5 fractions across two technical replicates (Supplementary Table S4). We focused our analysis on the LargeMid library to evaluate the impact of Singleton Correction as the ultra-deeply sequenced SmallDeep contained very few singletons. For sensitivity, we evaluated single nucleotide polymorphisms corresponding to the MOLM13 cell line spiked into the dilution series. We curated a list of heterozygous (40–60% AF) and homozygous (>95% AF) single nucleotide polymorphisms (SNPs) overlapping the targeted panel that were not present in the background cell line above 1% AF. There were 222 SNPs common between technical replicates with four SNPs identified only in one of the replicates as a result of our AF thresholds. Variant calls were generated for each sample using Varscan2 (v. 2.4.2) (20). We calculated sensitivity using the list of candidate SNPs across uncollapsed and consensus reads. When assessing specificity, we bootstrapped 222 positions across the 1.2 MB targeted panel excluding sites with potential variants from both cell lines. To prevent inflation of errors, we excluded regions with poor alignability scores (obtained from ENCODE). We enumerated false positives within randomly sampled positions across 1000 iterations to evaluate specificity.

Analysis of patient samples

In our analysis, we selected samples reported to have putative driver mutations of acute myeloid leukemia (AML) (Abelson *et al.* Supplementary Table S2.1) and healthy age- and sex matched controls. We obtained 291 BAM files of peripheral blood leukocyte samples from Abelson *et al.* (15). In addition, we received 10 BAM files of umbilical cord blood samples with hybrid capture using the same 1.2 Mb leukemia panel (xGen® Acute Myeloid Leukemia Cancer Panel, IDT) sequenced to similar depths as the peripheral blood samples. UMIs were previously extracted and appended to the query name of each file. The BAM files were aligned with BWA mem to the Genome Reference Consortium Human build 37 (GRCh37).

The 10 umbilical cord blood samples were obtained from Trillium Hospital (Mississauga, Ontario, Canada) with informed consent in accordance to guidelines approved by the University Health Network Research Ethics Board. Cord blood was processed 24–48 h post-delivery. Mononuclear cells were enriched using Ficoll-Paque followed by red blood cells lysis by ammonium chloride and CD34+ selection prior to DNA extraction. 100 ng genomic DNA from the umbilical cord blood samples was used for library preparation and target capture sequencing as described above.

We processed the reads using our duplex UMI method with or without Singleton Correction. We carried out consensus efficiency and error rate as described above. To assess variant detection performance, we used 391 pre-leukemic mutations reported by Abelson *et al.* as a gold standard list (Supplementary Table S2.1, excluding one mutation that was not present in our BAM files). Files were analyzed to detect single nucleotide variants (SNVs) and small indels

Downloaded from https://academic.oup.com/nar/article-abstract/47/15/e87/5486633 by guest on 11 March 2020

using Varscan2 (20). We calculated sensitivity using the 391 pre-leukemic mutations and the 224 samples (67 samples from pre-AML individuals and 157 samples from age- and sex-matched controls) in which they were reported. Additionally, we assessed specificity using the 391 pre-leukemic mutations in all 301 samples, excluding reported mutations from Abelson *et al.* Specificity was similar when only considering a subset (77) of the 301 samples, including the 10 umbilical cord blood samples and 67 control samples not found to have pre-leukemic mutations by Abelson *et al.* (Supplementary Table S5).

RESULTS

Low efficiency consensus sequence assembly with traditional UMI methods

To assess the potential impact of Singleton Correction across diverse datasets, we first calculated important metrics of consensus sequence assembly from prior landmark studies that used traditional UMI methods (Figure 1) (8–10). This revealed critical inefficiencies in constructing SSCSs (efficiency $\leq$ 25%) and DCSs (efficiency $\leq$ 2.5%) when singletons are excluded. To confirm this using newly generated data, we performed hybrid capture NGS on cancer cell line genomic DNA with either a large panel sequenced to moderately deep coverage (LargeMid; 1.2 Mb panel, 4223× average depth) or a small panel sequenced to ultra-deep coverage (SmallDeep; 13 kb panel, 186 312× average depth). With two or more redundant reads required to construct a consensus sequence, only two-thirds of all reads in LargeMid qualified for traditional error suppression; this corresponded to a 25% SSCS efficiency rate and 2% DCS efficiency rate (Figure 1B). Since two SSCSs are required to form a DCS, theoretically we expect the maximum frequency of DCS recovery to be half of total SSCSs. However, only 15% of the expected DCSs were observed in LargeMid, and the more deeply sequenced libraries had only modest gains in DCS recovery (SmallDeep and (8–10)).

Singleton Correction augments consensus sequence assembly efficiency

We reasoned that the low consensus efficiency and DCS recovery rates observed with traditional UMI methods could be attributed to the high rate of singletons. Indeed, when Singleton Correction was applied to the LargeMid dataset, efficiency increased to 33% for SSCS and 9% for DCS. This improvement in efficiency of consensus sequence assembly resulted in a 3.6-fold increase in DCS recovery (53%) compared to traditional duplex UMI methods. In contrast to the LargeMid dataset, the vast majority (98.7%) of reads in the SmallDeep dataset contributed to consensus sequences. With so few singletons available in SmallDeep, Singleton Correction had a negligible impact on SSCS and DCS formation (Figure 1B).

High quality error suppression using singletons

We next evaluated the quality of the singletons that participated in Singleton Correction to assess their suitability for error suppression. Singleton Correction reduced the per-base error rate of singletons by 25-fold from 0.028% to 0.0011% (Figure 2A). Error rates in DCSs augmented by Singleton Correction were comparable to traditional DCSs in our datasets (Figure 2 and Supplementary Figure S3) and those from previous reports (8–10,19) (Supplementary Figure S1). This suggests high quality error suppression can be achieved using singletons, challenging the fundamental notion of requiring redundant reads for correction in traditional UMI-based methods.

Influence of sequencing depth on the impact of Singleton Correction

Since we observed a much greater effect of Singleton Correction on consensus efficiency and DCS recovery in the LargeMid dataset compared with the SmallDeep dataset, next we formally assessed the influence of sequencing depth on the impact of Singleton Correction. We performed downsampling of SmallDeep and LargeMid sequencing reads to achieve sequencing depths between 500× and 128 000× and then applied consensus assembly with or without Singleton Correction. Both SmallDeep and LargeMid displayed similar trends in consensus efficiency and recovery with a greater proportion of singletons corrected as sequencing depth increased (Figure 3A–D). Peak Singleton Correction rate occurred at 8000× depth, where 21% of singletons were corrected. This high rate was nearly maintained up to 16 000×, but at $\geq$32 000× a smaller proportion of singletons underwent Singleton Correction, suggesting an increased prevalence of duplicate reads. Analysis of SSCSs revealed consistent trends, with decreased efficiency beyond 8000× depth, indicating saturation of unique molecules with duplicate reads (Figure 3B). While Singleton Correction contributed only minor improvements to SSCS efficiency, DCS efficiency improved >2-fold at sequencing depths where singletons were abundant (Figure 3C). Furthermore, Singleton Correction enhanced DCS recovery at every coverage interval we sampled (Figure 3D). Thus, Singleton Correction ameliorated the inefficiencies of traditional UMI methods and achieved optimal recovery of DCSs across a wide range of sequencing depths. The overall impact of Singleton Correction was muted at $\geq$32 000× depth due to saturation of unique molecules in the dataset.

Increasing sensitivity with Singleton Correction

Next, we compared the detection of 222 high-confidence germline variants from the MOLM13 cell line not found in SW48 (LargeMid dataset) using duplex UMI methods with and without Singleton Correction (Supplementary Figure S4A). Using mixed cancer cell lines, we emulated varying levels of mutation variant allele frequencies at 5-fold dilutions from 100% to 0.04% MOLM13 (Supplementary Figure S4B). Across all the dilutions, uncollapsed reads had the highest sensitivity (58–100%) and the lowest specificity (62–66%). Likewise, SSCSs displayed greater sensitivity than DCSs at the expense of reduced specificity (~97%). Although the inclusion of Singleton Correction resulted in minimal difference for SSCS, DCS sensitivity increased on average by 18% without a detriment in specificity (~99.5%). At 0.04% MOLM13, the lowest dilution point, Singleton

Downloaded from https://academic.oup.com/nar/article-abstract/47/15/e87/5486633 by guest on 11 March 2020

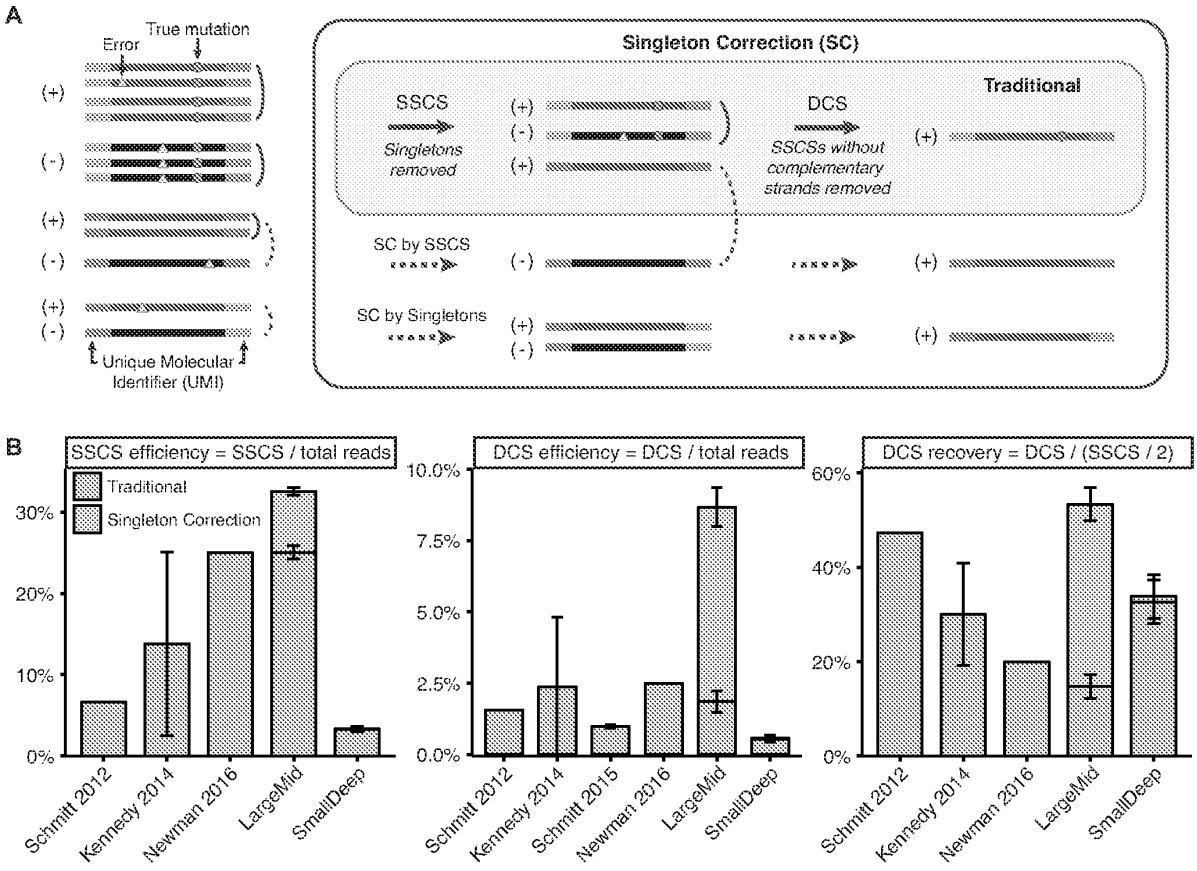


Figure 1. Singleton Correction improves traditional duplex UMI methods. (**A**) Singleton correction (SC) can be achieved through two strategies. (i) In the absence of redundant reads, singletons derived from complementary strands can be used to correct one another for *SC by Singletons*. (ii) If PCR duplicate reads are present only for one strand, they are first collapsed to form a single strand consensus sequence (SSCS). This can be subsequently used to correct the singleton of the complementary strand for *SC by SSCS*. Uncollapsed reads are not an accurate representation of molecular diversity and contain polymerase, sequencer, and oxidation errors. Traditional UMI methods of error suppression are restricted to molecules with redundant reads. Singleton Correction expands error suppression to duplex-matched singletons and enables error correction for a greater number of reads. (**B**) Comparisons of traditional duplex UMI methods from the indicated publications (8–10) and from this study (LargeMid, $n = 12$ libraries; SmallDeep, $n = 8$ libraries). Plot shows SSCS and duplex consensus sequence (DCS) efficiency and recovery for methods with traditional duplex UMI processing or with Singleton Correction. Efficiency is an assessment of over-sequencing relative to unique molecules, whereas recovery is an estimate of molecular retention after sequencing. Data are presented as mean ± S.D.

Correction produced an 8-fold increase in DCS sensitivity from 0.68% to 5.63% (Figure 4A, B). These results demonstrate the potential of Singleton Correction for high-confidence detection of low-frequency variants.

Validation of Singleton Correction performance in clinical samples

To investigate the impact of Singleton Correction in clinical samples, we next applied our method to a large study on pre-leukemia mutation detection from peripheral blood (15). Peripheral blood genomic DNA samples from 301 individuals were sequenced using the 1.2 Mb LargeMid panel to an average depth of 4746× (Figure 5A). This cohort consisted of 67 pre-leukemia patients and 224 age- and sex-matched individuals (controls) (15) as well as 10 umbilical cord blood samples that served as additional controls. Across the entire cohort, over half of all sequenced reads were unique molecules (singletons) with the remainder comprised of duplicate reads. With a traditional UMI correction method, the efficiency rate was on average 24% for SSCSs and 1% for DCSs (Figure 5B). Singleton Correction increased efficiency by 8% in SSCS and 6% in DCS and increased duplex recovery by 4-fold from 9.6% to 42%. We again observed a positive correlation between Singleton Correction and sequencing depth (Supplementary Figure S5A). Furthermore, we found consistent efficiency rates with the LargeMid cell line dilution experiment that employed a similar sequencing depth.

Singleton Correction expanded the number of reads corrected without inflating the overall error rate in patient samples. With a traditional UMI correction method, error rates were 0.01% in SSCSs and 0.0005% in DCSs. Our method reduced the error rate within singletons to 0.0007%, which

Downloaded from https://academic.oup.com/nar/article-abstract/47/15/e87/5486633 by guest on 11 March 2020

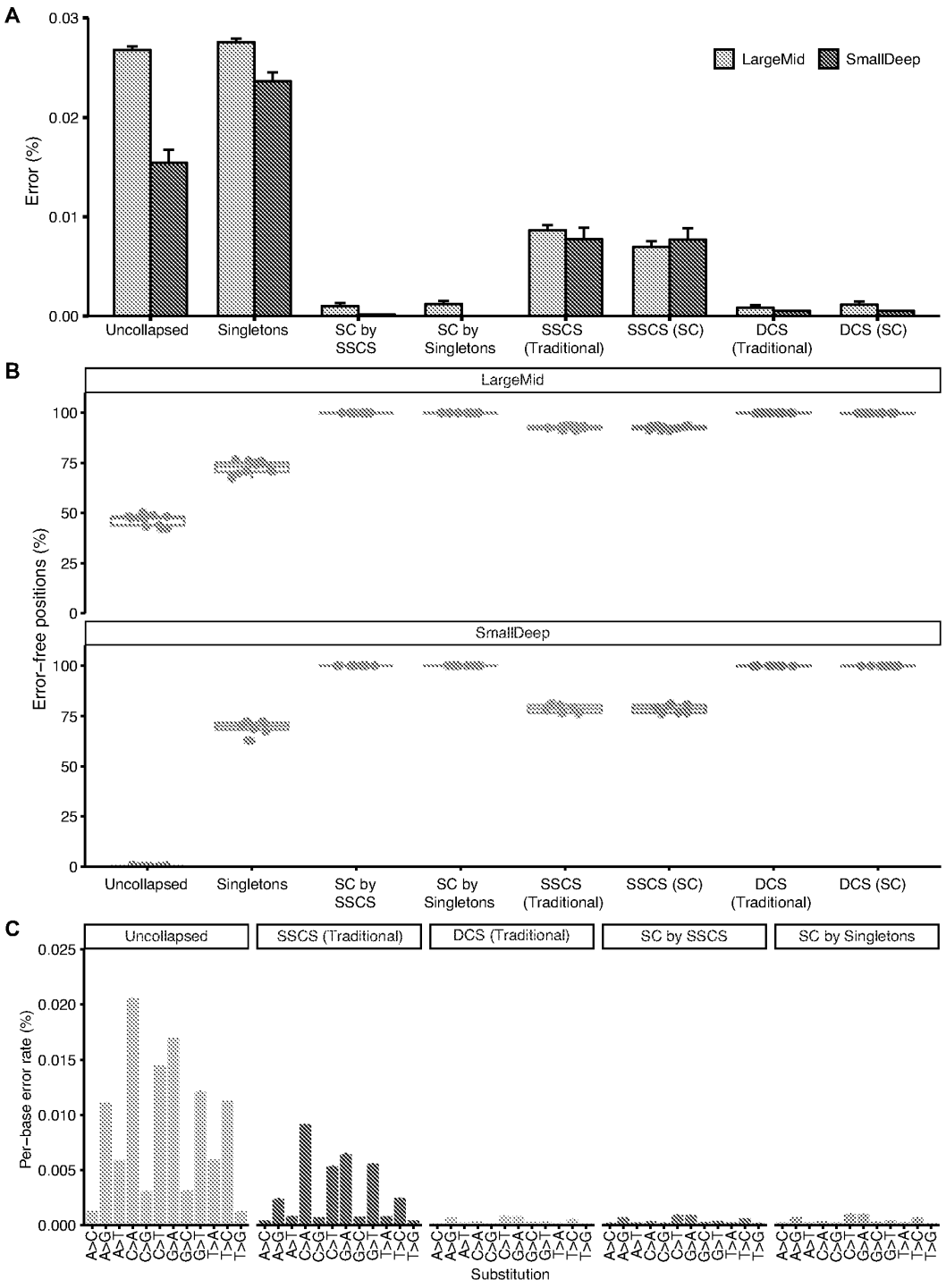


Figure 2. Corrected singletons have error profiles similar to high-quality Duplex Consensus Sequences. Comparisons between (**A**) selector-wide error rates and (**B**) error-free positions in the LargeMid (n = 12 libraries) and SmallDeep (n = 8 libraries) cell line datasets. Low rates of error and high frequency of error-free positions in Singleton Correction (SC) of SmallDeep may be attributed to the low presence of singletons within the sample and even fewer singletons corrected, as it had ultra-deep sequencing. As such, we focused our analysis on the LargeMid cell line dataset to assess (**C**) per-base error rates across different base substitutions to evaluate the error profile within corrected singletons (SC by SSCS and SC by Singletons). Data are presented as mean ± S.D.

Downloaded from https://academic.oup.com/nar/article-abstract/47/15/e87/5486633 by guest on 11 March 2020

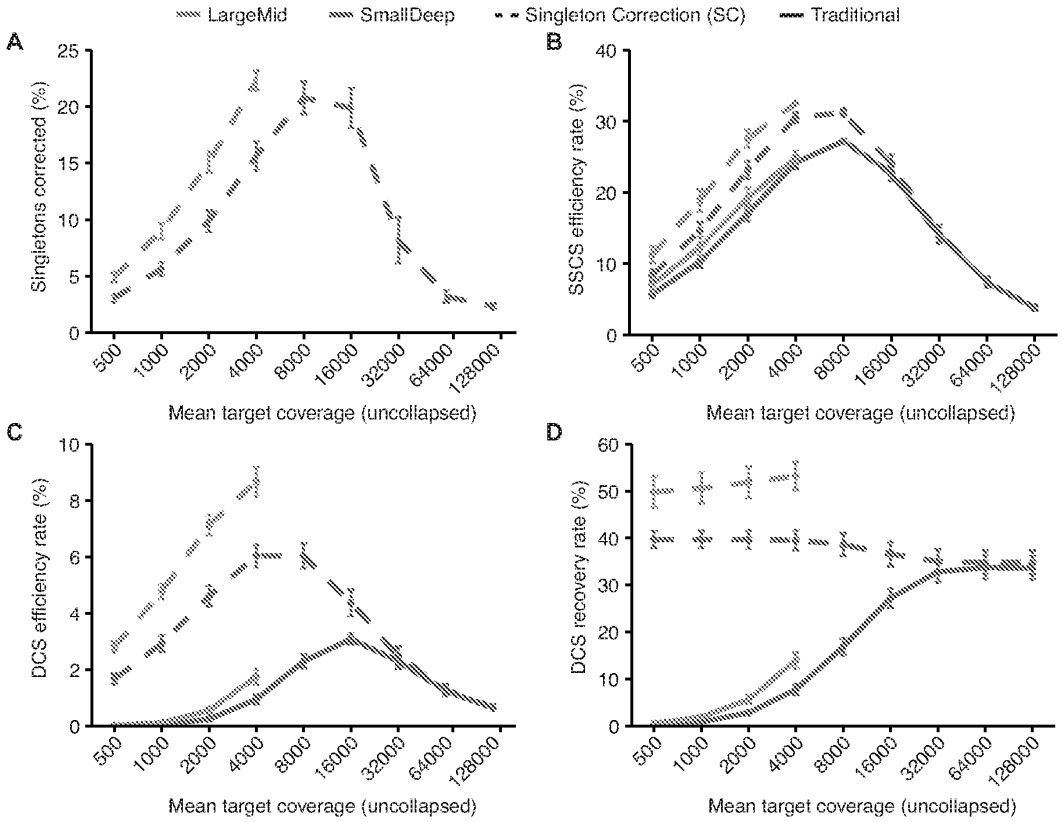


Figure 3. Singleton Correction impacts consensus formation and variant calling performance. (**A–D**) A traditional UMI approach is compared with Singleton Correction across nine coverage intervals. Plots are shown for LargeMid ($n = 12$ libraries) and SmallDeep ($n = 8$ libraries) cell line datasets that were downsampled from 128 000× to 500× depth, decreasing by half at each interval, across ten simulations. Mean target coverage is $\log_{10}$ transformed to show trends at lower range. Uncollapsed reads are defined as unprocessed reads that have not been collapsed into consensus sequences. See Methods for efficiency and recovery rate calculations.

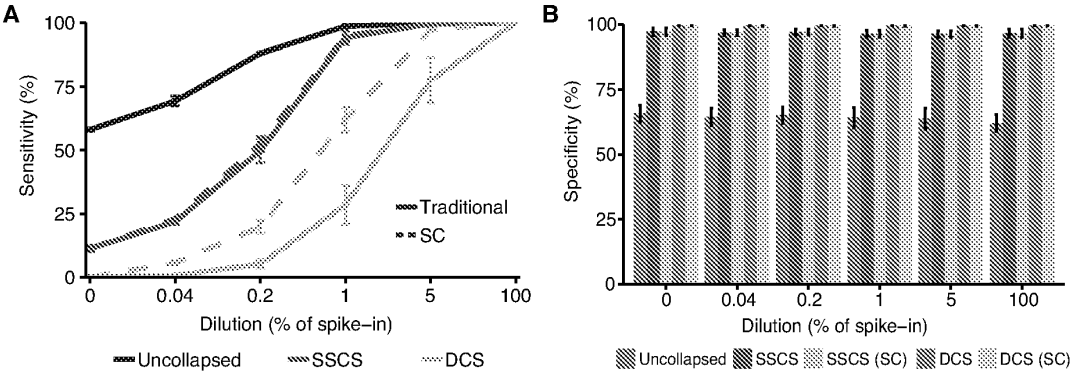


Figure 4. Variant detection sensitivity is improved with Singleton Correction. (**A**) Serial dilutions of cell lines MOLM13 spiked into SW48 (LargeMid, $n = 12$, 2 replicates per dilution). Sensitivity was assessed using 222 single nucleotide polymorphisms (SNPs) unique to the spike-in cell line that were not found in the background cell line at allele frequencies >1%. (**B**) Specificity was evaluated through bootstrapping 222 positions across the targeted panel (1000 iterations); regions of potential variants in either cell line were excluded from sampling. Data are presented as mean ± S.D.

Downloaded from https://academic.oup.com/nar/article-abstract/47/15/e87/5486633 by guest on 11 March 2020

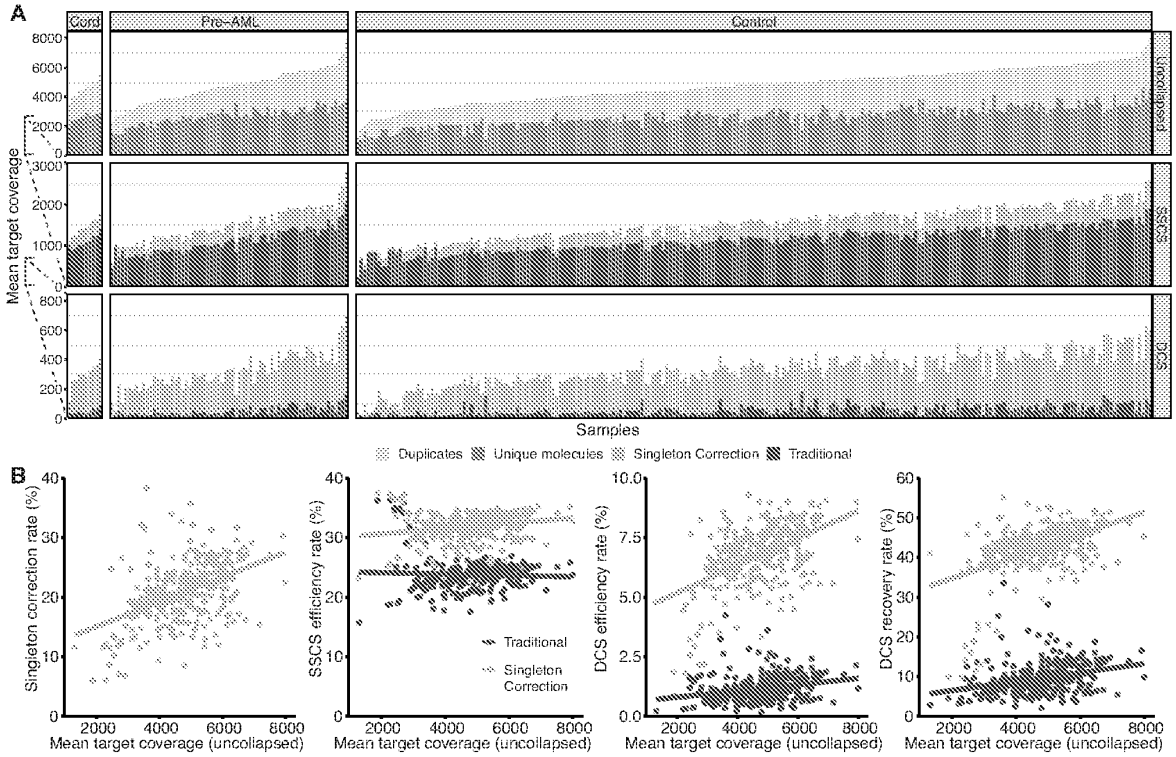


Figure 5. Performance of Singleton Correction in targeted sequencing of 301 peripheral blood leukocytes. (**A**) Plot is divided into columns corresponding to the three sample types (umbilical cord blood, blood from healthy volunteers, and blood from pre-leukemia individuals) and further separated into rows with each panel indicating the proportion of depth corresponding to uncollapsed, SSCS, and DCS reads. Each panel contrasts the reads derived from a traditional UMI strategy with reads from Singleton Correction. (**B**) Scatter plots of mean target coverage from uncollapsed reads versus Singleton Correction rate, SSCS efficiency rate, DCS efficiency rate, or DCS recovery rates; a traditional UMI approach is compared with Singleton Correction. See Methods for efficiency and recovery rate calculations and interpretations.

was comparable to the DCS error profile (Figure 6A) and to the cell line findings (Figure 2). Error substitution profiles reflected a signature of oxidative damage in reads without duplex correction (21). Notably, the characteristic imbalance between G>T and C>A substitutions was eliminated in singletons that underwent Singleton Correction (Figure 6B, Supplementary Figure S5B). These results validate our findings from cell lines and indicate that Singleton Correction is a generalizable approach that can improve the performance of UMI-based techniques in clinical samples.

Detection of low-allele-frequency variants in clinical samples

We next evaluated the effect of Singleton Correction on mutation detection accuracy in this cohort of clinical samples. Using 391 putative driver mutations of AML from Abelson *et al.* (Supplementary Figure S6A), we assessed sensitivity and specificity of duplex UMI methods. Within the different consensus data types, we evaluated performance across a range of variant count thresholds between 1 and 30; variant counts were used as opposed to variant allele fractions because of the skewed (overestimated) distribution of variant allele fractions often present within consensus sequences (Supplementary Figure S6B). Of the consensus data types, the aggregate of all unique molecules (i.e. merged DCSs, SSCSs and singletons) had the highest sensitivity but also low specificity due to inclusion of uncorrected singletons (Figure 6C). While traditional DCS had near perfect specificity without any additional filtering, sensitivity was less than half of SSCS. Singleton Correction improved sensitivity of DCS by 39% while maintaining specificity >99%.

DISCUSSION

The ability to detect low-allele-frequency variants with high-throughput sequencing technologies is dictated by the quantity of template DNA molecules, sequencing depth, and level of technical artefacts. Effective error suppression strategies are needed as errors determine the threshold at which true genetic variants can be discerned from false positives. False positive mutation calls are particularly problematic when the analysis space spans many thousands of bases, as is the case for some commercial sequencing services (32–34). Methods reported to date have not been capable of achieving high accuracy mutation detection at low thresholds without ultra-deep sequencing and/or sacrificing template DNA molecules, or without the use of large control cohorts for modeling background error rates (10). In this study, we present an enhanced UMI-based error cor-

Downloaded from https://academic.oup.com/nar/article-abstract/47/15/e87/5486633 by guest on 11 March 2020

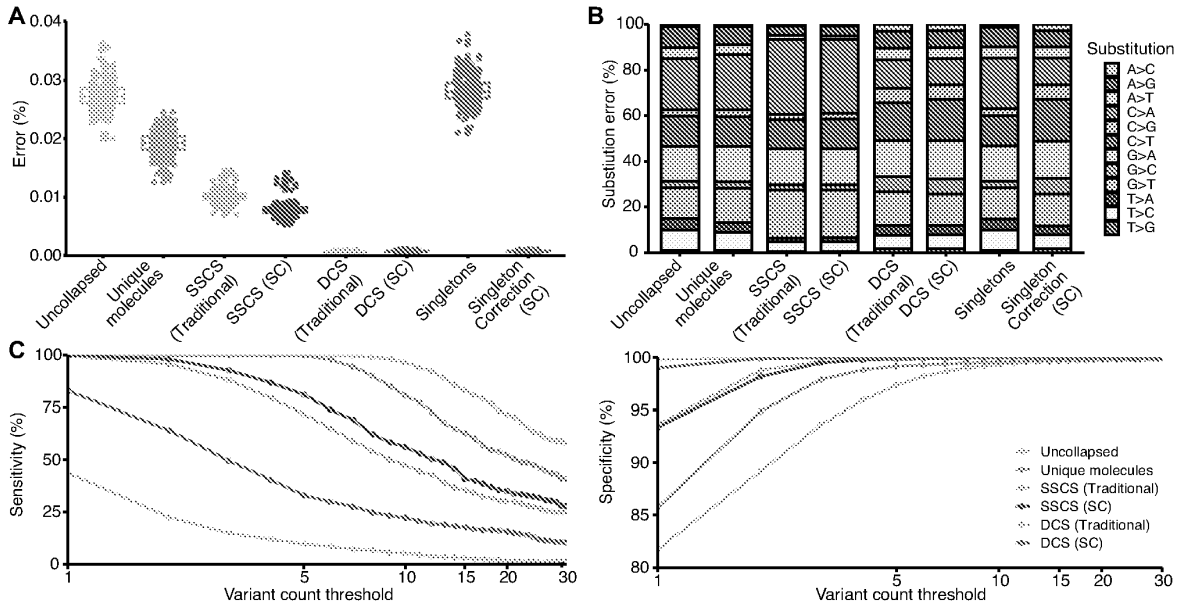


Figure 6. Error suppression and detection of low-frequency variants in clinical samples. (**A, B**) Selector-wide error rates and substitution profiles across reads with varying levels of error correction. Consensus sequences from a traditional UMI approach are compared with those derived from Singleton Correction. (**C**) Sensitivity and specificity of SNV calls at variant count thresholds from 1 to 30 for 391 putative driver mutations of acute myeloid leukaemia from the original study by Abelson *et al.* (15). Sensitivity was assessed in 224 samples in which the 391 mutations were reported. Additionally, we assessed specificity using the 391 mutations in all 301 samples, excluding exact matches from Abelson *et al.*

rection methodology aimed at addressing these important limitations.

Traditional UMI-based error correction methods require deep sequencing to achieve multiple redundant reads from the same template DNA molecule. For instance, Duplex Sequencing creates high quality DCSs with exceedingly low error rates but at the expense of inefficient processes leading to critical losses of template DNA molecules (14,15,18). Here, we demonstrate that Singleton Correction is a powerful extension for UMI-based error correction because it enables high quality error suppression across a greater number of reads. Indeed, through Singleton Correction we found that the benefits of duplex UMI methods can be extended to singletons, and therefore these reads no longer need to be categorically excluded from error suppression procedures (8–10,22). As a result, Singleton Correction results in higher consensus sequence efficiency and recovery compared to traditional methods.

Singleton Correction can be incorporated into any duplex UMI method (6,8–10,19). We used custom duplex UMI-containing adapters and sequenced on an Illumina platform, but other commercial and custom implementations of duplex UMIs for Illumina and alternative sequencing platforms would also benefit by incorporating Singleton Correction. We found the greatest benefit in hybrid capture NGS datasets with sequencing depths $\leq$16 000×. Amplicon NGS datasets would be expected to benefit less, since they generally contain fewer singletons compared with hybrid capture NGS.

Despite the gains in DCS recovery achieved using Singleton Correction compared with traditional UMI methods, still 40–50% of expected DCSs were not recovered. This could be explained by losses that are known to occur during upstream library preparation and sequencing (23), which cannot be completely overcome through over-sequencing or Singleton Correction. Further innovations in library preparation and/or sequencing methodologies may be required to realize even greater improvements in DCS recovery.

Based on our data, an important benefit of incorporating Singleton Correction is an increase in sensitivity for detecting low-frequency variants without compromising specificity. We confirmed this result using both a cell line dilution series as well as a large cohort of clinical samples that included individuals with pre-AML and/or age-related clonal hematopoiesis. High specificity is particularly important for noninvasive genotyping or screening applications (24), for instance in the setting of early detection of AML in otherwise healthy individuals (15), as false positive results may lead to unnecessary procedures and distress. Taken together, our results will inform future prospective studies in which NGS is conducted on peripheral blood or circulating DNA for early cancer detection and for other applications in oncology and precision medicine.

DATA AVAILABILITY

The dataset generated and analyzed during the current study are available in the NCBI Sequence Read Archive (SRA; https://www.ncbi.nlm.nih.gov/sra/) under access numbers SRP140497 and SRP141184. Software is

Downloaded from https://academic.oup.com/nar/article-abstract/47/15/e87/5486633 by guest on 11 March 2020

available as supplementary material and on GitHub under https://github.com/pughlab/ConsensusCruncher.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank the Genome Technologies and Genome Sequence Informatics teams at Ontario Institute for Cancer Research, the staff of the Princess Margaret Genomics Centre (www.pmgenomics.ca, Troy Ketela, Neil Winegarden, Julissa Tsao and Nick Khuu), and the Bioinformatics and High-Performance Computing Core (Carl Virtanen and Zhibin Lu) for their expertise in generating the sequencing data used in this study. We thank Marco Di Grappa for technical support and evaluation of the software. The authors gratefully acknowledge the support from the Princess Margaret Cancer Foundation.

FUNDING

Princess Margaret Cancer Foundation, Joe and Cara Finley Centre for Head & Neck Translational Research, a Canadian Cancer Society grant generously supported by the Lotte & John Hecht Memorial Foundation [704762]; Cancer Research Society [21282]; Conquer Cancer Foundation of ASCO Career Development Award (SVB). Any opinions, findings, and conclusions expressed in this material are those of the author(s) and do not necessarily reflect those of the American Society of Clinical Oncology or the Conquer Cancer Foundation. S.V.B. and T.J.P. are supported by the Gattuso-Slaight Personalized Cancer Medicine Fund at Princess Margaret Cancer Centre. Additional grant support to TJP from the Canada Research Chairs program; Canada Foundation for Innovation, Leaders Opportunity Fund [CFI #32383]; Ontario Ministry of Research and Innovation, Ontario Research Fund Small Infrastructure Program. Funding for open access charge: Canadian Cancer Society Research Institute.

Conflict of interest statement. S.V.B. is co-inventor on a patent 'Identification and use of circulating tumor markers' licensed to Roche Molecular Diagnostics.

REFERENCES

1. Abbosh,C., Birkbak,N.J. and Swanton,C. (2018) Early stage NSCLC — challenges to implementing ctDNA-based screening and MRD detection. *Nat. Rev. Clin. Oncol.*, **15**, 577–586.
2. Burgener,J.M., Rostami,A., De Carvalho,D.D. and Bratman,S.V. (2017) Cell-free DNA as a post-treatment surveillance strategy: current status. *Semin. Oncol.*, **44**, 330–346.
3. Fox,E.J., Reid-Bayliss,K.S., Emond,M.J. and Loeb,L.A. (2014) Accuracy of next generation sequencing platforms. *Gener. Seq. Appl.*, **1**, 1000106.
4. Schirmer,M., D'Amore,R., Ijaz,U.Z., Hall,N. and Quince,C. (2016) Illumina error profiles: resolving fine-scale variation in metagenomic sequencing data. *BMC Bioinformatics*, **17**, 125.
5. Kinde,I., Wu,J., Papadopoulos,N., Kinzler,K.W. and Vogelstein,B. (2011) Detection and quantification of rare mutations with massively parallel sequencing. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 9530–9535.
6. Alcaide,M., Yu,S., Davidson,J., Albuquerque,M., Bushell,K., Fornika,D., Arthur,S., Grande,B.M., McNamara,S., Tertre,M.C.D. *et al.* (2017) Targeted error-suppressed quantification of circulating tumor DNA using semi-degenerate barcoded adapters and biotinylated baits. *Sci. Rep.*, **7**, 10574.
7. Wang,K., Lai,S., Yang,X., Zhu,T., Lu,X., Wu,C.-I. and Ruan,J. (2017) Ultrasensitive and high-efficiency screen of *de novo* low-frequency mutations by o2n-seq. *Nat. Commun.*, **8**, 15335.
8. Schmitt,M.W., Kennedy,S.R., Salk,J.J., Fox,E.J., Hiatt,J.B. and Loeb,L.A. (2012) Detection of ultra-rare mutations by next-generation sequencing. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 14508–14513.
9. Kennedy,S.R., Schmitt,M.W., Fox,E.J., Kohrn,B.F., Salk,J.J., Ahn,E.H., Prindle,M.J., Kuong,K.J., Shen,J.-C., Risques,R.-A. *et al.* (2014) Detecting ultralow-frequency mutations by Duplex Sequencing. *Nat. Protoc.*, **9**, 2586–2606.
10. Newman,A.M., Lovejoy,A.F., Klass,D.M., Kurtz,D.M., Chabon,J.J., Scherer,F., Stehr,H., Liu,C.L., Bratman,S.V., Say,C. *et al.* (2016) Integrated digital error suppression for improved detection of circulating tumor DNA. *Nat. Biotechnol.*, **34**, 547–555.
11. Sloan,D.B., Broz,A.K., Sharbrough,J. and Wu,Z. (2018) Detecting rare mutations and DNA damage with sequencing-based methods. *Trends Biotechnol.*, **36**, 729–740.
12. Aird,D., Ross,M.G., Chen,W.-S., Danielsson,M., Fennell,T., Russ,C., Jaffe,D.B., Nusbaum,C. and Gnirke,A. (2011) Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol.*, **12**, R18.
13. van Dijk,E.L., Jaszczyszyn,Y. and Thermes,C. (2014) Library preparation methods for next-generation sequencing: Tone down the bias. *Exp. Cell Res.*, **322**, 12–20.
14. Kis,O., Kaedbey,R., Chow,S., Danesh,A., Dowar,M., Li,T., Li,Z., Liu,J., Mansour,M., Masih-Khan,E. *et al.* (2017) Circulating tumour DNA sequence analysis as an alternative to multiple myeloma bone marrow aspirates. *Nat. Commun.*, **8**, 15086.
15. Abelson,S., Collord,G., Ng,S.W.K., Weissbrod,O., Cohen,N.M., Niemeyer,E., Barda,N., Zuzarte,P.C., Heisler,L., Sundaravadanam,Y. *et al.* (2018) Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature*, **559**, 400–404.
16. Li,H. and Durbin,R. (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, **25**, 1754–1760.
17. DePristo,M.A., Banks,E., Poplin,R.E., Garimella,K.V., Maguire,J.R., Hartl,C., Philippakis,A.A., del Angel,G., Rivas,M., Hanna,M. *et al.* (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.*, **43**, 491–498.
18. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G. and Durbin,R. (2009) The sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
19. Schmitt,M.W., Fox,E.J., Prindle,M.J., Reid-Bayliss,K.S., True,L.D., Radich,J.P. and Loeb,L.A. (2015) Sequencing small genomic targets with high efficiency and extreme accuracy. *Nat. Methods*, **12**, 423–425.
20. Koboldt,D.C., Zhang,Q., Larson,D.E., Shen,D., McLellan,M.D., Lin,L., Miller,C.A., Mardis,E.R., Ding,L. and Wilson,R.K. (2012) VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.*, **22**, 568–576.
21. Costello,M., Pugh,T.J., Fennell,T.J., Stewart,C., Lichtenstein,L., Meldrim,J.C., Fostel,J.L., Friedrich,D.C., Perrin,D., Dionne,D. *et al.* (2013) Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation. *Nucleic Acids Res.*, **41**, e67.
22. Ahn,E. and Lee,S. (2019) Detection of Low-Frequency mutations and identification of Heat-Induced artifactual mutations using duplex sequencing. *Int. J. Mol. Sci.*, **20**, 199.
23. Fu,G.K., Xu,W., Wilhelmy,J., Mindrinos,M.N., Davis,R.W., Xiao,W. and Fodor,S.P. (2014) Molecular indexing enables quantitative targeted RNA sequencing and reveals poor efficiencies in standard library preparations. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, 1891–1896.
24. Newman,A.M., Bratman,S.V., To,J., Wynne,J.F., Eclov,N.C.W., Modlin,L.A., Liu,C.L., Neal,J.W., Wakelee,H.A., Merritt,R.E. *et al.* (2014) An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat. Med.*, **20**, 548–554.

Downloaded from https://academic.oup.com/nar/article-abstract/47/15/e87/5486633 by guest on 11 March 2020

TEMPUS EX1002.47

Electronic Patent Application Fee Transmittal

Application Number:	16575128
Filing Date:	18-Sep-2019
Title of Invention:	SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION
First Named Inventor/Applicant Name:	AmirAli TALASAZ
Filer:	Timothy A Hott/Michelle Chan
Attorney Docket Number:	42534-704.313

Filed as Large Entity

Filing Fees for Utility under 35 USC 111(a)

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Miscellaneous:				
SUBMISSION- INFORMATION DISCLOSURE STMT	1806	1	240	240
		Total in USD ($)		**240**

Electronic Acknowledgement Receipt

EFS ID:	39346841
Application Number:	16575128
International Application Number:	
Confirmation Number:	9440
Title of Invention:	SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION
First Named Inventor/Applicant Name:	AmirAli TALASAZ
Customer Number:	115823
Filer:	Timothy A Hott/Michelle Chan
Filer Authorized By:	Timothy A Hott
Attorney Docket Number:	42534-704.313
Receipt Date:	04-MAY-2020
Filing Date:	18-SEP-2019
Time Stamp:	18:36:16
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	DA
Payment was successfully received in RAM	$240
RAM confirmation Number	E202054I36331793
Deposit Account	
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		2020-05-04_GH0002US_CON14_OA1Response.pdf	160990 bd1bb3a428982294857309fbd59b497508c9126c	yes	10
	Multipart Description/PDF files in .zip description				
	Document Description		**Start**	**End**	
	Amendment/Req. Reconsideration-After Non-Final Reject		1	1	
	Specification		2	2	
	Claims		3	7	
	Applicant Arguments/Remarks Made in an Amendment		8	10	
Warnings:					
Information:					
2	Transmittal Letter	2020-05-04_GH0002US-CON14_IDSTrans.pdf	144436 bf60efa19b22f6296e5e3b87319dbf96dbecb328	no	4
Warnings:					
Information:					
3	Information Disclosure Statement (IDS) Form (SB08)	2020-05-04_GH0002US-CON14_SB08.pdf	1055111 1438210474d7af4b8cd12fc3b4657c07f71788d0	no	6
Warnings:					
Information:					
4	Non Patent Literature	ADEY_GenBiol_2010_R119.pdf	3747562 0a26979c6b8c13dd9d53cd4bf420c443980b33a8	no	17
Warnings:					
Information:					

5	Non Patent Literature	GH_FMI_PGDx_ReportRecommend_2020-01-07.pdf	368170 06bb538754e81ef546eaefebb5bb79e21324d97b	no	24
Warnings:					
Information:					
6	Non Patent Literature	GH_FMI_PGDx_MemorandumOrderReport_2020-03-23.pdf	184494 d272b4e0ea6577ac72671034fde4c6b9bf5e157d	no	10
Warnings:					
Information:					
7	Non Patent Literature	IPR2019-00634_Response_2019-12-04.pdf	551659 69dd360c33dd58dfb9859c1656ae63cee8485959	no	60
Warnings:					
Information:					
8	Non Patent Literature	IPR2019-00634_OppositionToAmend_2020-03-03.pdf	302829 871e4cc94c077e99918cc08c2c3b98aa8851bdbe	no	37
Warnings:					
Information:					
9	Non Patent Literature	IPR2019-00634_PetitionersReply_2020-03-03.pdf	513272 d99e8cba582414bd8c63e77dfc1c2f04836a4acc	no	40
Warnings:					
Information:					
10	Non Patent Literature	IPR2019-00652_Response_2019-12-04.pdf	494190 2d3c8c664bf1c8ace9a981b77d9b8eaf2acb7116	no	71
Warnings:					
Information:					
11	Non Patent Literature	IPR2019-00652_PetitionersReply_2020-03-03.pdf	1130839 73db1007bfe72f85ba0ff5ce11cd269d04f0e4b6	no	46
Warnings:					
Information:					

12	Non Patent Literature	KAMPS-HUGHES_PLOSOne_2018.pdf	5707894 ea4d5cf7239d69ce8d4cdfae92b25404809533e1	no	18
Warnings:					
Information:					
13	Non Patent Literature	LENNON_GenMed_2016_112.pdf	689512 67e86d14248058462070 40bcbbef63ab7e600b00	no	10
Warnings:					
Information:					
14	Non Patent Literature	MISHRA_MCB_2016_1050-1063.pdf	1216913 5f63e365b9b13c1c13c502d1372bc57258b43782	no	14
Warnings:					
Information:					
15	Non Patent Literature	MOENCH_CancTherAdv_2019.pdf	640654 cbedcb9c9350ebc56dce20b0e103f70bf26863c3	no	3
Warnings:					
Information:					
16	Non Patent Literature	NEWMAN_NatMed_2014_547-555.pdf	1377423 69ec2aa88500b798bf011f81a5bfe5b8de41be93	no	9
Warnings:					
Information:					
17	Non Patent Literature	ODEGAARD_ClinCancRes_2018_3539-3549.pdf	1062206 cb51a9f4833c005db46554e71fa08573f616a364	no	11
Warnings:					
Information:					
18	Non Patent Literature	OA_16601168_2020-02-18.pdf	4018735 cde09ae13240737208c8f6f88a1cc377634499e3	no	105
Warnings:					
Information:					

19	Non Patent Literature	OA_16672267_2020-03-03.pdf	4023116 d3dc04b0f208e979ea063dc71b34bdbc9ba7db01	no	105
Warnings:					
Information:					
20	Non Patent Literature	OA_16714579_2020-03-03.pdf	4039894 661125852728 4a9bbe173157cb72ff6274960cf1	no	106
Warnings:					
Information:					
21	Non Patent Literature	OU_ACSO_2018_978-997.pdf	664167 dde74d5088588343da8b04ede4fcf58466ae171e	no	20
Warnings:					
Information:					
22	Non Patent Literature	Sathirapongsasuti_BioInformatics_2011.pdf	484289 78ff01d10aa50d956f5f7a931c727355df943f9e	no	7
Warnings:					
Information:					
23	Non Patent Literature	TRAPNELL_NatBiotech_2009_455-457.pdf	183342 8d9e4daf2d810f2b71651c6febd734337a82b52a	no	3
Warnings:					
Information:					
24	Non Patent Literature	VAN_LOO_PNAS_2010_16910-16915.pdf	2450725 77581bb7574de0bfab9474778d4b67be852c428a	no	6
Warnings:					
Information:					
25	Non Patent Literature	WANG_NAR_2019.pdf	735593 e5227ff4b8dcb4d907196ad89e55c94d3b43e356	no	11
Warnings:					
Information:					

26	Fee Worksheet (SB06)	fee-info.pdf	30591 07c781c76298d315fa1d965070155dcb5a8b9a90	no	2
Warnings:					
Information:					
			Total Files Size (in bytes):	35978606	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PTO/SB/06 (09-11)
Approved for use through 1/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PATENT APPLICATION FEE DETERMINATION RECORD Substitute for Form PTO-875	Application or Docket Number 16/575,128	Filing Date 09/18/2019	☐ To be Mailed

ENTITY: ☑ LARGE ☐ SMALL ☐ MICRO

APPLICATION AS FILED - PART I

FOR	(Column 1) NUMBER FILED	(Column 2) NUMBER EXTRA		RATE ($)	FEE ($)
☐ BASIC FEE (37 CFR 1.16(a), (b), or (c))	N/A	N/A		N/A	
☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m))	N/A	N/A		N/A	
☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q))	N/A	N/A		N/A	
TOTAL CLAIMS (37 CFR 1.16(i))	minus 20 =	*		x $100 =	
INDEPENDENT CLAIMS (37 CFR 1.16(h))	minus 3 =	*		x $460 =	
☐ APPLICATION SIZE FEE (37 CFR 1.16(s))	If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).				
☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j))					
* If the difference in column 1 is less than zero, enter "0" in column 2.				TOTAL	

APPLICATION AS AMENDED - PART II

AMENDMENT	05/04/2020	(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA		RATE ($)	ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	* 30	Minus	** 84	= 0		x $100 =	0
	Independent (37 CFR 1.16(h))	* 3	Minus	*** 3	= 0		x $460 =	0
	☐ Application Size Fee (37 CFR 1.16(s))							
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							
							TOTAL ADD'L FEE	0

AMENDMENT		(Column 1) CLAIMS REMAINING AFTER AMENDMENT		(Column 2) HIGHEST NUMBER PREVIOUSLY PAID FOR	(Column 3) PRESENT EXTRA		RATE ($)	ADDITIONAL FEE ($)
	Total (37 CFR 1.16(i))	*	Minus	**	=		x $0 =	
	Independent (37 CFR 1.16(h))	*	Minus	***	=		x $0 =	
	☐ Application Size Fee (37 CFR 1.16(s))							
	☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j))							
							TOTAL ADD'L FEE	

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.	LIE
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".	/ROCHELLE C. GETER/
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".	
The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.	

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

NOTICE OF ALLOWANCE AND FEE(S) DUE

115823 7590 07/10/2020
Wilson Sonsini Goodrich & Rosati / Guardant Health
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER
HORLICK, KENNETH R

ART UNIT	PAPER NUMBER
1637	

DATE MAILED: 07/10/2020

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/575,128	09/18/2019	AmirAli TALASAZ	42534-704.313	9440

TITLE OF INVENTION: SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	UNDISCOUNTED	$1000	$0.00	$0.00	$1000	10/13/2020

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.

THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE DOES NOT REFLECT A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE IN THIS APPLICATION. IF AN ISSUE FEE HAS PREVIOUSLY BEEN PAID IN THIS APPLICATION (AS SHOWN ABOVE), THE RETURN OF PART B OF THIS FORM WILL BE CONSIDERED A REQUEST TO REAPPLY THE PREVIOUSLY PAID ISSUE FEE TOWARD THE ISSUE FEE NOW DUE.

HOW TO REPLY TO THIS NOTICE:

I. Review the ENTITY STATUS shown above. If the ENTITY STATUS is shown as SMALL or MICRO, verify whether entitlement to that entity status still applies.

If the ENTITY STATUS is the same as shown above, pay the TOTAL FEE(S) DUE shown above.

If the ENTITY STATUS is changed from that shown above, on PART B - FEE(S) TRANSMITTAL, complete section number 5 titled "Change in Entity Status (from status indicated above)".

For purposes of this notice, small entity fees are 1/2 the amount of undiscounted fees, and micro entity fees are 1/2 the amount of small entity fees.

II. PART B - FEE(S) TRANSMITTAL, or its equivalent, must be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed and an extra copy of the form should be submitted. If an equivalent of Part B is filed, a request to reapply a previously paid issue fee must be clearly made, and delays in processing may occur due to the difficulty in recognizing the paper as an equivalent of Part B.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Maintenance fees are due in utility patents issuing on applications filed on or after Dec. 12, 1980. It is patentee's responsibility to ensure timely payment of maintenance fees when due. More information is available at www.uspto.gov/PatentMaintenanceFees.

TEMPUS EX1002.57

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), by mail or fax, or via EFS-Web.

By mail, send to: Mail Stop ISSUE FEE
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

By fax, send to: (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

115823 7590 07/10/2020
Wilson Sonsini Goodrich & Rosati / Guardant Health
650 Page Mill Road
Palo Alto, CA 94304

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission

I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being transmitted to the USPTO via EFS-Web or by facsimile to (571) 273-2885, on the date below.

(Typed or printed name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/575,128	09/18/2019	AmirAli TALASAZ	42534-704.313	9440

TITLE OF INVENTION: SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	UNDISCOUNTED	$1000	$0.00	$0.00	$1000	10/13/2020

EXAMINER	ART UNIT	CLASS-SUBCLASS
HORLICK, KENNETH R	1637	435-006120

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).

❑ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.

❑ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-09 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) The names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) The name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1_____
2_____
3_____

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document must have been previously recorded, or filed for recordation, as set forth in 37 CFR 3.11 and 37 CFR 3.81(a). Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE (B) RESIDENCE: (CITY and STATE OR COUNTRY)

Please check the appropriate assignee category or categories (will not be printed on the patent) : ❑ Individual ❑ Corporation or other private group entity ❑ Government

4a. Fees submitted: ❑Issue Fee ❑Publication Fee (if required) ❑Advance Order - # of Copies _____

4b. Method of Payment: *(Please first reapply any previously paid fee shown above)*

❑ Electronic Payment via EFS-Web ❑ Enclosed check ❑ Non-electronic payment by credit card (Attach form PTO-2038)

❑ The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment to Deposit Account No. _____

5. **Change in Entity Status** (from status indicated above)

❑ Applicant certifying micro entity status. See 37 CFR 1.29

❑ Applicant asserting small entity status. See 37 CFR 1.27

❑ Applicant changing to regular undiscounted fee status.

NOTE: Absent a valid certification of Micro Entity Status (see forms PTO/SB/15A and 15B), issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment.
NOTE: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.
NOTE: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

NOTE: This form must be signed in accordance with 37 CFR 1.31 and 1.33. See 37 CFR 1.4 for signature requirements and certifications.

Authorized Signature _____ Date _____

Typed or printed name _____ Registration No. _____

TEMPUS EX1002.58

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/575,128	09/18/2019	AmirAli TALASAZ	42534-704.313	9440

115823 7590 07/10/2020

Wilson Sonsini Goodrich & Rosati / Guardant Health
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER	
HORLICK, KENNETH R	

ART UNIT	PAPER NUMBER
1637	

DATE MAILED: 07/10/2020

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(Applications filed on or after May 29, 2000)

The Office has discontinued providing a Patent Term Adjustment (PTA) calculation with the Notice of Allowance.

Section 1(h)(2) of the AIA Technical Corrections Act amended 35 U.S.C. 154(b)(3)(B)(i) to eliminate the requirement that the Office provide a patent term adjustment determination with the notice of allowance. See Revisions to Patent Term Adjustment, 78 Fed. Reg. 19416, 19417 (Apr. 1, 2013). Therefore, the Office is no longer providing an initial patent term adjustment determination with the notice of allowance. The Office will continue to provide a patent term adjustment determination with the Issue Notification Letter that is mailed to applicant approximately three weeks prior to the issue date of the patent, and will include the patent term adjustment on the patent. Any request for reconsideration of the patent term adjustment determination (or reinstatement of patent term adjustment) should follow the process outlined in 37 CFR 1.705.

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at 1-(888)-786-0101 or (571)-272-4200.

TEMPUS EX1002.59

OMB Clearance and PRA Burden Statement for PTOL-85 Part B

The Paperwork Reduction Act (PRA) of 1995 requires Federal agencies to obtain Office of Management and Budget approval before requesting most types of information from the public. When OMB approves an agency request to collect information from the public, OMB (i) provides a valid OMB Control Number and expiration date for the agency to display on the instrument that will be used to collect the information and (ii) requires the agency to inform the public about the OMB Control Number's legal significance in accordance with 5 CFR 1320.5(b).

The information collected by PTOL-85 Part B is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 30 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450. Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Notice of Allowability	Application No. 16/575,128	Applicant(s) TALASAZ, AmirAli	
	Examiner KENNETH R HORLICK	Art Unit 1637	AIA (FITF) Status Yes

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☑ This communication is responsive to the response filed 05/04/20.
   ☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2. ☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

3. ☑ The allowed claim(s) is/are 31-61 . As a result of the allowed claim(s), you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see **http://www.uspto.gov/patents/init_events/pph/index.jsp** or send an inquiry to **PPHfeedback@uspto.gov.**

4. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

   Certified copies:

   a) ☐ All b) ☐ Some *c) ☐ None of the:

   1. ☐ Certified copies of the priority documents have been received.
   2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
   3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____ .

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____ .

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).

6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☑ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date 5/4/20.
3. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material _____.
4. ☐ Interview Summary (PTO-413), Paper No./Mail Date. _____.
5. ☑ Examiner's Amendment/Comment
6. ☐ Examiner's Statement of Reasons for Allowance
7. ☐ Other _____.

/KENNETH R HORLICK/
Primary Examiner, Art Unit 1637

EXAMINER'S COMMENTS

1. The present application, filed on or after March 16, 2013, is being examined under the first inventor to file provisions of the AIA.

2. The terminal disclaimer filed on 05/04/20 disclaiming the terminal portion of any patent granted on this application which would extend beyond the expiration date of any patent issuing from U.S. application '099, '079, '633, or '653, or of patent '992, '678, or 810, has been reviewed and is accepted. The terminal disclaimer has been recorded.

3. Any inquiry concerning this communication or earlier communications from the examiner should be directed to KENNETH R HORLICK whose telephone number is (571)272-0784. The examiner can normally be reached on Mon. - Thurs. 8:30 - 6:30.

Examiner interviews are available via telephone, in-person, and video conferencing using a USPTO supplied web-based collaboration tool. To schedule an interview, applicant is encouraged to use the USPTO Automated Interview Request (AIR) at http://www.uspto.gov/interviewpractice.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary Benzion can be reached on 571-272-0782. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see https://ppair-my.uspto.gov/pair/PrivatePair. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

06/29/20

/KENNETH R HORLICK/
Primary Examiner, Art Unit 1637

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	16/575,128	TALASAZ, AmirAli
	Examiner	**Art Unit**
	KENNETH R HORLICK	1637

CPC							
Symbol						**Type**	**Version**
C12Q	/	1	/	6886		F	2013-01-01
G16B	/	30	/	00		I	2019-02-01
C12Q	/	1	/	6869		I	2013-01-01
C12N	/	15	/	1065		I	2013-01-01
C12Q	/	1	/	6806		I	2013-01-01
C12Q	/	1	/	6874		I	2013-01-01
C12Q	/	2600	/	118		A	2013-01-01
C12Q	/	2600	/	156		A	2013-01-01
C12Q	/	2600	/	16		A	2013-01-01
C12Q	/	2600	/	158		A	2013-01-01

CPC Combination Sets								
Symbol					**Type**	**Set**	**Ranking**	**Version**
C12Q	/	1	/	6869	A	1	1	2013-01-01
C12Q	/	2537	/	165	A	1	2	2013-01-01
C12Q	/	2545	/	114	A	1	3	2013-01-01

NONE		**Total Claims Allowed:**	
(Assistant Examiner)	(Date)	31	
/KENNETH R HORLICK/ Primary Examiner, Art Unit 1637	29 June 2020	O.G. Print Claim(s)	O.G. Print Figure
(Primary Examiner)	(Date)	1	NONE

U.S. Patent and Trademark Office — Part of Paper No.: 20200629

TEMPUS EX1002.63

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	16/575,128	TALASAZ, AmirAli
	Examiner	**Art Unit**
	KENNETH R HORLICK	1637

INTERNATIONAL CLASSIFICATION					
CLAIMED					
C12P	/	19	/	34	

NON-CLAIMED					
	/		/		

US ORIGINAL CLASSIFICATION	
CLASS	**SUBCLASS**
435	6.12

CROSS REFERENCES(S)						
CLASS	**SUBCLASS (ONE SUBCLASS PER BLOCK)**					
435	91.2					

NONE (Assistant Examiner)	(Date)	**Total Claims Allowed:** 31	
/KENNETH R HORLICK/ Primary Examiner, Art Unit 1637 (Primary Examiner)	29 June 2020 (Date)	O.G. Print Claim(s) 1	O.G. Print Figure NONE

U.S. Patent and Trademark Office — Part of Paper No.: 20200629

TEMPUS EX1002.64

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	16/575,128	TALASAZ, AmirAli
	Examiner	**Art Unit**
	KENNETH R HORLICK	1637

☑ Claims renumbered in the same order as presented by applicant ☐ CPA ☑ T.D. ☐ R.1.47

CLAIMS

Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original
1	31	10	40	19	49	28	58								
2	32	11	41	20	50	29	59								
3	33	12	42	21	51	30	60								
4	34	13	43	22	52	31	61								
5	35	14	44	23	53										
6	36	15	45	24	54										
7	37	16	46	25	55										
8	38	17	47	26	56										
9	39	18	48	27	57										

NONE		**Total Claims Allowed:**	
(Assistant Examiner)	(Date)	31	
/KENNETH R HORLICK/ Primary Examiner, Art Unit 1637	29 June 2020	O.G. Print Claim(s)	O.G. Print Figure
(Primary Examiner)	(Date)	1	NONE

U.S. Patent and Trademark Office — Part of Paper No.: 20200629

TEMPUS EX1002.65

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	16/575,128	TALASAZ, AmirAli
	Examiner	**Art Unit**
	KENNETH R HORLICK	1637

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

CLAIMS

☑ Claims renumbered in the same order as presented by applicant ☐ CPA ☑ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	02/14/2020	06/29/2020							
1	31	✓	=							
2	32	✓	=							
3	33	✓	=							
4	34	✓	=							
5	35	✓	=							
6	36	✓	=							
7	37	✓	=							
8	38	✓	=							
9	39	✓	=							
10	40	✓	=							
11	41	✓	=							
12	42	✓	=							
13	43	✓	=							
14	44	✓	=							
15	45	✓	=							
16	46	✓	=							
17	47	✓	=							
18	48	✓	=							
19	49	✓	=							
20	50	✓	=							
21	51	✓	=							
22	52	✓	=							
23	53	✓	=							
24	54	✓	=							
25	55	✓	=							
26	56	✓	=							
27	57	✓	=							
28	58	✓	=							
29	59	✓	=							
30	60	✓	=							
31	61		=							

Search Notes	Application/Control No.	Applicant(s)/Patent Under Reexamination
	16/575,128	TALASAZ, AmirAli
	Examiner	Art Unit
	KENNETH R HORLICK	1637

CPC - Searched*		
Symbol	Date	Examiner

CPC Combination Sets - Searched*		
Symbol	Date	Examiner

US Classification - Searched*			
Class	Subclass	Date	Examiner

* See search history printout included with this form or the SEARCH NOTES box below to determine the scope of the search.

Search Notes		
Search Notes	Date	Examiner
inventor name search	02/11/2020	KH
updated parent searches in USPAT and PGPUB	02/11/2020	KH
reviewed parent applications and references therein	02/11/2020	KH
updated in USPAT and PGPUB	06/29/2020	KH

Interference Search			
US Class/CPC Symbol	US Subclass/CPC Group	Date	Examiner
NONE	NONE	06/29/2020	KH

/KENNETH R HORLICK/ Primary Examiner, Art Unit 1637	

TEMPUS EX1002.67

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

BIB DATA SHEET

CONFIRMATION NO. 9440

SERIAL NUMBER	FILING or 371(c) DATE RULE	CLASS	GROUP ART UNIT	ATTORNEY DOCKET NO.
16/575,128	09/18/2019	435	1637	42534-704.313

APPLICANTS
GUARDANT HEALTH, INC., Redwood City, CA;

INVENTORS
AmirAli TALASAZ, Atherton, CA;

**** CONTINUING DATA ***************************
This application is a CON of 16/283,635 02/22/2019 PAT 10494678
which is a CON of 15/872,831 01/16/2018 PAT 10457995
which is a CON of 15/828,099 11/30/2017
which is a CON of 15/467,570 03/23/2017 PAT 9840743
which is a CON of 14/425,189 03/02/2015 PAT 10041127
which is a 371 of PCT/US2013/058061 09/04/2013
which claims benefit of 61/845,987 07/13/2013
and claims benefit of 61/793,997 03/15/2013
and claims benefit of 61/704,400 09/21/2012
and claims benefit of 61/696,734 09/04/2012

**** FOREIGN APPLICATIONS ***************************

**** IF REQUIRED, FOREIGN FILING LICENSE GRANTED ****
10/02/2019

Foreign Priority claimed ☐ Yes ☑ No 35 USC 119(a-d) conditions met ☐ Yes ☑ No Verified and Acknowledged /KENNETH R HORLICK/ Examiner's Signature	☐ Met after Allowance Initials	STATE OR COUNTRY	SHEETS DRAWINGS	TOTAL CLAIMS	INDEPENDENT CLAIMS
		CA	16	30	2

ADDRESS
Wilson Sonsini Goodrich & Rosati / Guardant Health
650 Page Mill Road
Palo Alto, CA 94304
UNITED STATES

TITLE
SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION

FILING FEE RECEIVED		
2720	FEES: Authority has been given in Paper No._____ to charge/credit DEPOSIT ACCOUNT No._____ for following:	☐ All Fees ☐ 1.16 Fees (Filing) ☐ 1.17 Fees (Processing Ext. of time) ☐ 1.18 Fees (Issue) ☐ Other _____ ☐ Credit

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (02-18)
Approved for use through 11/30/2020. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT
(Not for submission under 37 CFR 1.99)

Application Number	16575128
Filing Date	2019-09-18
First Named Inventor	AmirAli TALASAZ
Art Unit	
Examiner Name	
Attorney Docket Number	42534-704.313

U.S.PATENTS Remove

Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	10287631	B2	2019-05-14	Salk et al.	Entire Document
	2	10370713	B2	2019-08-06	Salk et al.	Entire Document
	3	10385393	B2	2019-08-20	Salk et al.	Entire Document
	4	10388403	B2	2019-08-20	Rava et al.	Entire Document
	5	10604804	B2	2020-03-31	Salk et al.	Entire Document

If you wish to add additional U.S. Patent citation information please click the Add button. Add

U.S.PATENT APPLICATION PUBLICATIONS Remove

Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1	20190271040	A1	2019-09-05	Salk et al.	Entire Document

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /K.R.H/

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	16575128
	Filing Date	2019-09-18
	First Named Inventor	AmirAli TALASAZ
	Art Unit	
	Examiner Name	
	Attorney Docket Number	42534-704.313

	2	20190292597	A1	2019-09-26	Salk et al.	Entire Document
	3	20190338358	A1	2019-11-07	Salk et al.	Entire Document
	4	20190352714	A1	2019-11-21	Salk et al.	Entire Document

If you wish to add additional U.S. Published Application citation information please click the Add button. Add

FOREIGN PATENT DOCUMENTS Remove

Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							

If you wish to add additional Foreign Patent Document citation information please click the Add button Add

NON-PATENT LITERATURE DOCUMENTS Remove

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]
	1	ADEY, et al. "Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition" Genome Biology (2010) 11:R119	
	2	Guardant Health, Inc. v. Foundation Medicine, Inc. Civil Action No. 17-1616-LPS-CJB and Guardant Health, Inc. v. Personal Genome Diagnostics Civil Action No. 17-1623-LPS-CJB, Report and Recommendation dated January 7, 2020	
	3	Guardant Health, Inc. v. Foundation Medicine, Inc. Civil Action No. 17-1616-LPS-CJB and Guardant Health, Inc. v. Personal Genome Diagnostics Civil Action No. 17-1623-LPS-CJB, Memorandum Order dated March 23, 2020	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /K.R.H/

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	16575128
	Filing Date	2019-09-18
	First Named Inventor	AmirAli TALASAZ
	Art Unit	
	Examiner Name	
	Attorney Docket Number	42534-704.313

	4	IPR2019-00634, Patent Owner Response, dated December 4, 2019	
	5	IPR2019-00634, Petitioner's Opposition to Patent Owner's Motion to Amend, dated March 3, 2020	
	6	IPR2019-00634, Petitioner's Reply, dated March 3, 2020	
	7	IPR2019-00652, Patent Owner Response, dated December 4, 2019	
	8	IPR2019-00652, Petitioner's Reply to Patent Owner Response, dated March 3, 2020	
	9	KAMPS-HUGHES, N. et al. "ERASE-Seq: Leveraging replicate measurements to enhance ultralow frequency variant detection in NGS data"PLOS One (2018)	
	10	LENNON, N.J. et al. "Technological considerations for genome-guided diagnosis and management of cancer" Gen Med (2016) 8:112	
	11	MISHRA, S. et al. "Different Facets of Copy Number Changes: Permanent, Transient, and Adaptive" Mol Cell Biol (2016) 36(7):1050-1063	
	12	MOENCH, S. "Genomic Profiling Using Guardant 360 Cell-Free DNA-Based Assay vs Tumor-Based Genotyping Assays in Advanced NSCLC, CANCER THERAPY ADVISOR (Feb. 28, 2019), https://www.cancertherapyadvisor.com/home/news/conferencecoverage/american-association-for-cancer-research-aacr/aacr-2019/genomic-profiling-using-guardant-360-cell-free-dna-based assay-vs-tumor-based-genotyping-assays-in-advanced-nsclc/ (lastaccessed Nov. 30, 2019)	
	13	NEWMAN, A. et al. "Integrated digital error suppression for improved detection of circulating tumor DNA" Nature Biotech (2016) 34(5):547-555	
	14	ODEGAARD, J.I. et al. "Validation of a Plasma-Based Comprehensive Cancer Genotyping Assay Utilizing Orthogonal Tissue- and Plasma-Based Methodologies" Clin Canc Res (2018) 24(15):3539-3549	

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /K.R.H/

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	16575128
	Filing Date	2019-09-18
	First Named Inventor	AmirAli TALASAZ
	Art Unit	
	Examiner Name	
	Attorney Docket Number	42534-704.313

	15	Office action dated 02/18/2020 for US Application No. 16/601,168.	
	16	Office action dated 03/03/2020 for US Application No. 16/672,267.	
	17	Office action dated 03/03/2020 for US Application No. 16/714,579.	
	18	OU, SHI et al. "Liquid Biopsy to Identify Actionable Genomic Alterations" Am Soc Clin Onc (2018) 978	
	19	SATHIRAPONGSASUTI, J.F. et al. "Exome sequencing-based copy-number variation and loss of heterozygosity detection: ExomeCNV" BioInformatics (2011) 27(19):2648-2654	
	20	TRAPNELL, C. et al. "How to map billions of short reads onto genomes" Nature Biotech (2009) 27(5):455-457	
	21	VAN LOO, P. et al. "Allele-specific copy number analysis of tumors" PNAS (2010) 107(39):16910-16915	
	22	WANG, T.T. et al. "High efficiency error suppression for accurate detection of low-frequency variants" NAR (2019) 47 (15):	

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature	/KENNETH R HORLICK/	Date Considered	06/18/2020

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /K.R.H/

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	16575128
	Filing Date	2019-09-18
	First Named Inventor	AmirAli TALASAZ
	Art Unit	
	Examiner Name	
	Attorney Docket Number	42534-704.313

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

See attached certification statement.

☒ The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☒ A certification statement is not submitted herewith.

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Timothy A. Hott/	Date (YYYY-MM-DD)	2020-05-04
Name/Print	Timothy A. Hott	Registration Number	67740

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /K.R.H/

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /K.R.H/

To: Patents@guardanthealth.com,patentdocket@wsgr.com,
From: PAIR_eOfficeAction@uspto.gov
Cc: PAIR_eOfficeAction@uspto.gov
Subject: Private PAIR Correspondence Notification for Customer Number 115823

Jul 10, 2020 04:19:40 AM

Dear PAIR Customer:

Wilson Sonsini Goodrich & Rosati / Guardant Health
650 Page Mill Road
Palo Alto, CA 94304
UNITED STATES

The following USPTO patent application(s) associated with your Customer Number, 115823 , have new outgoing correspondence. This correspondence is now available for viewing in Private PAIR.

The official date of notification of the outgoing correspondence will be indicated on the form PTOL-90 accompanying the correspondence.

Disclaimer:
The list of documents shown below is provided as a courtesy and is not part of the official file wrapper. The content of the images shown in PAIR is the official record.

Application	Document	Mailroom Date	Attorney Docket No.
16575128	NOA	07/10/2020	42534-704.313
	1449	07/10/2020	42534-704.313

To view your correspondence online or update your email addresses, please visit us anytime at https://sportal.uspto.gov/secure/myportal/privatepair.

If you have any questions, please email the Electronic Business Center (EBC) at EBC@uspto.gov with 'e-Office Action' on the subject line or call 1-866-217-9197 during the following hours:

Monday - Friday 6:00 a.m. to 12:00 a.m.

Thank you for prompt attention to this notice,

UNITED STATES PATENT AND TRADEMARK OFFICE

PATENT APPLICATION INFORMATION RETRIEVAL SYSTEM

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor(s): AmirAli TALASAZ Serial Number: 16/575,128 Filing Date: September 18, 2019 Title: SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION	Confirmation No.: 9440 Customer No.: 115823 Group Art Unit: 1637 Examiner: Kenneth R. HORLICK

Mail Stop Issue Fee
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

AMENDMENT UNDER 37 C.F.R. §1.312

Dear Commissioner:

This communication is submitted under 37 C.F.R. §1.312 after mailing of the Notice of Allowance on July 10, 2020.

Amendments to the Specification begin on page 2 of this paper.
Amendments to the Claims begin on page 3 of this paper.
Remarks and **Conclusion** begin on page 8 of this paper.

AMENDMENTS TO THE SPECIFICATION

Please replace paragraph [0001] with the following paragraph:

[0001] This application is a continuation of U.S. Patent Application 16/283,635, filed February 22, 2019 (now U.S. Patent 10,494,678, issued December 3, 2019), which is a continuation of U.S. Patent Application No. 15/872,831, filed January 16, 2018 (now U.S. Patent 10,457,995, issued October 29, 2019), which is a continuation application of U.S. Patent Application No. 15/828,099, filed November 30, 2017, which is a continuation application of U.S. Patent Application No. 15/467,570, filed March 23, 2017 (now U.S. Patent 9,840,743, issued December 12, 2017), which is a continuation application of U.S. Patent Application No. 14/425,189, filed March 2, 2015 (now U.S. Patent 10,041,127, issued August 7, 2018), which is a national stage entry of international Application No. PCT/US2013/058061, filed September 4, 2013, which claims priority to U.S. Provisional Patent Application No. 61/696,734, filed September 4, 2012, U.S. Provisional Patent Application No. 61/704,400, filed September 21, 2012, [[and]] U.S. Provisional Patent Application No. 61/793,997, filed March 15, 2013, and U.S. Provisional Patent Application No. 61/845,987, filed July 13, 2013, each of which is entirely incorporated herein by reference for all purposes.

AMENDMENTS TO THE CLAIMS

This listing of claims will replace all prior versions and listings in the above-referenced patent application. The foregoing amendments are without prejudice and do not constitute an admission regarding the patentability of the amended subject matter and should not so be construed. Applicant reserves the right to pursue the subject matter of the canceled claims in this or any other appropriate patent application.

Listing of Claims:

1. - 30. (Cancelled).

31. (Currently amended): A method for inferring a quantitative measure of microsatellite changes in a sample of cell-free nucleic acid molecules from a subject having a cancer, the method comprising:

(a) tagging a plurality of the cell-free nucleic acid molecules from the sample with molecular barcodes from a set of molecular barcodes to produce tagged parent polynucleotides;

(b) amplifying a plurality of the tagged parent polynucleotides to produce amplified tagged progeny polynucleotides;

(c) sequencing a plurality of the amplified tagged progeny polynucleotides to produce a set of sequencing reads;

(d) grouping a plurality of sequencing reads from the set of sequencing reads into families based at least on sequence information of the molecular barcodes, each family having sequencing reads amplified from [[the]] a same tagged parent polynucleotide from among the tagged parent polynucleotides;

(e) determining a quantitative measure of ~~sequence~~ sequencing reads in a plurality of the families and a quantitative measure of unique families from among a plurality of the families;

(f) inferring a ~~measure~~ number of the tagged parent polynucleotides ~~in the set of tagged parent polynucleotides~~ based on the quantitative measure of the ~~sequence~~ sequencing reads in the plurality of the families and the quantitative measure of unique families from among the plurality of the families;

(g) determining a quantitative measure of microsatellite changes from among a plurality of the families; and

(h) based on the determined quantitative measure of microsatellite changes from among the plurality of the families, inferring a quantitative measure of microsatellite changes in the inferred number of ~~inferred~~ the tagged parent polynucleotides.

32. (Previously presented): The method of claim 31, wherein the sample is a bodily fluid sample selected from the group consisting of blood, plasma, and serum.

33. (Currently amended): The method of claim 31, wherein a molecular barcode from the set of molecular barcodes is attached on both ends of a molecule of the plurality of the cell-free nucleic acid molecules.

34. (Previously presented): The method of claim 31, wherein the molecular barcodes from the set of molecular barcodes have 2 to 1,000 different molecular barcode sequences.

35. (Previously presented): The method of claim 31, wherein the tagging comprises performing blunt-end ligation or sticky-end ligation.

36. (Currently amended): The method of claim 31, further comprising selectively enriching at least a portion of the amplified tagged progeny polynucleotides for target regions associated with cancer prior to the sequencing.

37. (Currently amended): The method of claim 31, further comprising filtering out sequencing reads from among the set of sequencing reads that fail to meet a quality threshold.

38. (Previously presented): The method of claim 31, further comprising mapping a plurality of sequencing reads from the set of sequencing reads to a reference sequence.

39. (Currently amended): The method of claim 38, wherein the grouping is based at least on a combination of the sequence information of the molecular barcodes and a beginning base position and an end base position of the sequencing reads that map to the reference sequence.

40. (Currently amended): The method of claim 31, further comprising, for a plurality of the families, collapsing [[the]] sequencing reads [[in]] within a given family of the plurality of the families to generate consensus sequences ~~reads~~.

41. (Currently amended): The method of claim 40, further comprising identifying one or more consensus sequences [[reads]] having a somatic mutation as compared to a reference sequence.

TEMPUS EX1002.79

42. (Currently amended): The method of claim 31, further comprising determining a treatment regimen for the subject based on the inferred quantitative measure of microsatellite changes in the inferred number of ~~inferred~~ the tagged parent polynucleotides.

43. (Currently amended): A method for detecting a genetic variation in one or more microsatellite regions in a sample of cell-free nucleic acid molecules from a subject having a cancer, the method comprising:

(a) ligating molecular barcodes from a set of molecular barcodes having 2 to 1,000,000 different molecular barcode sequences to a ~~population~~ plurality of the cell-free nucleic acid[[s]] molecules ~~obtained~~ from the sample to produce tagged parent polynucleotides;

(b) amplifying a plurality of the tagged parent polynucleotides to produce amplified tagged progeny polynucleotides;

(c) sequencing a plurality of the amplified tagged progeny polynucleotides to produce a set of sequencing reads; and

(d) determining, from among a plurality of sequencing reads in the set of sequencing reads, a quantitative measure of polymorphic forms comprising microsatellite changes in the one or more microsatellite regions based at least on sequence information [[from]] of the molecular barcodes, thereby detecting the genetic variation in the one or more microsatellite regions.

44. (Previously presented): The method of claim 43, wherein the genetic variation in the one or more microsatellite regions comprises an insertion or deletion (indel).

45. (Previously presented): The method of claim 43, wherein the sample is a bodily fluid sample selected from the group consisting of blood, plasma, and serum.

46. (Currently amended): The method of claim 43, wherein a molecular barcode from the set of molecular barcodes is attached on both ends of a molecule of the plurality of the cell-free nucleic acid molecules.

47. (Currently amended): The method of claim 43, wherein the molecular barcodes from the set of molecular barcodes ~~comprises molecular barcodes having~~ have 2 to 1,000 different molecular barcode sequences.

TEMPUS EX1002.80

48. (Currently amended): The method of claim 43, wherein the molecular barcodes are ligated to the plurality of the cell-free nucleic acid molecules by blunt-end ligation or sticky-end ligation.
49. (Currently amended): The method of claim 43, further comprising selectively enriching at least a portion of the amplified tagged progeny polynucleotides for target regions associated with cancer prior to the sequencing.
50. (Currently amended): The method of claim 43, further comprising filtering out sequencing reads from among the set of sequencing reads that fail to meet a quality threshold.
51. (Previously presented): The method of claim 43, further comprising mapping a plurality of sequencing reads from the set of sequencing reads to a reference sequence.
52. (Currently amended): The method of claim 51, further comprising grouping a subset of sequencing reads into families based on sequence information of the molecular barcodes and (1) a start base position of a given sequencing read from among the subset of sequencing reads at which the given sequencing read is determined to start mapping to the reference sequence [[and/]]or (2) a stop base position of the given sequencing read at which the given sequencing read is determined to stop mapping to the reference sequence.
53. (Currently amended): The method of claim 52, further comprising, for a plurality of the families, collapsing [[the]] sequencing reads [[in]] within a given family of the plurality of the families to generate consensus sequences ~~reads~~.
54. (Currently amended): The method of claim 53, further comprising identifying one or more consensus sequences [[reads]] having the genetic variation as compared to the reference sequence.
55. (Currently amended): The method of claim 43, wherein the ~~population~~ sample of cell-free nucleic acid[[s]] molecules comprises 1 nanogram (ng) to 100 ng of cell-free nucleic acid molecules.
56. (Currently amended): The method of claim 43, wherein the cell-free nucleic acid[[s]] molecules comprise [[are]] cell-free deoxyribonucleic acid (cfDNA) molecules.
57. (Cancelled).

58. (Currently amended): The method of claim 43, further comprising determining a confidence score based on ~~the presence or absence of~~ the genetic variation detected in the one or more microsatellite regions.

59. (Previously presented): The method of claim 43, further comprising determining a treatment regimen for the subject based on the quantitative measure of polymorphic forms comprising microsatellite changes in the one or more microsatellite regions.

60. (Previously presented): The method of claim 59, further comprising administering the treatment regimen to the subject.

61. (Currently amended): A method for detecting a genetic variation in one or more microsatellite regions in a sample of cell-free deoxyribonucleic acid (cfDNA) molecules from a subject having a cancer, the method comprising:

(a) tagging a ~~population~~ plurality of ~~cell-free nucleic acids obtained~~ the cfDNA molecules from the sample with molecular barcodes from a set of molecular barcodes to produce tagged parent polynucleotides, wherein a molecular barcode from the set of molecular barcodes [[are]] is attached to both ends of a molecule of the plurality of the cfDNA molecules ~~from the population of cfDNA molecules~~, and wherein a plurality of the tagged parent polynucleotides has identical molecular barcode sequences;

(b) amplifying the tagged parent polynucleotides to produce amplified tagged progeny polynucleotides;

(c) sequencing a plurality of the amplified tagged progeny polynucleotides to produce a set of sequencing reads;

(d) identifying, from among a plurality of sequencing reads in the set of sequencing reads, a plurality of unique ~~sequence~~ sequencing reads based at least on sequence information [[from]] of the molecular barcodes, wherein a unique ~~sequence~~ sequencing read corresponds to a nucleic acid sequence of a cfDNA molecule from among the tagged parent polynucleotides; and

(e) determining, from among [[a]] the plurality of unique ~~sequence~~ sequencing reads, a quantitative measure of polymorphic forms comprising microsatellite changes in the one or more microsatellite regions, thereby detecting the genetic variation in the one or more microsatellite regions.

REMARKS

Claims 31-56 and 58-61 were allowed prior to entry of the abovementioned amendments. Claims 31, 33, 36-37, 39-43, 46-50, 52-56, 58 and 61 are amended herein for clarity and/or antecedent basis purposes. Applicant submits that the amendments are fully supported by the application as filed. These amendments do not add any new matter or raise new issues. Accordingly, Applicant respectfully requests entry of these amendments under 37 C.F.R. §1.312.

CONCLUSION

Should the Examiner have any question, the Examiner is respectfully encouraged to telephone the undersigned attorney. The Commissioner is authorized to charge any underpayment or credit any overpayment to Deposit Account No. 60-2231 (Attorney Docket No. GH0002US-CON14).

Respectfully submitted,
GUARDANT HEALTH, INC.

Date: July 13, 2020

By: /Timothy A. Hott/

Timothy A. Hott
Registration No.: 67740

GUARDANT HEALTH, INC.
505 Penobscot Drive
Redwood City, CA 94063
Customer No. 115823

Electronic Acknowledgement Receipt

EFS ID:	39980632
Application Number:	16575128
International Application Number:	
Confirmation Number:	9440
Title of Invention:	SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION
First Named Inventor/Applicant Name:	AmirAli TALASAZ
Customer Number:	115823
Filer:	Timothy A Hott/Michelle Chan
Filer Authorized By:	Timothy A Hott
Attorney Docket Number:	42534-704.313
Receipt Date:	13-JUL-2020
Filing Date:	18-SEP-2019
Time Stamp:	13:54:48
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	no

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1		2020-07-13_GH0002US-CON14_312Amend.pdf	195603 f77b20d3b97f42fad43f3539978126732107a346	yes	8

	Multipart Description/PDF files in .zip description		
	Document Description	**Start**	**End**
	Amendment after Notice of Allowance (Rule 312)	1	1
	Specification	2	2
	Claims	3	7
	Applicant Arguments/Remarks Made in an Amendment	8	8
Warnings:			
Information:			
	Total Files Size (in bytes):	195603	

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.
National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.
New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

Doc Code: IFEE PTOL/85B-EFS

Document Description: Issue Fee Payment (PTO-85B)

Issue Fee Transmittal Form

Application Number	Filing Date	First Named Inventor	Atty. Docket No.	Confirmation No.
16575128	18-Sep-2019	AmirAli TALASAZ	42534-704.313	9440

TITLE OF INVENTION :

SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION

Entity Status	Application Type	Art Unit	Class - Subclass	EXAMINER
Regular Undiscounted	Utility under 35 USC 111(a)	1637	006120	KENNETH HORLICK

Issue Fee Due	Publication Due	Total Fee(s) Due	Date Due	Prev. Paid Fee
$1000	$0	$1000	13-Oct-2020	$0

1.Change of Correspondence Address and/or Indication Of Fee Address (37 CFR 1.33 & 1.363)

Current Correspondence Address:	Current Indicated Fee Address :
115823 Wilson Sonsini Goodrich & Rosati / Guardant Health 650 Page Mill Road Palo Alto CA 94304 UNITED STATES 650-493-9300 Patents@guardanthealth.com	
☐ Change of correspondence address requested, system generated AIA/122-EFS form attached	☐ Fee Address indication requested, system generated SB/47-EFS form attached

2.Entity Status

Change in Entity Status

○ Applicant certifying micro entity status; system generated Micro Entity certification form attached. See 37 CFR 1.29.
Note: Absent a valid certification of micro entity status, issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment. If this box is checked, you will be prompted to choose a micro entity status on the gross income basis (37 CFR 1.29(a)) or the institution of higher education basis (37 CFR 1.29(d)), and make the applicable certification online.

○ Applicant asserting small entity status. See 37 CFR 1.27.
Note: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.

◉ Applicant changing to regular undiscounted fee status.
Note: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

Doc Code: IFEE PTOL/85B-EFS

Document Description: Issue Fee Payment (PTO-85B)

3.The Following Fee(s) Are Submitted:

☒ Issue Fee

☐ I authorize USPTO to apply my previously paid issue fee to the current fees due

☐ Publication Fee

☐ The Director is hereby authorized to apply my previously paid issue fee to the current fee due and to charge deficient fees to Deposit Account Number _____

☐ Advance Order - # of copies _____

☒ If i**n addition to** the payment of the issue fee amount submitted with this form, there are any discrepancies in any amount(s) due, the Director is authorized to charge any deficiency, or credit any overpayment, to Deposit Account Number 602231 . The **issue fee must be submitted** with this form. **If payment of the issue fee does not accompany this form, checking this box and providing a deposit account number will NOT be effective to satisfy full payment of the fee(s) due.**

4.Firm and/or Attorney Names To Be Printed

NOTE: If no name is listed, no name will be printed
For printing on the patent front page, list to be displayed as entered

1. TIMOTHY HOTT
2.
3.

5.Assignee Name(s) and Residence Data To Be Printed

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

Name	City	State	Country	Category
GUARDANT HEALTH, INC.	Redwood City	CALIFORNIA	united states	corporation

6.Signature

I certify, in accordance with 37 CFR 1.4(d)(4) that I am an attorney or agent registered to practice before the Patent and Trademark Office who has filed and has been granted power of attorney in this application. I also certify that this Fee(s) Transmittal form is being transmitted to the USPTO via EFS-WEB on the date indicated below.

Signature	/Timothy A. Hott/	**Date**	07-15-2020
Name	Timothy A Hott	**Registration Number**	67740

WEB IFEE 1.0

Electronic Patent Application Fee Transmittal

Application Number:	16575128
Filing Date:	18-Sep-2019
Title of Invention:	SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION
First Named Inventor/Applicant Name:	AmirAli TALASAZ
Filer:	Timothy A Hott/Michelle Chan
Attorney Docket Number:	42534-704.313

Filed as Large Entity

Filing Fees for Utility under 35 USC 111(a)

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
UTILITY APPL ISSUE FEE	1501	1	1000	1000
PUBL. FEE- EARLY, VOLUNTARY, OR NORMAL	1504	1	0	0
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Miscellaneous:				
		Total in USD ($)		**1000**

Electronic Acknowledgement Receipt

EFS ID:	40012327
Application Number:	16575128
International Application Number:	
Confirmation Number:	9440
Title of Invention:	SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION
First Named Inventor/Applicant Name:	AmirAli TALASAZ
Customer Number:	115823
Filer:	Timothy A Hott/Michelle Chan
Filer Authorized By:	Timothy A Hott
Attorney Docket Number:	42534-704.313
Receipt Date:	15-JUL-2020
Filing Date:	18-SEP-2019
Time Stamp:	17:21:00
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	DA
Payment was successfully received in RAM	$1000
RAM confirmation Number	E20207EH20574074
Deposit Account	
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Issue Fee Payment (PTO-85B)	Web85b.pdf	46290 66149aa6ffb0de4fa55150308e8e3961951f032a	no	2
Warnings:					
Information:					
2	Fee Worksheet (SB06)	fee-info.pdf	31976 34b5f363f72c13ff9c9ccd296d398abd650c5feb	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	78266		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

PRINTER RUSH
(PTO ASSISTANCE)

Application: 16575128	**Examiner:** Horlick	**GAU:** 1637
From:	**Location:** RTFM	**Creation Date:** 07/30/2020

Week Date: 7/6/2020

DOC CODE	DOC DATE	MISCELLANEOUS
☐ 1449		☐ Continuing Data
☐ IDS		☐ Foreign Priority
☐ CLM		☐ Document Legibility
☐ IIFW/FWCLM		☐ Fees
☐ SRFW		☐ Petition (TC)
☐ DRW		☐ Other
☐ OATH		
☒ 312	07/13/2020	
☐ SPEC		

[RUSH] Message:

Please respond to the 07/13/2020 A.NA.

Thank You

[XRUSH] Response:

Initials: _____

Examiner: PUBS contact - James Leija, (703) 756-1197
NOTE: This form will be included as part of the official USPTO record with the response document coded as XRUSH.
REV: AUG 16

TEMPUS EX1002.92

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/575,128	09/18/2019	AmirAli TALASAZ	42534-704.313	9440

115823 7590 08/07/2020

Wilson Sonsini Goodrich & Rosati / Guardant Health
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER
HORLICK, KENNETH R

ART UNIT	PAPER NUMBER
1637	

NOTIFICATION DATE	DELIVERY MODE
08/07/2020	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

Patents@guardanthealth.com
patentdocket@wsgr.com

CORRECTED ***Notice of Allowability***	**Application No.** 16/575,128	**Applicant(s)** TALASAZ, AmirAli	
	Examiner KENNETH R HORLICK	**Art Unit** 1637	**AIA (FITF) Status** Yes

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☑ This communication is responsive to the response filed 07/13/20.
   ☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2. ☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

3. ☑ The allowed claim(s) is/are 31-56 and 58-61 . As a result of the allowed claim(s), you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information , please see **http://www.uspto.gov/patents/init_events/pph/index.jsp** or send an inquiry to **PPHfeedback@uspto.gov.**

4. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

   Certified copies:

   a) ☐ All b) ☐ Some *c) ☐ None of the:

   1. ☐ Certified copies of the priority documents have been received.
   2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
   3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____ .

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____ .

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).

6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☐ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date _____.
3. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material _____.
4. ☐ Interview Summary (PTO-413), Paper No./Mail Date. _____.
5. ☐ Examiner's Amendment/Comment
6. ☐ Examiner's Statement of Reasons for Allowance
7. ☑ Other PTO-271.

/KENNETH R HORLICK/
Primary Examiner, Art Unit 1637

Response to Rule 312 Communication	Application No.	Applicant(s)	
	16/575,128	TALASAZ, AmirAli	
	Examiner	Art Unit	AIA (FITF) Status
	KENNETH R HORLICK	1637	Yes

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

1. ☑ The amendment filed on 13 July 2020 under 37 CFR 1.312 has been considered, and has been:

   a) ☐ entered.

   b) ☑ entered as directed to matters of form not affecting the scope of the invention.

   c) ☐ disapproved because the amendment was filed after the payment of the issue fee.

   Any amendment filed after the date the issue fee is paid must be accompanied by a petition under 37 CFR 1.313(c)(1) and the required fee to withdraw the application from issue.

   d) ☐ disapproved. See explanation below.

   e) ☐ entered in part. See explanation below.

/KENNETH R HORLICK/
Primary Examiner, Art Unit 1637

Index of Claims	Application/Control No.	Applicant(s)/Patent Under Reexamination
	16/575,128	TALASAZ, AmirAli
	Examiner	**Art Unit**
	KENNETH R HORLICK	1637

✓	Rejected	-	Cancelled	N	Non-Elected	A	Appeal
=	Allowed	÷	Restricted	I	Interference	O	Objected

CLAIMS

☑ Claims renumbered in the same order as presented by applicant ☐ CPA ☑ T.D. ☐ R.1.47

CLAIM		DATE								
Final	Original	02/14/2020	06/29/2020							
1	31	✓	=							
2	32	✓	=							
3	33	✓	=							
4	34	✓	=							
5	35	✓	=							
6	36	✓	=							
7	37	✓	=							
8	38	✓	=							
9	39	✓	=							
10	40	✓	=							
11	41	✓	=							
12	42	✓	=							
13	43	✓	=							
14	44	✓	=							
15	45	✓	=							
16	46	✓	=							
17	47	✓	=							
18	48	✓	=							
19	49	✓	=							
20	50	✓	=							
21	51	✓	=							
22	52	✓	=							
23	53	✓	=							
24	54	✓	=							
25	55	✓	=							
26	56	✓	=							
	57	✓	-							
27	58	✓	=							
28	59	✓	=							
29	60	✓	=							
30	61		=							

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	16/575,128	TALASAZ, AmirAli
	Examiner	**Art Unit**
	KENNETH R HORLICK	1637

CPC							
Symbol						**Type**	**Version**
C12Q	/	1	/	6886		F	2013-01-01
G16B	/	30	/	00		I	2019-02-01
C12Q	/	1	/	6869		I	2013-01-01
C12N	/	15	/	1065		I	2013-01-01
C12Q	/	1	/	6806		I	2013-01-01
C12Q	/	1	/	6874		I	2013-01-01
C12Q	/	2600	/	118		A	2013-01-01
C12Q	/	2600	/	156		A	2013-01-01
C12Q	/	2600	/	16		A	2013-01-01
C12Q	/	2600	/	158		A	2013-01-01

CPC Combination Sets								
Symbol					**Type**	**Set**	**Ranking**	**Version**
C12Q	/	1	/	6869	A	1	1	2013-01-01
C12Q	/	2537	/	165	A	1	2	2013-01-01
C12Q	/	2545	/	114	A	1	3	2013-01-01

NONE		**Total Claims Allowed:**	
(Assistant Examiner)	(Date)	30	
/KENNETH R HORLICK/ Primary Examiner, Art Unit 1637	03 August 2020	O.G. Print Claim(s)	O.G. Print Figure
(Primary Examiner)	(Date)	1	NONE

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	16/575,128	TALASAZ, AmirAli
	Examiner	**Art Unit**
	KENNETH R HORLICK	1637

INTERNATIONAL CLASSIFICATION				
CLAIMED				
C12P	/	19	/	34

NON-CLAIMED				
	/		/	

US ORIGINAL CLASSIFICATION	
CLASS	**SUBCLASS**
435	6.12

CROSS REFERENCES(S)						
CLASS	**SUBCLASS (ONE SUBCLASS PER BLOCK)**					
435	91.2					

NONE (Assistant Examiner)	(Date)	**Total Claims Allowed:** 30	
/KENNETH R HORLICK/ Primary Examiner, Art Unit 1637 (Primary Examiner)	03 August 2020 (Date)	O.G. Print Claim(s) 1	O.G. Print Figure NONE

U.S. Patent and Trademark Office — Part of Paper No.: 20200803a

TEMPUS EX1002.98

Issue Classification	Application/Control No.	Applicant(s)/Patent Under Reexamination
	16/575,128	TALASAZ, AmirAli
	Examiner	**Art Unit**
	KENNETH R HORLICK	1637

☑ Claims renumbered in the same order as presented by applicant ☐ CPA ☑ T.D. ☐ R.1.47

CLAIMS

Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original	Final	Original
1	31	10	40	19	49	27	58								
2	32	11	41	20	50	28	59								
3	33	12	42	21	51	29	60								
4	34	13	43	22	52	30	61								
5	35	14	44	23	53										
6	36	15	45	24	54										
7	37	16	46	25	55										
8	38	17	47	26	56										
9	39	18	48		57										

NONE		**Total Claims Allowed:**	
(Assistant Examiner)	(Date)	30	
/KENNETH R HORLICK/ Primary Examiner, Art Unit 1637	03 August 2020	O.G. Print Claim(s)	O.G. Print Figure
(Primary Examiner)	(Date)	1	NONE

U.S. Patent and Trademark Office — Part of Paper No.: 20200803a

TEMPUS EX1002.99

PRINTER RUSH
(PTO ASSISTANCE)

Application: 16575128	**Examiner:** Horlick	**GAU:** 1637
From:	**Location:** RTFM	**Creation Date:** 07/30/2020

Week Date: 7/6/2020

DOC CODE	DOC DATE	MISCELLANEOUS
☐ 1449		☐ Continuing Data
☐ IDS		☐ Foreign Priority
☐ CLM		☐ Document Legibility
☐ IIFW/FWCLM		☐ Fees
☐ SRFW		☐ Petition (TC)
☐ DRW		☐ Other
☐ OATH		
☒ 312	07/13/2020	
☐ SPEC		

[RUSH] Message:

Please respond to the 07/13/2020 A.NA.

Thank You

[XRUSH] Response:

see attached

Initials: KH

Examiner: PUBS contact - James Leija, (703) 756-1197
NOTE: This form will be included as part of the official USPTO record with the response document coded as XRUSH.
REV: AUG 16

TEMPUS EX1002.100

Attorney Docket: 704.313/GH0002US-CON14

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventor(s): AmirAli TALASAZ Serial Number: 16/575,128 Filing Date: September 18, 2019 Title: SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION	Confirmation No.: 9440 Customer No.: 115823 Group Art Unit: 1637 Examiner: Kenneth R. HORLICK

Mail Stop Issue Fee
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

OK TO ENTER: /K.R.H/

/K.R.H/ (08/03/2020)

AMENDMENT UNDER 37 C.F.R. §1.312

Dear Commissioner:

This communication is submitted under 37 C.F.R. §1.312 after mailing of the Notice of Allowance on July 10, 2020.

Amendments to the Specification begin on page 2 of this paper.

Amendments to the Claims begin on page 3 of this paper.

Remarks and **Conclusion** begin on page 8 of this paper.

To: Patents@guardanthealth.com,patentdocket@wsgr.com,
From: PAIR_eOfficeAction@uspto.gov
Cc: PAIR_eOfficeAction@uspto.gov
Subject: Private PAIR Correspondence Notification for Customer Number 115823

Aug 07, 2020 03:51:36 AM

Dear PAIR Customer:

Wilson Sonsini Goodrich & Rosati / Guardant Health
650 Page Mill Road
Palo Alto, CA 94304
UNITED STATES

The following USPTO patent application(s) associated with your Customer Number, 115823 , have new outgoing correspondence. This correspondence is now available for viewing in Private PAIR.

The official date of notification of the outgoing correspondence will be indicated on the form PTOL-90 accompanying the correspondence.

Disclaimer:
The list of documents shown below is provided as a courtesy and is not part of the official file wrapper. The content of the images shown in PAIR is the official record.

Application	Document	Mailroom Date	Attorney Docket No.
16575128	NOA	08/07/2020	42534-704.313
	N271	08/07/2020	42534-704.313

To view your correspondence online or update your email addresses, please visit us anytime at https://sportal.uspto.gov/secure/myportal/privatepair.

If you have any questions, please email the Electronic Business Center (EBC) at EBC@uspto.gov with 'e-Office Action' on the subject line or call 1-866-217-9197 during the following hours:

Monday - Friday 6:00 a.m. to 12:00 a.m.

Thank you for prompt attention to this notice,

UNITED STATES PATENT AND TRADEMARK OFFICE

PATENT APPLICATION INFORMATION RETRIEVAL SYSTEM

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	ISSUE DATE	PATENT NO.	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/575,128	09/08/2020	10767233	42534-704.313	9440

115823 7590 08/19/2020

Wilson Sonsini Goodrich & Rosati / Guardant Health
650 Page Mill Road
Palo Alto, CA 94304

ISSUE NOTIFICATION

The projected patent number and issue date are specified above.

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(application filed on or after May 29, 2000)

The Patent Term Adjustment is 0 day(s). Any patent to issue from the above-identified application will include an indication of the adjustment on the front page.

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Application Assistance Unit (AAU) of the Office of Data Management (ODM) at (571)-272-4200.

APPLICANT(s) (Please see PAIR WEB site http://pair.uspto.gov for additional applicants):

AmirAli TALASAZ, Atherton, CA;
GUARDANT HEALTH, INC., Redwood City, CA;

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The USA offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to encourage and facilitate business investment. To learn more about why the USA is the best country in the world to develop technology, manufacture products, and grow your business, visit SelectUSA.gov.

To: Patents@guardanthealth.com,patentdocket@wsgr.com,
From: PAIR_eOfficeAction@uspto.gov
Cc: PAIR_eOfficeAction@uspto.gov
Subject: Private PAIR Correspondence Notification for Customer Number 115823

Aug 20, 2020 04:08:43 AM

Dear PAIR Customer:

Wilson Sonsini Goodrich & Rosati / Guardant Health
650 Page Mill Road
Palo Alto, CA 94304
UNITED STATES

The following USPTO patent application(s) associated with your Customer Number, 115823 , have new outgoing correspondence. This correspondence is now available for viewing in Private PAIR.

The official date of notification of the outgoing correspondence will be indicated on the form PTOL-90 accompanying the correspondence.

Disclaimer:
The list of documents shown below is provided as a courtesy and is not part of the official file wrapper. The content of the images shown in PAIR is the official record.

Application	Document	Mailroom Date	Attorney Docket No.
16575128	ISSUE.NTF	08/19/2020	42534-704.313

To view your correspondence online or update your email addresses, please visit us anytime at https://sportal.uspto.gov/secure/myportal/privatepair.

If you have any questions, please email the Electronic Business Center (EBC) at EBC@uspto.gov with 'e-Office Action' on the subject line or call 1-866-217-9197 during the following hours:

Monday - Friday 6:00 a.m. to 12:00 a.m.

Thank you for prompt attention to this notice,

UNITED STATES PATENT AND TRADEMARK OFFICE

PATENT APPLICATION INFORMATION RETRIEVAL SYSTEM

Doc code: RCEX
Doc description: Request for Continued Examination (RCE)

PTO/SB/30EFS (02-18)
Approved for use through 11/30/2020. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

REQUEST FOR CONTINUED EXAMINATION(RCE)TRANSMITTAL
(Submitted Only via EFS-Web)

Application Number	16/575,128	Filing Date	2019-09-18	Docket Number (if applicable)	42534-704.313	Art Unit	1637
First Named Inventor	AmirAli TALASAZ			Examiner Name	Kenneth R. HORLICK		

This is a Request for Continued Examination (RCE) under 37 CFR 1.114 of the above-identified application.
Request for Continued Examination (RCE) practice under 37 CFR 1.114 does not apply to any utility or plant application filed prior to June 8, 1995, or to any design application. The Instruction Sheet for this form is located at WWW.USPTO.GOV

SUBMISSION REQUIRED UNDER 37 CFR 1.114

Note: If the RCE is proper, any previously filed unentered amendments and amendments enclosed with the RCE will be entered in the order in which they were filed unless applicant instructs otherwise. If applicant does not wish to have any previously filed unentered amendment(s) entered, applicant must request non-entry of such amendment(s).

☐ Previously submitted. If a final Office action is outstanding, any amendments filed after the final Office action may be considered as a submission even if this box is not checked.

☐ Consider the arguments in the Appeal Brief or Reply Brief previously filed on _____

☐ Other _____

☒ Enclosed

☐ Amendment/Reply

☒ Information Disclosure Statement (IDS)

☐ Affidavit(s)/ Declaration(s)

☐ Other _____

MISCELLANEOUS

☐ Suspension of action on the above-identified application is requested under 37 CFR 1.103(c) for a period of months _____
(Period of suspension shall not exceed 3 months; Fee under 37 CFR 1.17(i) required)

☐ Other _____

FEES

☒ **The RCE fee under 37 CFR 1.17(e) is required by 37 CFR 1.114 when the RCE is filed.**
The Director is hereby authorized to charge any underpayment of fees, or credit any overpayments, to Deposit Account No 602231

SIGNATURE OF APPLICANT, ATTORNEY, OR AGENT REQUIRED

☒ Patent Practitioner Signature

Applicant Signature

PTO/SB/30EFS (02-18)
Approved for use through 11/30/2020. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

Signature of Registered U.S. Patent Practitioner			
Signature	/Timothy A. Hott/	Date (YYYY-MM-DD)	2020-08-28
Name	Timothy A. Hott	Registration Number	67740

This collection of information is required by 37 CFR 1.114. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Doc Code: QPIDS.REQ
Document Description: QuickPath Information Disclosure Statement

PTO/SB/09 (09-18)

CERTIFICATION AND REQUEST FOR CONSIDERATION OF AN INFORMATION DISCLOSURE STATEMENT FILED AFTER PAYMENT OF THE ISSUE FEE UNDER THE QPIDS PROGRAM

Non-Provisional Application Number: 16/575,128	Filing Date: 2019-09-18
First Named Inventor: AmirAli TALASAZ	Title of Invention: SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION

THE UNDERSIGNED HEREBY CERTIFIES AND REQUESTS THE FOLLOWING FOR THE ABOVE-IDENTIFIED APPLICATION.

1. Consideration is requested of the information disclosure statement (IDS) submitted herewith, which is being filed after payment of the issue fee.
2. Check the box next to the appropriate selection:

   ☐ Each item of information contained in the IDS was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the IDS. See 37 CFR 1.97(e)(1).

   OR

   ☑ No item of information contained in the IDS was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the IDS was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the IDS. See 37 CFR 1.97(e)(2).

   OR

   ☐ See attached certification statement in compliance with 37 CFR 1.97(e).
3. Please charge the IDS fee set forth in 37 CFR 1.17(p) to Deposit Account No. 602231.
4. A Petition to Withdraw from Issue After Payment of the Issue Fee (37 CFR 1.313(c)(2)), including the petition fee set forth in 37 CFR 1.17(h), is submitted herewith as a **Web-based ePetition**. **WARNING:** Do not submit the petition as a follow-on paper via EFS-Web. Submit the petition as a Web-based ePetition by signing on to EFS-Web as a registered user, selecting the radio button next to "Existing application/patent," and then selecting the radio button next to "ePetition (for automatic processing and immediate grant, if all petitions requirements are met)." Failure to use the Web-based ePetition interface will result in automatic entry of the RCE.
5. A request for continued examination (RCE) under 37 CFR 1.114 and the RCE fee under 37 CFR 1.17(e) are submitted herewith.
6. The RCE will be treated as a "conditional" RCE. In the event the examiner determines that any item of information contained in the IDS necessitates the reopening of prosecution in the application, the undersigned understands that (i) the RCE will be processed and treated as an RCE under 37 CFR 1.114 and therefore (ii) the IDS fee under 37 CFR 1.17(p) will be returned in accordance with 37 CFR 1.97(b)(4). In the event that no item of information in the IDS necessitates reopening prosecution, the undersigned understands that the RCE will not be processed and the RCE fee under 37 CFR 1.17(e) will be returned.
7. This certification and request is being filed as a **Web-based ePetition** and is not accompanied by an amendment to the application. Inclusion of an amendment will result in automatic entry of the RCE.

Signature /Timothy A. Hott/	Date 2020-08-28
Name (Print/Typed) Timothy A. Hott	Practitioner Registration Number (If applicable) 67740

Note: *Signatures of all the inventors or assignees of record of the entire interest or their representative(s) are required in accordance with 37 CFR 1.33 and 11.18. Please see 37 CFR 1.4(d) for the form of the signature. If necessary, submit multiple forms for more than one signature, see below.**

☐ *Total of _____ forms are submitted.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

Privacy Act Statement

The **Privacy Act of 1974 (P.L. 93-579)** requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (*i.e.*, GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (02-18)
Approved for use through 11/30/2020. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Not for submission under 37 CFR 1.99)

Application Number	16575128
Filing Date	2019-09-18
First Named Inventor	AmirAli TALASAZ
Art Unit	1637
Examiner Name	Kenneth R. HORLICK
Attorney Docket Number	42534-704.313

U.S.PATENTS Remove

Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					

If you wish to add additional U.S. Patent citation information please click the Add button. Add

U.S.PATENT APPLICATION PUBLICATIONS Remove

Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					

If you wish to add additional U.S. Published Application citation information please click the Add button. Add

FOREIGN PATENT DOCUMENTS Remove

Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							

If you wish to add additional Foreign Patent Document citation information please click the Add button Add

NON-PATENT LITERATURE DOCUMENTS Remove

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	16575128
	Filing Date	2019-09-18
	First Named Inventor	AmirAli TALASAZ
	Art Unit	1637
	Examiner Name	Kenneth R. HORLICK
	Attorney Docket Number	42534-704.313

	1	IPR2019-00634, Final Written Decision of U.S. Patent 9,840,743, dated August 18, 2020	
	2	IPR2019-00652, Final Written Decision of U.S. Patent 9,834,822, dated August 18, 2020	

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature		Date Considered	

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (**Not for submission under 37 CFR 1.99)**		
	Application Number	16575128
	Filing Date	2019-09-18
	First Named Inventor	AmirAli TALASAZ
	Art Unit	1637
	Examiner Name	Kenneth R. HORLICK
	Attorney Docket Number	42534-704.313

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☒ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

See attached certification statement.

☒ The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

A certification statement is not submitted herewith.

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Timothy A. Hott/	Date (YYYY-MM-DD)	2020-08-28
Name/Print	Timothy A. Hott	Registration Number	67740

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Trials@uspto.gov
571-272-7822

Paper 52
Date: August 18, 2020

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

FOUNDATION MEDICINE, INC.,
Petitioner,

v.

GUARDANT HEALTH, INC.,
Patent Owner.

IPR2019-00634
Patent 9,840,743 B2

Before TINA E. HULSE, JOHN E. SCHNEIDER,
and KRISTI L. R. SAWERT, *Administrative Patent Judges.*

SCHNEIDER, *Administrative Patent Judge*.

JUDGMENT
Final Written Decision
Determining Some Challenged Claims Unpatentable
Denying Patent Owner's Motion to Amend
Denying Patent Owner's Motion to Exclude
Denying Petitioner's Motion to Exclude
35 U.S.C. § 318(a)

I. INTRODUCTION

This is a Final Written Decision in an *inter partes* review challenging the patentability of claims 1–26 of U.S. Patent 9,840,743 B2 ("the '743 patent," Ex. 1001). We have jurisdiction under 35 U.S.C. § 6, and enter this decision pursuant to 35 U.S.C. § 318(a) and 37 C.F.R. § 42.73. For the reasons set forth below, we determine that Foundation Medicine, Inc., ("Petitioner") has shown by a preponderance of the evidence that claims 1–9, 20, and 22–26 are unpatentable, but has not shown that claims 10–19 or 21 are unpatentable. Additionally, we deny the contingent Motion to Amend filed by Guardant Health, Inc. ("Patent Owner").

A. Procedural History

Petitioner filed a petition for *inter partes* review of the challenged claims under 35 U.S.C. § 311. Paper 2 ("Pet."). Petitioner supported the Petition with the Declaration of Dr. Stacy Gabriel. Ex. 1002. Patent Owner filed a Preliminary Response. Paper 6 ("Prelim. Resp.). Petitioner submitted an authorized Reply to the Preliminary Response. Paper 9

On August 19, 2019, pursuant to 35 U.S.C. § 314(a), we instituted trial to determine whether any of the challenged claims are unpatentable on the grounds raised in the Petition. Paper 10. ("Inst. Dec."). In our Decision, we declined to exercise our discretion to deny the Petition under 35 U.S.C. § 314(a). Inst. Dec. 33–42.

Patent Owner filed a Response to the Petition. Paper 25 ("PO Resp."). Patent Owner supported the Response with the Declaration of Carlos D. Bustamante, Ph.D. (Ex. 2023) and the Declaration of Dr. John Quackenbush (Ex. 2026). Petitioner filed a Reply to the Patent Owner Response. Paper 33 ("Pet. Reply"). Patent Owner filed a Sur-Reply. Paper 37 ("PO Sur-Reply").

Patent Owner filed a Contingent Motion to Amend. Paper 26 ("MTA"). Patent Owner supported the Motion to Amend with the Declaration of Carlos D. Bustamante, Ph.D. (Ex. 2024). Petitioner filed an Opposition to the Motion to Amend. Paper 34 ("MTA Opp."). Petitioner supported the Opposition with the Declaration of Dr. Stacy Gabriel (Ex. 1060). Patent Owner filed a Reply to Petitioner's opposition to the motion. Paper 36 ("MTA Reply"). Petitioner filed a Sur-Reply. Paper 44 ("MTA Sur-Reply").

On May 13, 2020, the parties presented argument at an oral hearing. Paper 41. The hearing transcript has been entered into the record. Paper 51 ("Tr.").

B. Real Parties in Interest

Petitioner identifies the real parties in interest as Foundation Medicine, Inc. and Roche Holdings, Inc., Roche Finance Ltd., and Roche Holding Ltd. Pet. 73. Patent Owner identifies Guardant Health, Inc. as the real party in interest.

C. Related Matters

Petitioner represents that the '743 patent has been asserted in two infringement actions, *Guardant Health, Inc. v. Foundation Medicine, Inc.*, C.A. No. 17-cv-1616-LPS-CJB (D. Del.) and *Guardant Health, Inc. v. Personal Genome Diagnostics, Inc.*, C.A. No. 17-cv-1623-LPS-CJB (D. Del.). Pet. 73–74.

Petitioner has also filed petitions challenging Patent Owner's related patents as follows: IPR2019-00636 (challenging U.S. Patent No. 9,902,992); IPR2019-00637 (also challenging U.S. Patent No. 9,902,992); IPR2019-00652 (challenging U.S. Patent No. 9,834,822); IPR2019-00653 (also

challenging U.S. Patent No. 9,834,822); and IPR209-00130 (challenging U.S. Patent 9,598,731). Paper 3.

Personal Genome Diagnostics, Inc. ("PGDx") also filed a petition seeking post-grant review of the '743 patent, designated PGR2018-00057. That proceeding was terminated without final decision on July 5, 2018. PGR2018-00057, Paper 9.

D. The '743 Patent

The '743 patent, titled "Systems and Methods to Detect Rare Mutations and Copy Number Variation," issued on December 12, 2017, from U.S. Patent Application No. 15/467,570, filed on March 23, 2017. Ex. 1001, codes (54), (45), (21), (22). The '743 patent is a continuation of U.S. Patent Application No. 14/425,189, filed on September 4, 2013. The '743 patent claims priority to U.S. Provisional Application No. 61/845,987, filed on July 13, 2013; U S. Provisional Application No. 61/793,997, filed on March 15, 2013; U.S Provisional Application 61/704,400, filed on September 12, 2012; and U.S. Provisional Application No. 61/696,734, filed on September 4, 2012. *Id.*, code (60).

The '743 patent involves a system and method for detecting rare mutations and copy number variations in cell free polynucleotides. Ex. 1001, code (57). The '743 patent teaches that cell-free DNA ("cfDNA") found in different types of bodily fluids may be used to detect and monitor disease. For instance, cfDNA may contain genetic aberrations like a change in copy number variation and/or single or multiple nucleic acid sequence variation associated with a particular disease that may be used to detect or monitor such disease. *Id.* at col. 1, ll. 29–45, col. 30, ll. 9–12.

The '743 patent states that the disclosed system and method generally "comprise sample preparation, or the extraction and isolation of cell free

polynucleotide sequences from a bodily fluid; subsequent sequencing of cell free polynucleotides by techniques known in the art; and application of bioinformatics tools to detect rare mutations and copy number variations as compared to a reference." *Id.* at col. 30, ll. 11–20.

E. Illustrative Claims

Of the challenged claims, claims 1 and 10 are independent. Claims 2–9, 20, and 22–26 depend, directly or indirectly, from claim 1. Claims 11–19 and 21 depend, directly or indirectly, from claim 10. Claims 1 and 10 read as follows:

> 1. A method for detecting copy number variation, comprising:
>
> a) sequencing extracellular polynucleotides from a bodily sample from a subject, wherein each of the extracellular polynucleotides generates a plurality of sequence reads;
>
> b) filtering out reads that fail to meet a set accuracy, quality score, or mapping score threshold;
>
> c) mapping the plurality of sequence reads to a reference sequence;
>
> d) quantifying mapped reads or unique sequence reads in a plurality of predefined regions of the reference sequence; and
>
> e) determining copy number variation in one or more of the plurality of predefined regions by:
>
> i) normalizing a number of reads in the plurality of predefined regions to each other, or a number of unique sequence reads in the plurality of predefined regions to each other; and/or
>
> ii) processing a number of reads in the plurality of predefined regions or a number of unique sequence reads in the plurality of predefined regions with numbers obtained from a control sample.
>
> 10. A method for detecting a rare mutation in a cell-free or substantially cell-free sample obtained from a subject, comprising:

TEMPUS EX1002.118

> a) sequencing extracellular polynucleotides from a bodily sample from the subject, wherein each of the extracellular polynucleotides generates a plurality of sequence reads;
>
> b) filtering out reads that fail to meet a set accuracy, quality score, or mapping score threshold;
>
> c) mapping sequence reads derived from the sequencing onto a reference sequence;
>
> d) determining unique sequence reads corresponding to the extracellular polynucleotides from among the sequence reads;
>
> e) identifying a subset of mapped unique sequence reads that include a variant as compared to the reference sequence at each mappable base position;
>
> f) for each mappable base position, calculating a ratio of (a) a number of mapped unique sequence reads that include a variant as compared to the reference sequence, to (b) a number of total unique sequence reads for each mappable base position; and
>
> g) processing the ratio with a similarly derived number from a reference sample.

Ex. 1001, col. 62, ll. 43–65, col. 63, ll. 26–49.

F. Evidence

Petitioner relies on the following references:

Chiu et al., *Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma*, 105 PROC. NAT. ACAD. SCI.. 20458 (2008) ("Chiu") Ex. 1022.

Kinde et al., *Detection and quantification of rare mutations with massively parallel sequencing*, 108 PROC. NAT. ACAD. SCI. 9539 (2011) ("Kinde 2011") Ex. 1023.

Kinde et al., *FAST-SeqS: A Simple and Efficient Method for the Detection of Aneuploidy by Massively Parallel Sequencing*, 7 PLOS ONE e41162 (2012) ("Kinde 2012") Ex. 1041.

TEMPUS EX1002.119

Forshew et al., *Noninvasive Identification and Monitoring of Cancer Mutations by Targeted Deep Sequencing of Plasma DNA*, 4 SCI. TRANSL. MED. 1 (2012) ("Forshew") Ex. 1035.

G. Asserted Grounds

Petitioner asserts that claims 1–26 would have been unpatentable on the following grounds:

Claim(s) Challenged	35 U.S.C. §	Reference(s)/Basis
1–3, 22–26	102(a)[1]	Chiu
1–9, 20, 22–26	103	Kinde 2011, Kinde 2012
10–13, 15–19, 21	103	Kinde 2011, Forshew
14	103	Kinde 2011, Kinde 2012, Forshew

II. ANALYSIS

A. Legal Standards

To prevail in this *inter partes* review, Petitioner must demonstrate by a preponderance of the evidence that the claims of the '743 patent are unpatentable. 35 U.S.C. § 316(e); 37 C.F.R. § 42.1(d) (2017). "In an [*inter partes* review], the petitioner has the burden from the onset to show with particularity why the patent it challenges is unpatentable." *Harmonic Inc. v. Avid. Tech., Inc.*, 815 F.3d 1356, 1363 (Fed. Cir. 2016); *see also* 35 U.S.C. § 312(a)(3) (requiring *inter partes* review petitions to identify "with particularity . . . the evidence that supports the grounds for the challenge to each claim"). That burden of persuasion never shifts to Patent Owner.

[1] The provisions of the America Invents Act regarding novelty and obviousness apply to patents containing at least one claim having an effective filing date on or after March 16, 2013. Pub L. 112–29. The '743 patent is subject to the novelty and obviousness provisions of the AIA. Ex. 1036, 4.

Dynamic Drinkware, LLC v. Nat'l Graphics, Inc., 800 F.3d 1375, 1378 (Fed. Cir. 2015); *see also In re Magnum Oil Tools Int'l, Ltd.*, 829 F.3d 1364, 1375–78 (Fed. Cir. 2016) (discussing the burden of proof in *inter partes* review).

"Anticipation requires that all of the claim elements and their limitations are shown in a single prior art reference." *In re Skvorecz*, 580 F.3d 1262, 1266 (Fed. Cir. 2009). "[U]nless a [prior art] reference discloses within the four corners of the document not only all of the limitations claimed but also all of the limitations arranged or combined in the same way as recited in the claim, it cannot be said to prove prior invention of the thing claimed and, thus, cannot anticipate under 35 U.S.C. § 102." *Net MoneyIN, Inc. v. VeriSign, Inc.*, 545 F.3d 1359, 1371 (Fed. Cir. 2008).

A claim is unpatentable for obviousness if, to one of ordinary skill in the pertinent art, "the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious before the effective filing date of the invention." 35 U.S.C. § 103 (2013); *see also KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007). The question of obviousness is resolved on the basis of underlying factual determinations including the scope and content of the prior art, any differences between the claimed subject matter and the prior art, the level of ordinary skill in the art, and objective evidence of nonobviousness. *Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966). A petitioner cannot satisfy its burden of proving obviousness by employing "mere conclusory statements." *Magnum Oil*, 829 F.3d at 1380. Moreover, a decision on the ground of obviousness must include "articulated reasoning with some rational underpinning to support the legal conclusion of obviousness." *KSR*, 550 U.S. at 418 (citing *In re Kahn*, 441 F.3d 977, 988

(Fed. Cir. 2006)). "An obviousness determination requires finding both 'that a skilled artisan would have been motivated to combine the teachings of the prior art references to achieve the claimed invention, and that the skilled artisan would have had a reasonable expectation of success in doing so.'" *CRFD Research, Inc. v. Matal*, 876 F.3d 1330, 1340 (Fed. Cir. 2017) (quoting *Intelligent Bio-Sys., Inc. v. Illumina Cambridge Ltd.*, 821 F.3d 1359, 1367–68 (Fed. Cir. 2016)). "The reasonable expectation of success requirement refers to the likelihood of success in combining references to meet the limitations of the claimed invention." *Intelligent Bio-Sys., Inc.*, 821 F.3d at 1367. A reasonable expectation of success "does not require absolute predictability of success . . . all that is required is a reasonable expectation of success." *In re Kubin*, 561 F.3d 1351, 1360 (Fed. Cir. 2009) (quoting *In re O'Farrell*, 853 F.2d 894, 903–04 (Fed. Cir. 1988)).

We analyze Petitioner's asserted grounds of unpatentability in accordance with the above-stated principles.

B. Level of Ordinary Skill in the Art

The level of ordinary skill in the art is a factual determination that provides a primary guarantee of objectivity in an obviousness analysis. *Al-Site Corp. v. VSI Int'l Inc.*, 174 F.3d 1308, 1324 (Fed. Cir. 1999) (citing *Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966); *Ryko Mfg. Co. v. Nu-Star, Inc.*, 950 F.2d 714, 718 (Fed. Cir. 1991)).

Petitioner contends that a person of ordinary skill in the art at the time the '743 patent was filed would have had "a Ph.D. in genetics, molecular biology, bioinformatics or a related field, and at least five years of research in an academic or industry setting, including at least two to three years of research experience in the field of cancer genomics." Pet. 16 (citing Ex. 1002 ¶ 66). Petitioner also contends that the person of ordinary skill in

TEMPUS EX1002.122

the art would "have had knowledge of DNA sequencing, including next generation sequencing (NGS) and related sequencing methods, and related sample preparation techniques; bioinformatics methods for grouping and comparing sequence reads and mapping sequence reads onto genomes; and methods for identifying genetic variants in a sample." *Id.*

Patent Owner has not proposed a definition of a person of ordinary skill in the art although Dr. Bustamante testified that the definition above describes "a person of exceptional skill, not a person of ordinary skill in the art." Ex. 2023 ¶ 28. Nevertheless, Dr. Bustamante adopted Petitioner's definition in his analysis of the claims. *Id.* ¶ 29.

Accordingly, for this Decision, we adopt Petitioner's definition, while maintaining that the prior art demonstrates the appropriate level of ordinary skill in the art. *See Okajima v. Bourdeau*, 261 F.3d 1350, 1355 (Fed. Cir. 2001) (the prior art, itself, can reflect appropriate level of ordinary skill in art). Moreover, we have reviewed the credentials of Drs. Gabriel, Bustamante, and Quackenbush, and consider each of them to be qualified to provide their opinion on the level of skill and the knowledge of a person of ordinary skill in the art at the time of the invention.

C. Claim Construction

We interpret a claim "using the same claim construction standard that would be used to construe the claim in a civil action under 35 U.S.C. 282(b)." 37 C.F.R. § 42.100(b). Under this standard, we construe the claim "in accordance with the ordinary and customary meaning of such claim as understood by one of ordinary skill in the art and the prosecution history pertaining to the patent." *Id.* Furthermore, we only need to construe terms that are in controversy, and to the extent necessary to resolve the controversy. *Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868

F.3d 1013, 1017 (Fed. Cir. 2017); *Vivid Techs., Inc. v. Am. Sci. & Eng'g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999).

While Petitioner contends that all the claim terms should be given their ordinary meaning and that no claim construction is needed, Patent Owner has identified three terms that require construction: "predefined regions," "copy number variation," and "unique sequence read." PO Resp. 9–22. We will address each of these terms in turn.

1. Predefined Regions

In our Decision to Institute, we determined that the term "predefined regions" is not limited to any specific number of base pairs, but encompasses any predefined region of a genome, including an entire chromosome. Inst. Dec. 8.

Patent Owner contends that the construction we adopted is in error as it does not make sense in the context of the claims. PO Resp. 11. Patent Owner contends that the term in step (e) of claim 1 calling for normalizing the number of reads "in the plurality of predefined regions" does not make sense if the predefined region is a chromosome. *Id.* According to Patent Owner, step (e) excludes reading the term to include chromosomes as the term calls for looking for copy number variation *within* the chromosome, not *between* chromosomes. *Id.* at 11–12. Patent Owner contends that the step of normalizing the reads within a chromosome would yield information about the relative size of the chromosomes, but would not yield information about the number of chromosomes. *Id.* at 12–13 (citing Ex. 2023 ¶ 43).

Patent Owner also contends that the Specification supports a narrower construction of the term "predefined regions." PO Resp. 12–13. Patent Owner points to the passage in the Specification where it states that "[i]n some embodiments the predefined regions are uniform or substantially

uniform in size. In some embodiments, the predefined regions are at least about 10 kb, 20 kb, 30 kb[,] 40 kb, 50 kb, 60 kb, 70 kb, 80 kb, 90 kb, or 100 kb in size. In some embodiments, at least 50, 100, 200, 500, 1000, 2000, 5000, 10,000, 20,000, or 50,000 regions are analyzed." PO Resp. 13; Ex. 1001, col. 16, ll. 7–12. Patent Owner acknowledges that the Specification refers to these kilobase ranges as specific embodiments but contends that, when read in context, the Specification supports Patent Owner's construction. PO Resp. 14.

Patent Owner contends that while the portions of the Specification cited in our Decision to Institute refer to regions in the genome without any limitation on the number of base pairs, those portions of the Specification do not support interpreting the term to embrace entire chromosomes. *Id.* at 14–15.

Petitioner responds by contending that Patent Owner's proposed construction would result in reading a limitation from a specific embodiment into the claim, something our reviewing court has held to be improper. Pet. Reply 10 (citing *Phillips v. AWH Corp.*, 415 F.3d 1303, 1323 (Fed. Cir. 2005); *Cont'l Circuits LLC v. Intel Corp.*, 915 F.3d 788, 796–96 (Fed. Cir. 2019)).

Petitioner also contends that the Board's preliminary construction was proper and supported by the Specification. Pet. Reply 10–11. Petitioner contends that the definition adopted by the Board is consistent with a person of ordinary skill in the art's understanding of the term. *Id.* at 11 (citing Ex. 1002 ¶¶ 88, 96; Ex. 2022, 71–72).

Petitioner contends that the language of the claims does not support Patent Owner's contention that "predefined region" cannot be a complete chromosome. Pet. Reply 11–12. Petitioner points out that the actual

language of the claims refers to normalizing reads "in a plurality of predefined regions" and that this would include a set of chromosomes. *Id.* Petitioner contends that Dr. Bustamante testified that step (e)(i) does not require normalization to a single region but includes determining copy number variation among the plurality of regions. *Id.* (citing Ex 1050, 168).

We have considered the arguments advanced by the parties and the evidence of record and we conclude that we did not err in our original construction of the term "predefined regions." The portions cited by Patent Owner in support of its construction make clear that the base pair discussion is directed to specific embodiments and not the invention as a whole. *See, e.g.*, Ex. 1001, col. 16, ll. 9–10 (stating "[i]n some embodiments" the predefined regions are specific sizes). Patent Owner does not point to, nor do we discern, anything in the intrinsic evidence to support a conclusion that Patent Owner intended this narrow definition to apply to all embodiments. Rather, as we found in our Decision to Institute, the Specification teaches that predefined regions can be any region in a genome. Inst. Dec. 7–8 (citing Ex. 1001, col. 5, l. 22, col. 10, l. 50; col. 17, l. 32, col. 25, ll. 2–3). And Patent Owner acknowledges that a chromosome is a part of a genome. PO Resp. 14–15.

We are unpersuaded by Patent Owner's argument that interpreting the term "predefined regions" would not make sense in the context of step (e)(i), which calls for normalization of the reads in the plurality of predefined regions to each other. PO Resp. 11–12. As Petitioner points out, Patent Owner's argument is premised on normalizing reads in each of the predefined regions to each other, i.e., on the same chromosome. Pet. Reply 12; PO Resp. 11–13. Claim 1, however, reads normalizing "a number of reads in the plurality of predefined regions to each other," which includes

TEMPUS EX1002.126

normalizing among multiple predefined regions, such as a set of chromosomes. Indeed, Dr. Bustamante admits that step (e)(i) is not limited to normalizing in a single region but includes determining copy number variation among a plurality of regions. Ex. 1050, 168.

2. Copy Number Variation

In our Decision to Institute, we construed the term "copy number variation" (CNV) to include aneuploidy. Inst. Dec. 11.

Patent Owner contends that we erred in construing the term to include aneuploidy, because the Specification teaches that CNV and aneuploidy are two separate and distinct genetic variations. PO Resp. 17. Patent Owner also contends that construing the term to exclude aneuploidy is consistent with Patent Owner's proposed construction of predefined region being a region smaller than a chromosome. *Id.*

Patent Owner contends that its proposed construction is consistent with how the term is defined in the art. *Id.* at 20–21. Patent Owner contends that the COSMIC[2] database teaches that copy number variations occur on a chromosome. *Id.* at 20 (citing Ex. 2035, 1). Patent Owner argues that Loo[3] distinguishes between ploidy and copy number variation. *Id.* at 21.

We have considered the arguments of the parties and the evidence of record and conclude that we did not err in construing the term "copy number variation" to include aneuploidy.

---

[2] The Catalogue of Somatic Mutations in Cancer (COSMIC) is a database containing information on somatic mutations in human cancer. Ex. 1058, 1.
[3] Loo et al., *Allele-specific copy number analysis of tumors,* 107 PROC. NAT. ACAD. SCI. 16910 (2010) (Ex. 2036, "Loo").

The Specification refers to aneuploidy as a type of copy number variation that can be detected by the disclosed method. The '743 patent states:

> For example, for quantitative analysis, a measure, e.g., a count, of the amount of tagged parent molecules mapping to a reference sequence is determined. This measure can be compared with a measure of tagged parent molecules mapping to a different genomic region. That is, the amount of tagged parent molecules mapping to a first location or mappable position in a reference sequence, such as the human genome, can be compared with a measure of tagged parent molecules mapping to a second location or mappable position in a reference sequence. This comparison can reveal, for example, the relative amounts of parent molecules mapping to each region. *This, in turn, provides an indication of copy number variation for molecules mapping to a particular region*. For example, if the measure of polynucleotides mapping to a first reference sequence is greater than the measure of polynucleotides mapping to a second reference sequence, *this may indicate that the parent population, and by extension the original sample, included polynucleotides from cells exhibiting aneuploidy*.

Ex. 1001, col. 35, ll. 8–31 (emphases added).

During oral argument counsel for Patent Owner was asked about whether this passage taught that aneuploidy fell within the definition of copy number variation as the term is used in the '743 patent. Tr. 47–49. Counsel for Patent Owner responded that this section of the '743 patent actually distinguishes between aneuploidy and copy number variation. Tr. 72–73. Specifically, Patent Owner explained:

> [I]t's a comparison. So, the reasoning is that if you have two chromosomes and you have a predefined region of 50 being mapped to Chromosome A and you have counts for 100 being mapped to Chromosome B, you don't know if it's a predefined region or if it's simply because Chromosome B is duplicated or

> aneuploidy. So that is what that is trying to do, is a comparison of the relative amounts of the parent molecules, which provides the copy number variation for the molecules mapping to a particular region. And then, it goes on to measure the polynucleotides. They're comparing, then, to the referenced sequences. And I think the goal there is, by extension, the original sample included polynucleotides from cells exhibiting aneuploidy. So, that's the point, is to exclude that possibility.

Id.

We are unpersuaded by this argument. We find that the Specification teaches that the method described in the '743 patent can be used to detect aneuploidy, which is an example of copy number variation. The Specification explains that comparing the relative amounts of parent molecules mapping to two different locations can "provide[] an indication of copy number variation for molecules mapping to a particular region." Ex. 1001, col. 35, ll. 20–22. The Specification then provides an example of this, where the amount of polynucleotides mapping to a first reference sequence is greater than the amount of polynucleotides mapping to a second reference sequence. *Id.* at col. 35, ll. 22–25. According to the Specification, "this may indicate that the parent population, and by extension the original sample, included polynucleotides from cells exhibiting aneuploidy." *Id.* at col. 35, ll. 25–27. Thus, we read this paragraph as teaching that comparing relative amounts of parent molecules can provide an indication of copy number variation, and an example of this may indicate that the parent population includes cells with aneuploidy. In other words, the Specification teaches that aneuploidy is a type of copy number variation.

Patent Owner argues that the Specification teaches that copy number variation and aneuploidy are distinct types of mutations. PO Resp. 18. In support of this contention, Patent Owner cites to the passages in the

Specification that list various genetic aberrations that can be detected by the disclosed method. *Id.* (citing Ex. 1001, col. 52, ll. 15–20). Patent Owner contends that, because copy number variation and aneuploidy are listed separately, one skilled in the art would consider them to be two distinct disorders. *Id.*

We are not persuaded by this argument. While the Specification lists copy number variation and aneuploidy separately, as we noted in our Decision to Institute, many of the terms listed in the cited paragraph are synonyms or subsets of each other. Inst. Dec. 9–11. The Specification states that the method disclosed therein may detect any number of genetic aberrations, including, but not limited to:

> mutations, rare mutations, indels, copy number variations, transversions, translocations, inversion, deletions, aneuploidy, partial aneuploidy, polyploidy, chromosomal instability, chromosomal structure alterations, gene fusions, chromosome fusions, gene truncations, gene amplification, gene duplications, chromosomal lesions, DNA lesions, abnormal changes in nucleic acid chemical modifications, abnormal changes in epigenetic patterns, abnormal changes in nucleic acid methylation infection and cancer.

Ex. 1001, 52:17–26. Arguably, most, if not all, of the genetic aberrations listed after "mutations" are types of mutations. And Patent Owner admitted that copy number variations are a type of rare mutation. Prelim. Resp. 16 ("Claims 1 and 10 . . . are directed to methods for detecting rare mutation, such as copy number variation . . . ."). Thus, Patent Owner's argument that the Specification distinguishes between aneuploidy and copy number variation because it lists them as separate genetic aberrations is not persuasive.

The extrinsic evidence cited by Patent Owner does not compel a different interpretation. The page from the COSMIC database is an explanation of a specific tab used in the COSMIC system and is not a definition of the term CNV. Ex. 2035, 1 ("This tab shows an overview of the data for the specified CNV."). Loo discusses an analysis of a subset of copy number events—allele-specific copy number variation which include aneuploidy. Ex. 2036, 1; *see also* Ex. 1050, 39–40. We agree with Petitioner that neither of these references excludes aneuploidy from the definition of copy number variation. Pet. Reply 9.

Moreover, the evidence of record shows that those skilled in the art interpret the term copy number variation to include aneuploidy. For example, Sehnert[4] teaches, "Because we are collecting information across the entire genome, this method is capable of detecting any aneuploidy or other copy number variation including insertions and deletions." Ex. 1026, 1048. Similarly, Casbon[5] states, "Copy number variations can include genomic rearrangements such as deletions, duplications, inversions, and translocations or whole chromosome aneuploidies." Ex. 1056, 24. For these reasons, we find that the intrinsic and extrinsic evidence support our construction of "copy number variation" as including aneuploidy.

3. Unique Sequence Read

While neither party initially proposed a construction for the term "unique sequence read," in its Preliminary Response, Patent Owner argued

---

[4] Sehnert et al., *Optimal Detection of Fetal Chromosomal Abnormalities by Massively Parallel DNA sequencing of Cell-Free Fetal DAN from maternal Blood,* 57 CLIN CHEM 1042 (2011) (Ex. 1026 "Sehnert").
[5] Casbon et al., WO 2012/038839 A2, published March 29, 2012, (Ex. 1056 "Casbon").

TEMPUS EX1002.131

that the terms "sequence reads" and "unique sequence reads," as used in the claims, are directed to two separate embodiments. Prelim. Resp. 23–28. In the Decision to Institute, we agreed with Patent Owner and construed the term "unique sequence read" to refer to "a sequence read that contains unique sequences, such that the sequences can be identified without the need for a barcode." Inst. Dec. 12.

Petitioner contends that we erred in our construction, specifically the requirement that the term excludes the use of a barcode to identify the sequence. Pet. Reply. 12–15. Petitioner contends that the term should be construed to mean a sequence that can be mapped to just one location. *Id.* at 14.

We have considered Petitioner's argument and the evidence of record and conclude that we did not err in our construction of the term "unique sequence read."

As we noted in our Decision to Institute, while the Specification does not define "unique sequence reads," the Specification teaches:

> Additionally, individual sequences may be identified based upon characteristics of sequence data for the read themselves. For example, the detection of unique sequence data at the beginning (start) and end (stop) portions of individual sequencing reads may be used, alone or in combination, with the length, or number of base pairs of each sequence read unique sequence to assign unique identities to individual molecules. Fragments from a single strand of nucleic acid, having been assigned a unique identity, may thereby permit subsequent identification of fragments from the parent strand. This can be used in conjunction with bottlenecking the initial starting genetic material to limit diversity.
>
> Further, using unique sequence data at the beginning (start) and end (stop) portions of individual sequencing reads

> and sequencing read length may be used, alone or [in] combination, with the use of barcodes.

Ex. 1001, col. 37, ll. 37–53. The discussion above, that the unique sequence data can be used alone or with barcodes, supports our conclusion that a unique sequence read is distinct from a sequence with a barcode.

Based on the foregoing we conclude that our initial construction of the term "unique sequence reads" was correct. "Unique sequence reads" as used in the '743 patent means a sequence read that contains unique sequences, such that the sequences can be identified without the need for a barcode.

D. Ground 1 – Anticipation by Chiu

Petitioner contends that claims 1–3 and 22–26 are unpatentable under 35 U.S.C. § 102(a) as anticipated by Chiu. Pet. 20–31.

1. Chiu

Chiu discloses the use of "massively parallel genomic sequencing to quantify maternal plasma DNA sequences for noninvasive prenatal detection of fetal trisomy 21." Ex. 1022, Abst. Chiu discloses a method comprising the steps of: obtaining a sample of human DNA from maternal plasma; sequencing the isolated DNA; aligning the DNA sequences to the human genome; counting the number of unique sequences; calculating the percentage of unique sequences; and calculating the Z-score for each chromosome and test sample. Ex. 1022, Fig. 1.

2. Analysis of Claim 1

Claim 1 is directed to a method for detecting copy number variation. Petitioner contends that each of the method steps is disclosed by Chiu. Pet. 20. Petitioner points to the chart shown in Figure 1 of Chiu as showing the steps recited in claim 1. *Id.*

TEMPUS EX1002.133

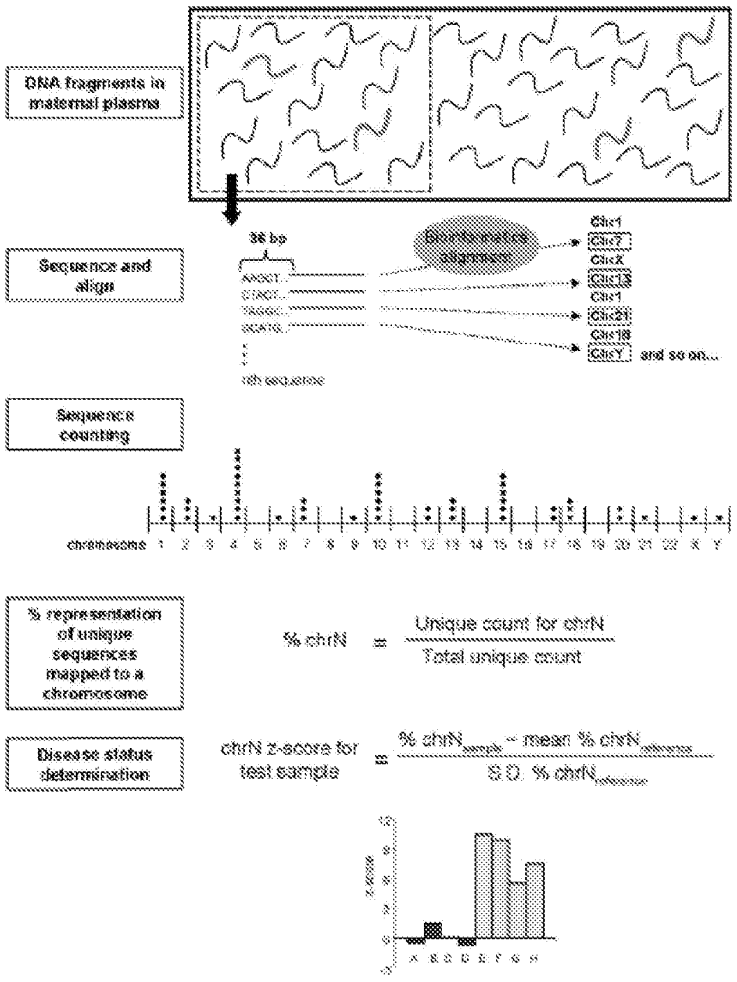


Figure 1 of Chiu outlines the method described therein. Ex. 1022, Fig. 1.

a) Method for detecting copy number variation

Petitioner contends that this limitation is met in that Chiu discloses a method for detecting aneuploidy. Pet. 17, 22.

Patent Owner contends that this limitation has not been met because the term copy number variation, as used in the '743 patent, does not include aneuploidy. PO Resp. 22–24. Patent Owner reiterates the arguments made with respect to the construction of the term copy number variation discussed above. *Id.*

As discussed above, we have construed the term copy number variation to include aneuploidy. Section II.C.2 *supra.* Chiu discloses a method for detecting fetal chromosomal aneuploidy. Ex. 1022, 1. We find that Chiu discloses the recited limitation.

b) Sequencing extracellular polynucleotides from a bodily sample from a subject, wherein each of the extracellular polynucleotides generates a plurality of sequence reads

Petitioner contends that Chiu discloses this limitation as Chiu discloses amplifying the DNA fragments prior to sequencing such that multiple copies of the fragments are generated. Pet. 23 (citing Ex. 1002 ¶ 80; Ex. 1022, SI1).

Patent Owner does not contest that Chiu discloses this limitation. *See* PO Resp. 23–25.

We find that Chiu discloses this limitation. Chiu discloses extraction of DNA from plasma, which is an extracellular polynucleotide. Ex. 1022, SI1; Ex. 1002 ¶ 80. Chiu goes on to disclose that the DNA was then amplified using 15-cycle PCR and that the DNA was sequenced. Ex. 1022, SI1; Ex. 1002 ¶ 80.

c) Filtering out reads that fail to meet a set accuracy, quality score, or mapping score threshold

Petitioner contends that Chiu discloses this limitation. Pet. 23–24. Petitioner argues that Chiu discloses that only sequences that could be mapped to just one location were counted and that this meets the criteria for the filtering step. *Id.* (citing Ex. 1002, 20459; Ex. 1002 ¶¶ 83–84).

Patent Owner does not contest that Chiu discloses this limitation. *See* PO Resp. 23–25.

We find that Chiu discloses this limitation. As Dr. Gabriel testifies and we agree, one skilled in the art would have understood that a mapping score is "a metric used to indicate whether a sequence read maps to a single

location in the genome or multiple locations in the genome" and that this definition is consistent with the teachings of the '743 patent. Ex. 1002 ¶ 82.

As Chiu discloses

> In this study, we counted only sequences that could be mapped to just one location in the repeat-masked reference genome with no mismatch, i.e., deemed as a "unique" sequence in the human genome. We termed these sequences as U0-1-0-0 on the basis of values in a number of fields in the data output files of the ELAND sequence alignment software (Illumina).

Ex. 1022, 20459. Dr. Gabriel testifies, and we agree, that one skilled in the art would have understood this passage of Chiu to describe the filtering step recited in claim 1. *See* Ex. 1002 ¶ 84.

d) Mapping the plurality of sequence reads to a reference sequence

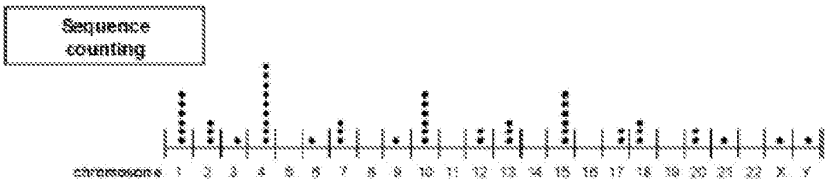
Petitioner contends that Chiu discloses this limitation. Petitioner contends that Chiu discloses identifying the chromosomal origin of each sequence through mapping to a human reference genome. Pet. 31 (citing Ex. 1022, Fig. 1).

Patent Owner does not contest that Chiu discloses this limitation. *See* PO Resp. 23–25.

We find that Chiu discloses this limitation. Chiu states, "the chromosomal origin of each 36-bp sequence was identified through mapping to the human reference genome by bioinformatics analysis." Ex. 1022, 20459, Fig. 1. Dr. Gabriel testifies, and we agree, that this disclosure meets the claim limitation calling for mapping the sequence reads to a reference sequence. Ex. 1002 ¶ 85.

e) Quantifying mapped reads or unique sequence reads in a plurality of predefined regions of the reference sequence

Petitioner contends that Chiu discloses this limitation. Pet. 25–26. Referring to part of Figure 1 reproduced below, Petitioner contends that Chiu teaches "the number of sequence reads originating from any particular chromosome was then counted and tabulated for each human chromosome." *Id.* (quoting Ex. 1022, 20459.)



Portion of Fig. 1 of Chiu showing Sequence Counting.

Patent Owner contends that Chiu does not disclose this step, arguing that Petitioner has improperly equated a predefined region with an entire chromosome. PO Resp. 23–24.

We have considered the arguments presented by the parties and the evidence of record and find that Chiu discloses quantifying mapped reads or unique sequence reads in a plurality of predefined regions of the reference sequence. Chiu teaches that the sequence reads, which originate from a particular chromosome, are counted and tabulated. Ex. 1022, 20459. Chiu also states that the method described in the article is designed to "quantify the amount of chr21 sequences in maternal plasma." *Id.* We agree with Dr. Gabriel that the counting and tabulating procedure described in Chiu meets the claim limitation calling for quantifying mapped regions in a plurality of predefined regions of the reference sequence. Ex. 1002 ¶¶ 86–88.

With respect to Patent Owner's contention that a predefined region does not include whole chromosomes, as discussed above we do not find that the term is so limited.

f) Determining copy number variation in one or more of the plurality of predefined regions by:

Petitioner contends that Chiu discloses this limitation in that Chiu is directed to a method of diagnosing aneuploidy using both a normalizing step and a processing step. Pet. 26. Petitioner refers back to its discussion of the construction of the term "predefined regions". *Id.*

Patent Owner contends that this limitation is not disclosed by Chiu because Chiu does not determine copy number variation in one or more of the plurality of predefined regions. PO Resp. 24. Patent Owner contends that the term "in one or more of the plurality of predefined regions" limits the claim to copy number variations within a region and thus excludes copy number variation based on an entire chromosome. *Id.*

We have considered the arguments advanced by the parties and the evidence of record and find that Chiu discloses this limitation. The limitation at issue reads "determining copy number variation in one *or* more of the plurality of predefined regions." Ex. 1001, 62:56–57 (emphasis added). As discussed above, we have construed the term "predefined regions" to include whole chromosomes. See Section II.C.1, *supra.* Thus, the limitation at issue embraces determining copy number variation both within a chromosome and between or among chromosomes. Chiu discloses determining copy number variation among chromosomes, e.g., aneuploidy. Ex. 1022, 20459; *see* Ex. 1002 ¶ 89.

TEMPUS EX1002.138

g) Normalizing a number of reads in the plurality of predefined regions to each other, or a number of unique sequence reads in the plurality of predefined regions to each other; and/or

Petitioner contends that Chiu discloses this limitation. Petitioner contends that Chiu discloses:

> We then determined the percentage contribution of unique sequences mapped to each chromosome by dividing the U0-1-0-0 count of a specific chromosome by the total number of U0-1-0-0 sequence reads generated in the sequencing run for the tested sample to generate a value termed % chrN, when the chromosome of interest is chrN.

Pet. 27 (quoting Ex. 1022, 20459). Petitioner contends that this represents a normalization of the data so that unique sequence reads can be compared to one another. *Id.*

Patent Owner contends that Chiu does not disclose this limitation. PO Resp. 25–26. Patent Owner contends:

> When different chromosomes are normalized to each other (each being a "predefined region" according to the petition), even with Z-score processing, information relates to the population of chromosomes in general, and not to any one specific chromosome. Thus, meaningful information regarding aneuploidy would not be obtained. EX2023, ¶61. Petitioner has not demonstrated otherwise.

Id.

We have considered the arguments advanced by the parties and the evidence of record and conclude that Chiu discloses the normalization step of claim 1. Chiu discloses calculating a % chrN value for each chromosome by comparing the number of sequence reads for a given chromosome to the total number of sequence reads. Ex. 1022, 20459. This allows the sequence reads for each chromosome to be compared with each other as well as allowing comparison to the total sequence reads. Ex. 1002 ¶ 92. This

comparison step meets the requirement of normalization. *See* Ex. 1050, 52–53, 151.

Patent Owner's argument that Chiu does not teach the normalization step is unpersuasive. Patent Owner contends that if chromosomes were the predefined regions, the normalization step would not yield any meaningful data regarding aneuploidy. PO Resp. 25. However, Chiu specifically teaches that the data from the normalization step can be used to determine if aneuploidy is present. *See* Ex. 1022, 29459 (stating normalization data was then used to calculate a z-score, which is then used to determine aneuploidy). We find that this disclosure adequately teaches the "normalizing" step.

h) Processing a number of reads in the plurality of predefined regions or a number of unique sequence reads in the plurality of predefined regions with numbers obtained from a control sample

Petitioner contends that Chiu discloses this limitation. Pet. 28–30. Petitioner contends that Chiu discloses calculating the z-score of % chr21 of a test sample, which represents the number of standard deviations from the mean of a reference data set. *Id.* at 29. Petitioner contends that this is equivalent to processing a number of unique reads in a plurality of predefined regions with the numbers obtained from a control sample. *Id.* (citing Ex. 1002 ¶ 96).

Patent Owner contends that Chiu does not disclose this limitation. PO Resp. 25–26. Patent Owner reiterates the arguments made with respect to step (e)(i) (the "normalizing" step) above. *Id.*

We have considered the arguments advanced by the parties and the evidence of record and find that Chiu discloses the processing step of claim 1. Chiu discloses:

TEMPUS EX1002.140

> To determine if a tested maternal plasma sample belonged to a T21 pregnancy, we calculated the *z*-score of % chr21 of the tested sample. The *z*-score refers to the number of standard deviations from the mean of a reference data set. Hence, for a T21 fetus, a high *z*-score for % chr21 was expected when compared with the mean and standard deviation of % chr21 values obtained from maternal plasma of euploid pregnancies.

Ex. 1022, 20459. We agree with Dr. Gabriel that "[b]y comparing that value in a tested sample to that value in a control sample, Chiu describes 'processing a number of unique sequence reads in the plurality of predefined regions with numbers obtained from a control sample.'" Ex. 1002 ¶ 96.

Patent Owner's arguments regarding the processing step are unpersuasive for the reasons stated above regarding the normalization step.

3. Claim 2

Claim 2 reads:

> The method of claim 1, further comprising isolating extracellular polynucleotides from the bodily sample.

Ex. 1001, col. 62, ll. 66–67.

Petitioner contends that Chiu discloses this limitation. Pet. 30.

Patent Owner offers no separate argument regarding the patentability of claim 2. *See* PO Resp. 22–28.

We find that Chiu discloses the step of claim 2. Chiu states, "DNA from maternal plasma was extracted with the blood and bodily fluid protocol of the QIAmp DSP DNA blood mini kit (Qiagen)." Ex. 1022, SI1; Ex. 1002 ¶ 97. We find that this disclosure adequately teaches the claimed limitation of "isolating extracellular polynucleotides from the bodily sample."

4. Claim 3

Claim 3 reads:

TEMPUS EX1002.141

> The method of claim 1, further comprising generating copies of the extracellular polynucleotides prior to sequencing.

Ex. 1001, col. 63, ll. 1–3.

Petitioner contends that Chiu discloses this step. Pet. 30.

Patent Owner offers no separate argument regarding the patentability of claim 3. *See* PO Resp. 22–28.

We find that Chiu discloses the step of claim 3. Chiu discloses that the selected DNA libraries were amplified, followed by sequencing. Ex. 1022, SI1, 20459 Figure 1; Ex. 1002 ¶ 98. Thus, we find that this disclosure adequately teaches "generating copies of the extracellular polynucleotides prior to sequencing."

5. Claim 22

Claim 22 reads:

> The method of claim 1, wherein e) comprises i) and ii).

Ex. 1001, col. 64, ll. 33–34.

Petitioner contends that Chiu discloses the method of claim 22 in that Chiu discloses both steps i) and ii).

Patent Owner offers no separate argument regarding the patentability of claim 22. *See* PO Resp. 22–28.

We find that Chiu discloses the method of claim 22. As discussed above, Chiu discloses normalization of the sequence reads followed by processing the reads with numbers from a reference sample. Ex. 1022, 20459.

6. Claims 23–26

Claims 23 through 26 read:

> 23. The method of claim 1, wherein e) comprises i) normalizing the number of reads in the plurality of predefined regions to each other.
>
> 24. The method of claim 1, wherein e) comprises i) normalizing the number of unique sequence reads in the plurality of predefined regions to each other.
>
> 25. The method of claim 1, wherein e) comprises ii) processing the number of reads in the plurality of predefined regions with numbers obtained from the control sample.
>
> 26. The method of claim 1, wherein e) comprises ii) processing the number of unique sequence reads in the plurality of predefined regions with numbers obtained from the control sample.

Ex. 1001, col. 64, ll. 38–47.

Petitioner contends that Chiu discloses the elements of claims 23–26 in that Chiu teaches the use of sequence reads and unique sequence reads. Pet. 31.

Patent Owner offers no separate argument regarding the patentability of claims 23–26. *See* PO Resp. 22–28.

We find that Chiu discloses the elements of claims 23–26. Chiu discloses that the method described uses only sequences that could be mapped to a single location and describes these sequences as unique. Ex. 1022, 20459, Fig. 1. Dr. Gabriel testifies, and we agree, that the filtering process used in Chiu produces unique sequences. Ex. 1002 ¶ 84. Unique sequence reads are a subset of reads, and thus the sequence reads used in Chiu satisfy both the requirement for sequence reads and unique sequence reads. *See* Ex. 1002 ¶ 96 (unique sequence reads are reads in the plurality of predefined regions).

7. Summary

For all the above reasons, we determine that Petitioner has shown by a preponderance of the evidence that Chiu discloses each and every limitation of claims 1–3 and 22–26. Thus, we find claims 1–3 and 22–26 unpatentable for anticipation by Chiu.

E. Ground 2 – Obviousness based on Kinde 2012 and Kinde 2011

Petitioner contends that the subject matter of claims 1–9, 20, and 22–26 would have been obvious to a person of ordinary skill in the art at the time the invention was made over Kinde 2012 combined with Kinde 2011.

1. Kinde 2012

Kinde 2012 teaches a method for detecting fetal aneuploidy using cell-free DNA from a sample of maternal plasma. Ex. 1041, Abst. The method of Kinde 2012 includes selecting DNA fragments from chromosomal regions, amplifying and sequencing the DNA fragments to generate multiple sequence reads for each fragment, filtering the reads for quality, mapping the reads to the human genome, counting the reads that map to each chromosomal region, normalizing the reads across regions, and comparing the normalized figures to a control sample. Ex. 1041, 2–5.

2. Kinde 2011

Kinde 2011 teaches a method for "identif[ying] mutations that are present in a small fraction of DNA templates." Ex. 1023, Abst. In the method of Kinde 2011, each DNA fragment to be analyzed is assigned a unique sequence identifier. *Id.* at 9530. The tagged DNA fragments are then amplified to create a number of daughter fragments. *Id.* The daughter fragments are then sequenced. *Id.*

3. Analysis of Claim 1

Claim 1 is directed to a method for detecting copy number variations. Petitioner contends that Kinde 2012 combined with Kinde 2011 teach or suggest all of the steps of the method recited in claim 1.

a) A method for detecting copy number variation comprising:

Petitioner contends that Kinde 2012 teaches a method for detecting fetal aneuploidy, which is a type of copy number variation. Pet. 34. Petitioner also contends that it would have been obvious to one skilled in the art to use the method of Kinde 2012 to detect other forms of copy number variation. *Id.*

Patent Owner contends that Kinde 2012 is not directed to detecting copy number variation because the term copy number variation, as used in the context of the '743 patent, does not encompass aneuploidy. PO Resp. 33–35. Patent Owner reiterates the same arguments presented concerning construction of the term copy number variation. *Id.*

We have considered the argument presented by the parties and the evidence of record and conclude that Kinde 2012 teaches a method for determining copy number variation. Kinde 2012 teaches an improved sequencing method that can be used as part of a "safe and effective screening method for fetal aneuploidies." Ex. 1041, Abstr. As discussed above, for purposes of this Decision, we have construed the term "copy number variation" to include aneuploidy. Therefore, Kinde 2012 teaches a method for determining copy number variation. Ex. 1002 ¶ 107.

TEMPUS EX1002.145

b) Sequencing extracellular polynucleotides from a bodily sample from a subject, wherein each of the extracellular polynucleotides generates a plurality of sequence reads

Petitioner contends that Kinde 2012 combined with Kinde 2011 teaches this step of the claim. Pet. 34–35. Petitioner contends that Kinde 2012 teaches an improved method for sequencing cell-free, maternal DNA. *Id.* at 34. Petitioner contends that Kinde 2012 uses the amplification technique disclosed in Kinde 2011. *Id.* at 34–35.

Patent Owner does not present a separate argument concerning this step of the claim. *See* PO Resp. 28–46.

We find that Kinde 2012 combined with Kinde 2011 teaches this step of the claimed method. Kinde 2012 teaches a method for sequencing cell-free, maternal plasma DNA, which is an extracellular polynucleotide. Ex. 1041, Abst.; Ex. 1002 ¶ 108. Kinde 2012 teaches that the DNA fragments were amplified using the method described in Kinde 2011. Ex. 1041, 2.

As shown below, Figure 1 of Kinde 2011 teaches that the fragments were first assigned a unique identification DNA sequence followed by amplifying each unique fragment to create a plurality of fragments. Ex. 1023, 9530, Fig. 1.

TEMPUS EX1002.146

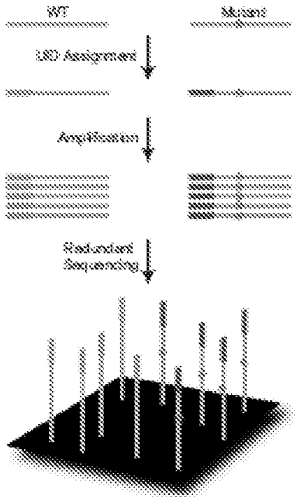


Fig. 1. Essential elements of Safe-Seq. In the first step, each fragment to be analyzed is assigned a unique identification (UID) DNA sequence (green or blue bars). In the second step, the uniquely tagged fragments are amplified, producing UID families, each member of which has the same UID. A supermutant is defined as a UID family in which ≥95% of family members have the same mutation.

Figure 1 from Kinde 2011, Ex. 1023 showing the tagging, amplification and filtering steps.

Based on the foregoing, we agree with Dr. Gabriel that Kinde 2012 combined with Kinde 2011 teaches the step of "sequencing extracellular polynucleotides from a bodily sample from a subject, wherein each of the extracellular polynucleotides generates a plurality of sequence reads." Ex. 1002 ¶¶ 108–110.

c) Filtering out reads that fail to meet a set accuracy, quality score, or a mapping score threshold

Petitioner contends that the disclosure of Kinde 2011 satisfies this limitation because Kinde 2011 teaches that only high quality reads that meet a set criteria are used in the analysis. Pet. 43 (citing Ex. 1009, SI2).

Patent Owner does not present a separate argument concerning this step of the claim. *See* PO Resp. 28–46.

We find that Kinde 2011 teaches this step of claim 1. Kinde 2011 teaches:

> Only high-quality reads meeting the following criteria were used for subsequent analysis: (i) the first 25 bases passed the standard Illumina chastity filter; (ii) every base in the read had a quality score ≥20; and (iii) ≤3 mismatches to expected sequences. For the exogenous UID libraries, we additionally required the UIDs to have a quality score ≥30.

Ex. 1009, SI2. We agree with Dr. Gabriel that this passage describes filtering out reads that fail to meet a set quality score threshold. Ex. 1002 ¶ 111.

d) Mapping the plurality of sequence reads to a reference sequence

Petitioner contends that Kinde 2012 teaches this step of the method of claim 1 in that Kinde 2012 teaches that the distinct sequences were mapped to the human genome. Pet. 37–38.

Patent Owner contends that Kinde 2012 does not teach or suggest this step of the claimed method. PO Resp. 36–38. Patent Owner contends that Kinde 2012 teaches that it does not map the sequences to the human genome but instead maps the sequences to a discrete set of chromosomal positions. *Id.* at 37. Patent Owner contends that this is mapping the sequences to predefined regions and not mapping to a reference sequence. *Id.*

We have considered the arguments of the parties and the evidence of record and conclude that Kinde 2012 teaches the step of mapping the plurality of sequences to a reference sequence. While we agree with Patent Owner that Kinde 2012 does not map the sequences against the entire human genome, claim 1 only requires mapping to "a reference sequence." Ex. 1001, col. 62, ll. 51–52. The Specification of the '743 patent refers to "the human genome" as an example of "a reference sequence," and states that reads can

be mapped to particular locations or positions in the reference sequence. Ex. 1001, col. 35, ll. 14–22. Thus, nothing in claim 1 or the Specification requires the "reference sequence" to be the entire human genome. Similarly, Kinde 2012 teaches that the sequences are aligned against a discrete set of chromosomal positions within the reference sequence. Ex. 1041, 2. Thus, we find Kinde 2012 teaches mapping a plurality of sequence reads to the reference sequence (i.e., the human genome) by aligning reads to a discrete set of chromosomal positions within the genome. Ex. 1042, 2.

e) Quantifying mapped reads or unique sequence reads in a plurality of predefined regions of the reference sequence

Petitioner contends that Kinde 2012 teaches this step of the method claimed. Pet. 38. Petitioner contends that Kinde 2012 teaches generating counts for each chromosomal position. *Id.*

Patent Owner contends that Petitioner's argument concerning the definition of a reference sequence is inconsistent with its position of a predefined region used in other steps of Kinde 2012's method. PO Resp. 29. Patent Owner contends that, for this step, Petitioner relies on the L1 retrotransposon regions as the plurality of predefined regions, but with respect to the normalizing and processing steps, Petitioner relies on entire chromosomes as the predefined regions. *Id.* at 29–34. Patent Owner contends that this inconsistency precludes a finding that Kinde 2012 teaches the claimed method. *Id.*

We have considered the arguments advanced by the parties and the evidence of record and find that Kinde 2012 teaches the quantification step of claim 1. Kinde 2012 teaches "generating a table of tag counts for each sample where each row represents a unique position (note that all tables will be of equal length as only the shared positions in each experiment were

analyzed)." Ex. 1041, 4. The generation of tag counts is the same as quantifying mapped reads in a plurality of predefined regions of the reference sequence. Ex. 1002 ¶ 118.

We have considered Patent Owner's argument regarding the alleged inconsistency of Petitioner's analysis and are unpersuaded. As discussed above, the Specification teaches that mapping can be performed using only a portion of the reference sequence. Ex. 1001, col. 35, ll. 14–22 (teaching measuring the amount of polynucleotides mapping to a first and second location in a particular region of a human genome). Similarly, Kinde 2012 teaches mapping sequences to the retrotransposon regions, i.e., a portion of the chromosomes, because mapping sequences to a discrete set of chromosomal positions is simpler than aligning to the entire genome. Ex. 1041, 2; Ex. 1002 ¶ 115. Kinde 2012, therefore, uses portions of the reference sequence or predefined region to determine copy number variation. Ex.1041, 5. Thus, we find Kinde 2012 teaches quantifying mapped reads of unique sequence reads in a plurality of predefined regions of the reference sequence.

f) Determining copy number variation in one or more of the plurality of predefined regions by: i) normalizing a number of reads in the plurality of predefined regions to each other, or a number of unique sequence reads in the plurality of predefined regions to each other; and/or ii) processing a number of reads in the plurality of predefined regions or a number of unique sequence reads in the plurality of predefined regions with number obtained from a control sample

Petitioner contends that Kinde 2012 teaches these steps. Pet. 37–42. Petitioner contends that Kinde 2012 teaches both normalization and processing of the sequence read counts. *Id.*

Patent Owner contends that Kinde 2012 does not teach either of the steps. PO Resp. 29–33, 36–39. Patent Owner contends that Kinde 2012 does

not teach normalizing a plurality of predefined regions to each other but teaches normalizing the tag count across samples. PO Resp. 29–30.

We have considered the arguments presented by the parties and the evidence of record and find that Kinde 2012 teaches both the normalization and processing steps of claim 1. Kinde 2012 teaches that normalization was performed using the following steps:

> 1) generating a table of tag counts for each sample where each row represents a unique position (note that all tables will be of equal length as only the shared positions in each experiment were analyzed);
> 2) sorting the rows in each table based on tag counts, resulting in each table having a different order of positions;
> 3) determining the mean tag count for each row across all samples;
> 4) replacing an individual sample's tag count with the mean tag count for all samples at each row; and
> 5) sorting the tag counts for each sample's table back to their original order based on position.

Ex. 1041, 4. As stated above, each row used in the normalization process represents a unique position and mean tag counts are determined for each row or unique position. *Id.* We agree with Dr. Gabriel that normalizing tag counts for each position is the same as normalizing a number of reads in a plurality of defined regions to each other. Ex. 1002 ¶¶ 117–119. This conclusion is also supported by the testimony of Dr. Bustamante. Ex. 1050, 123–125 (testifying that Kinde 2012 teaches comparing tag counts within each sample, which is the same as normalizing the number of reads in the plurality of defined regions to each other).

Kinde 2012 goes on to teach

> A common method of determining the aneuploidy status of a particular sample in Digital Karyotyping-based assays is by comparison of z-scores. Through this method, one determines

> the mean and standard deviation of tag counts lying within a chromosome of interest in a group of reference samples (e.g., samples with known euploid content), and then creates a standardized score (i.e., z-score) for a chromosome of interest for each sample.

Ex. 1041, 4–5 (citations omitted). We agree with Dr. Gabriel that by teaching the steps of calculating z-scores for both the test sample and the control sample and then comparing the z-scores, Kinde 2012 teaches processing a number of reads or unique sequence reads in the plurality of predefined regions with numbers obtained from a control sample. Ex. 1002 ¶¶ 120–123.

We are unpersuaded by Patent Owner's argument that Kinde 2012 does not teach normalizing tag counts to each other but instead teaches normalizing the counts across samples. PO Resp. 29–30. As seen from the passage of Kinde 2012 quoted above, normalization is done for rows, which represent a unique sequence. Ex. 1041, 4. Dr. Gabriel testifies, and we agree, that normalizing tag counts for each position is the same as normalizing a number of reads in a plurality of predefined regions to each other. Ex. 1002 ¶¶ 117–119.

We are also unpersuaded by Patent Owner's argument that Petitioner has failed to show that Kinde 2012 teaches steps c), d), and e) of claim 1. As shown by the analysis discussed above, Kinde 2012 teaches all three steps. *See* Section II.E. 3.d–f, *supra.*

g) Motivation to combine/Reasonable expectation of success

Petitioner contends that one skilled in the art would have been motivated to combine the teachings of Kinde 2012 with those of Kinde 2011. Pet. 31–34. Petitioner contends that both Kinde references have the same authors and teach similar methods for detecting rare mutations and

copy number variations using massively parallel sequencing. *Id.* at 31. Petitioner also contends that Kinde 2012 incorporates the teachings of Kinde 2011. *Id.* at 32.

Patent Owner contends that Petitioner has failed to establish a motivation to combine the references. PO Resp. 39–41. Patent Owner contends that Petitioner has not shown why one skilled in the art would use the filtering technique of Kinde 2011 in the method of Kinde 2012. *Id.* at 40. Patent Owner also contends that Kinde 2011 discloses several different embodiments and there is nothing in either reference to teach or suggest which embodiment of Kinde 2011 to use. *Id.* at 41.

We have considered the arguments advanced by the parties and the evidence of record and we find that one skilled in the art would have been motivated to incorporate the teachings of Kinde 2011 into the method of Kinde 2012. Both Kinde 2011 and Kinde 2012 teach methods for detecting rare mutations and copy number variations using massively parallel sequencing. Ex. 1023, 9530; Ex. 1041, Abstract. Kinde 2012 specifically teaches using the library preparation techniques of Kinde 2011, including Kinde 2011's amplification and sequencing techniques. Ex. 1041, 2, 7. We also find that a person of ordinary skill in the art would have had a reasonable expectation of success. Specifically, Kinde 2012 successfully incorporated certain steps of Kinde 2011 into the method taught in Kinde 2012. Ex. 1002 ¶ 106. We find that this teaching of the successful incorporation of Kinde 2011 into Kinde 2012 would lead one skilled in the art to have a reasonable expectation of successfully accomplishing the same incorporation of Kinde 2011 into Kinde 2012 with respect to the steps of amplifying and tagging the polynucleotides.

TEMPUS EX1002.153

h) Conclusion

Based on the foregoing, we conclude that Petitioner has demonstrated by a preponderance of the evidence that the subject matter of claim 1 would have been obvious to one of ordinary skill in the art at the time of the invention and is therefore unpatentable.

4. Claim 2

Claim 2 reads

> The method of claim 1, further comprising isolating extracellular polynucleotides from the bodily sample.

Ex. 1001, col. 62, ll. 66–67

Petitioner contends that Kinde 2012 teaches this claim element. Pet. 42. Petitioner contends that Kinde 2012 teaches obtaining plasma samples from patients followed by preparing DNA for sequencing and that preparing the DNA for sequencing necessarily involves isolating the DNA from the plasma. *Id.*

Patent Owner offers no separate argument regarding the patentability of claim 2. *See* PO Resp. 41–46.

We find that Kinde 2012 teaches the step of isolating extracellular polynucleotides from a bodily sample. Kinde 2012 refers to testing maternal plasma DNA for aneuploidy. Ex. 1041, 2. Dr. Gabriel testifies, and we agree, that plasma is a bodily sample and one skilled in the art would have understood that in order to sequence the DNA in the plasma, the DNA would need to be isolated from the plasma. Ex. 1002 ¶ 127.

As discussed above, Petitioner has demonstrated a motivation to combine the Kinde references with a reasonable expectation of success. We, therefore, conclude that Petitioner has shown by a preponderance of the

evidence that the subject matter of claim 2 would have been unpatentable for obviousness.

5. Claim 3.

Claim 3 reads:

> The method of claim 1, further comprising generating copies of the extracellular polynucleotides prior to sequencing.

Ex. 1001, col. 63, ll. 1–3.

Petitioner contends that Kinde 2011 discloses this step in that it teaches making copies of the DNA fragments before sequencing. Pet. 42.

Patent Owner contends that Kinde 2011 does not teach or suggest generating copies of extracellular DNA prior to sequencing. PO Resp. 42. Patent Owner contends that Kinde 2011 relates to cellular DNA and not extracellular DNA. *Id.*

We have considered the arguments advanced by the parties and the evidence of record and conclude that Kinde 2012 in combination with Kinde 2011 teaches or suggests the step recited in claim 3. As shown in Figure 1 of Kinde 2011 reproduced above, Kinde 2011 teaches amplifying the DNA fragment prior to sequencing. Ex. 1023, 9531.

While we agree with Patent Owner that Kinde 2011 is directed to cellular DNA, Kinde 2012 is directed to extracellular DNA. *See* Ex. 1041, 1. Thus, it is the combination of the Kinde references that teaches or suggests the subject matter of claim 3.

As discussed above, Petitioner has demonstrated a motivation to combine the Kinde references with a reasonable expectation of success.

We conclude that Petitioner has established by a preponderance of the evidence that claim 3 is unpatentable for obviousness.

TEMPUS EX1002.155

6. Claim 4

Claim 4 reads:

> The method of claim 1, further comprising determining a percent of sequences having copy number variation or rare mutation or variant in the bodily sample.

Ex. 1001, col. 63, ll. 4–6.

Petitioner contends that this step is taught by Kinde 2011. Pet. 43. Petitioner contends that Kinde 2011 teaches creating oligonucleotide mixtures with predetermined percentages of mutant and normal fractions and then using the method of Kinde 2011 to analyze the mixtures. *Id.* The resulting ratios were reported in a percent of the mutant present in the mixture. *Id.*

Patent Owner contends that Petitioner has not explained why one skilled in the art would have used the percentage calculation disclosed in Kinde 2011 in the method disclosed in Kinde 2012. PO Resp. 43. Patent Owner also contends that Kinde 2012 does not relate to either copy number variations or rare mutations. *Id.*

We have considered the arguments presented by the parties and the evidence of record, and find that the combination of the Kinde references teaches the step recited in claim 4.

Kinde 2011 teaches:

> Does Safe-SeqS preserve the ratio of mutant:normal sequences in the original templates? To address this question, we synthesized two 31- base oligonucleotides of identical sequence with the exception of nucleotide 15 . . . and mixed them at nominal mutant/normal fractions of 3.3% and 0.33%. Through Safe-SeqS analysis of the oligonucleotide mixtures, we found that the ratios were 2.8% and 0.27%, respectively. We conclude that the UID assignment and amplification procedures used in Safe-SeqS do not greatly alter the proportion of variant

> sequences and thereby provide a reliable estimate of that proportion when unknown.

Ex. 1021, 9533. Dr. Gabriel testifies that one skilled in the art would have known that mutation frequency can be expressed as a percentage of the sequences containing a mutation. Ex. 1002 ¶ 133. Dr. Gabriel testifies, and we agree, that the percentage of mutations can be easily calculated from the data in Kinde 2011 "simply by dividing the number of 'total supermutants identified' by the number of 'UID families.'" Ex. 1002 ¶ 134.

With respect to the motivation to use the percentage calculation of Kinde 2011 in the method of Kinde 2012, as explained above, the Kinde articles address the same problem and Kinde 2012 specifically teaches using the method of Kinde 2011. Moreover, as Dr. Gabriel testifies and we agree, calculating the percentage of sequences containing a mutation was a common way of expressing mutation frequency at the time the invention was made. Ex. 1002 ¶¶ 133–135.

As to Patent Owner's argument regarding "a rare mutation," we note that the language of claim 4 recites "copy number variation or rare mutation or variant." Ex. 1001, col. 63, ll. 5–6. The method of claim 4 need only determine the percentage of sequences having one of these types of mutations to satisfy the claim step. As discussed above, Kinde 2012 determines aneuploidy, a form of copy number variation. Section II, E, 3 *supra*. Also, Kinde 2011 determines the percentage of mutant (or variant) sequences. Ex. 1021, 9533.

Based on the foregoing, we conclude that Petitioner has established by a preponderance of the evidence that the subject matter of claim 4 would have been obvious over Kinde 2012 combined with Kinde 2011.

7. Claims 5, 6, and 9

Claims 5, 6, and 9 read:

> 5. The method of claim 1, further comprising attaching one or more barcodes to the extracellular polynucleotides or fragments thereof prior to sequencing.
>
> 6. The method of claim 5, wherein each barcode attached to the extracellular polynucleotides or fragments thereof prior to sequencing is not unique.
>
> 9. The method of claim 1, further comprising attaching one or more barcodes to the extracellular polynucleotides or fragments thereof prior to an amplification or enrichment step.

Ex. 1001, col. 63, ll. 7–12 and 21–25.

Petitioner contends that Kinde 2011 teaches the steps recited in claims 5, 6, and 9 in that Kinde 2011 teaches attaching barcodes to DNA fragments prior to sequencing and amplification, and that the barcodes are not unique. Pet. 44.

Patent Owner contends that Kinde 2011 does not teach the steps of claim 5, 6, and 9 because the claims call for the barcodes to be attached to extracellular DNA and Kinde 2011 is directed to cellular DNA. PO Resp. 43–44. Patent Owner also contends that the embodiment of Kinde 2011 that teaches non-unique barcodes is not applicable to the DNA of Kinde 2012, because the embodiment of Kinde 2011 requires the presence of sheared ends and the extracellular DNA of Kinde 2012 does not possess sheared ends. *Id.* at 44–45 (citing Ex. 2020 ¶ 36).

We have considered the arguments presented by the parties and the evidence of record and find that Kinde 2012 combined with Kinde 2011 teach the steps of claims 5, 6, and 9.

Kinde 2012 relates to analysis of extracellular DNA to determine aneuploidy. Ex. 1041, Abstr. Kinde 2012 also teaches that the DNA

fragments were "tagged with a unique identifier DNA sequence" and that amplification was done using the method of Kinde 2011. Ex. 1041, 2. Kinde 2011 teaches that tagging was done before amplification. Ex. 1023, Figure 1. Kinde 2011 also teaches that the unique identifier DNA is also known as a barcode. *Id.* at 9531. Thus, the combination of Kinde references teaches attaching a barcode to extracellular sequences or fragments prior to amplification as recited in claims 5 and 9.

Claim 6 calls for the use of non-unique barcodes. We agree with Petitioner that Kinde 2011 teaches this limitation. In the inverse PCR experiment reported in Kinde 2011, Kinde analyzed over 1000 different polynucleotides using just 12 different sequences for the UIDs of barcodes. Ex. 1012, 9531, SI10, table 4. Given the number of barcodes used in the analysis was significantly less than the number of polynucleotides analyzed, the barcodes used were not unique. Ex. 1002 ¶ 145.

Patent Owner contends that one skilled in the art would not have been motivated to combine the teachings of Kinde 2011 with those of Kinde 2012 to produce the method of claim 6 because Kinde 2011 requires that the polynucleotides have sheared ends to create a UID and the method of Kinde 2012 does not produce sheared ends. PO Resp. (citing Ex. 2026 ¶ 36).

We are not persuaded by Patent Owner's argument with respect to claim 6. Kinde 2012 teaches that the amplification and tagging process of Kinde 2011 was used in the method disclosed in Kinde 2011. Ex. 1041, 4. As discussed above, Kinde 2011 discloses using a set of UIDs that are not unique.

We conclude that Petitioner has shown that the subject matter of claims 5, 6, and 9 would have been obvious over Kinde 2012 combined with Kinde 2011.

8. Claim 7

Claim 7 reads

> The method of claim 5, wherein each barcode comprises a fixed or semi-random oligonucleotide sequence that in combination with a diversity of molecules sequenced from a selected region enables identification of unique molecules.

Ex. 1001, col. 63, ll. 13–17.

Petitioner contends that the step recited in claim 7 is taught by the combination of Kinde 2011 and Kinde 2012. Pet. 47–50. Petitioner contends that the index sequences used in one embodiment of Kinde 2011 comprise fixed sequences as the term is used in the '743 patent. Pet. 48–49. Petitioner contends that Kinde 2011 teaches that the index sequences can be used to create unique identifier DNA. *Id.* Petitioner contends that the combination of Kinde 2012 and Kinde 2011 teaches the use of the index sequences in combination with a diversity of molecules sequenced from a selected region for the identification of unique molecules. Pet. 50.

Patent Owner contends that like claim 6, claim 7 calls for the use of a non-unique barcode and that Kinde 2011 does not teach the use of a non-unique barcode. PO Resp. 44–45. Patent Owner also contends that Petitioner has not explained why one skilled in the art would use the specific embodiment in Kinde 2011 of using index sequences in the method disclosed in Kinde 2012. *Id.*

We have considered the arguments presented by the parties and the evidence of record and conclude that Petitioner has shown that the subject matter of claim 7 would have been obvious over Kinde 2012 combined with Kinde 2011. We agree with Petitioner that Kinde 2011 teaches the use of a

TEMPUS EX1002.160

non-unique barcode. Further, Kinde 2012 teaches using the UIDs of Kinde 2011 in the method of Kinde 2012. Ex. 1041, 4.

9. Claim 8

Claim 8 reads

> The method of claim 1, further comprising selectively enriching regions from a genome or transcriptome of the subject prior to sequencing.

Ex. 1001, col. 63, ll. 18–20.

Petitioner contends that the step of claim 8 is taught by Kinde 2012 in that Kinde 2012 teaches the use of specific primers that anneal to a subset of repeated regions dispersed throughout the genome prior to sequencing. Pet. 51 (citing Ex. 1041, 5). Petitioner contends that this step is the same as selective enrichment required by claim 8. *Id.*; Ex. 1002 ¶¶ 156–158.

Patent Owner does not present a separate argument regarding the patentability of claim 8. *See* PO Resp. 41–46.

We find that Petitioner has shown by a preponderance of the evidence that the subject matter of claim 8 would have been obvious to one skilled in the art over Kinde 2012 combined with Kinde 2011. Kinde 2012 teaches "the use of specific primers that anneal to a subset of repeated regions dispersed throughout the genome" prior to sequencing. EX1041, 5. Dr. Gabriel testifies, and we agree, that one skilled in the art would have understood selective enrichment to mean "selecting and/or enriching particular portions of genomic regions of interest for analysis." Ex. 1002 ¶ 156 (citing Ex. 1031, 374–375). Dr. Gabriel also testifies, and we agree. that using that definition of selective enhancement, the subject step of selective enhancement would have been obvious to one skilled in the art based on the teachings of Kinde 2012. Ex. 1002 ¶ 158.

TEMPUS EX1002.161

10. Claim 20

Claim 20 reads

> The method of claim 1, wherein each of the plurality of predefined regions is a single base.

Ex. 1001, col. 64, ll. 29–30.

Petitioner contends that Kinde 2011 teaches this limitation. Pet. 51–52. Petitioner contends that Kinde 2011 teaches identifying mutations in the CTNNB1 gene wherein each mutation is identified as a single mutation. *Id.* (citing Ex. 1023, 9533, Table S1). Petitioner also contends that one skilled in the art would have understood that next generation sequencing methods such as those taught in Kinde 2011 and Kinde 2012 "offer substantial benefits for copy number analysis, including higher resolution (up to the level of the single-base insertion or deletion)." *Id.* (quoting Ex. 1042, 690–691). Petitioner contends that based on the teaching of Kinde 2011 and the general knowledge in the art, it would have been obvious to one skilled in the art to analyze a plurality of regions each consisting of a single base pair. *Id.* at 52 (citing Ex. 1002 ¶ 161).

Patent Owner does not present a separate argument regarding the patentability of claim 20. *See* PO Resp. 41–46

We find that Petitioner has shown by a preponderance of the evidence that the subject matter of claim 20 would have been obvious to one of ordinary skill in the art over Kinde 2012 combined with Kinde 2011. Kinde 2011 teaches that the disclosed method can be used to detect single point mutations. Ex. 1023, 9533; Ex. 1002 ¶ 159. Dr. Gabriel testifies, and we agree, that given the state of the art and the teachings of Kinde 2011, it would have been obvious to analyze a plurality of regions each consisting of

TEMPUS EX1002.162

a single base pair. Ex. 1002 ¶ 161. For example, Meyerson[6] teaches "[n]ucleotide substitution mutations are the most common known somatic genomic alteration in cancer" and that next-generation sequencing methods "offer substantial benefits for copy number analysis, including higher resolution (up to the level of the single base insertion or deletion)." Ex. 1042, 690–691. Thus, we agree with Dr. Gabriel that it would have been routine for person of ordinary skill in the art to analyze a plurality of regions.

11. Claim 22

Claim 22 reads

The method of claim 1, wherein e) comprises i) and ii).

Ex. 1001, col. 64, ll. 33-34.

Petitioner contends that Kinde 2012 combined with Kinde 2011 teach the method of claim 22 as the combination of references teach performing both the normalization step and the processing step recited in step e). Pet. 52.

Patent Owner does not present a separate argument regarding the patentability of claim 22. *See* PO Resp. 41–46.

As discussed above, we find that Kinde 2012 combined with Kinde 2011 teach the steps of normalization followed by processing. Section II.E. 3.f. We conclude that Petitioner has shown by a preponderance of evidence that the subject matter of claim 22 would have been obvious to one skilled in the art over Kinde 2012 combined with Kinde 2011.

12. Claims 23–26

Claims 23 to 26 read

---

[6] Meyerson et al., *Advanced in understanding cancer genome through second-generation sequencing,* 11 NATURE REVIEWS 685 (2010) (Ex. 1042, "Meyerson").

TEMPUS EX1002.163

> 23. The method of claim 1, wherein e) comprises i) normalizing the number of reads in the plurality of predefined regions to each other.
>
> 24. The method of claim 1, wherein e) comprises i) normalizing the number of unique sequence reads in the plurality of predefined regions to each other.
>
> 25. The method of claim 1, wherein e) comprises ii) processing the number of reads in the plurality of predefined regions with numbers obtained from the control sample.
>
> 26. The method of claim 1, wherein e) comprises ii) processing the number of unique sequence reads in the plurality of predefined regions with numbers obtained from the control sample.

Ex. 1001, col. 64, ll. 36–48.

Petitioner contends that it would have been obvious to apply the normalization and processing steps of step e) to both reads and unique sequence reads in view of the teachings of Kinde 2012 and Kinde 2011. Pet. 52.

Patent Owner does not present a separate argument regarding the patentability of claims 23–26. *See* PO Resp. 41–46.

While we agree with Petitioner that Kinde 2011 and Kinde 2012 teach the application of both normalization and processing to "reads" as recited in claims 23 and 25, we do not agree that they teach the application of those steps to "unique sequence reads" as recited in claims 24 and 26. As discussed above, we have construed the term "unique sequence reads" to exclude the use of barcodes. Both Kinde 2012 and Kinde 2011 teach the use of barcodes as a part of their process. Ex. 1023, 9531; Ex. 1041, 2. Petitioner has not pointed to, nor have we discerned, any teaching in the record to show that one skilled in the art would have applied the methods of Kinde 2012 or Kinde 2011 to sequences lacking a barcode.

We conclude that Petitioner has shown by a preponderance of the evidence that the subject matter of claims 23 and 25 would have been obvious over Kinde 2012 combined with Kinde 2011, but that Petitioner has not shown that the subject matter of claims 24 and 26 would have been obvious.

13. Conclusion

Based on the foregoing we conclude that Petitioner has demonstrated by a preponderance of the evidence that claims 1–9, 20, 22, 23, and 25 are unpatentable as obvious over Kinde 2011 combined with Kinde 2012. We also conclude that Petitioner has not demonstrated by a preponderance of the evidence that claims 24 and 26 are unpatentable as obvious over Kinde 2011 combined with Kinde 2012.

F. Ground 3 – Obviousness based on Kinde 2011 and Forshew.

Petitioner contends that claims 10–13, 15–19, and 21 would have been obvious over the teachings of Kinde 2011 combined with Forshew.

1. Forshew

Forshew teaches a method for detecting rare "cancer mutations present in circulating [cell-free] DNA." Ex. 1035, Abstract. Forshew teaches targeting DNA fragments from genomic regions, amplifying and sequencing the fragments, and then aligning the sequences to a reference sequence to identify mutations. *Id.*, Fig. 1, Supplement 1–4. The method can be used to monitor mutation frequencies over time using samples from a patient and to compare mutation frequencies in a sample to a reference sample. *Id.* at 8–9, Fig. 4.

2. Analysis

Petitioner contends that Kinde 2011 combined with Forshew teaches the step in claim 10 calling for the detection of a rare mutation in a "cell-free

or substantially cell-free sample obtained from a subject." Pet. 56. In support of this contention, Petitioner points to the statement in Kinde 2011 that the method can be used "to determine the prevalence of rare mutations in a small region of the CTNNB1 gene," and that the method can be used to identify rare mutations. Pet. 56 (citing Ex. 1023, 9533). Petitioner also relies on Forshew, which describes a method of detecting rare mutations in cell-free tumor DNA. *Id.* (citing Ex. 1035, 1).

Petitioner contends that Kinde 2011 also teaches the filtering step of claim 10. Petitioner cites to Kinde 2011 as follows:

> Only high-quality reads meeting the following criteria were used for subsequent analysis: (i) the first 25 bases passed the standard Illumina chastity filter; (ii) every base in the read had a quality score ≥20; and (iii) ≤3 mismatches to expected sequences. For the exogenous UID libraries, we additionally required the UIDs to have a quality score ≥30.

Ex. 1023, SI2; Ex. 1002 ¶ 181.

Petitioner contends that Kinde 2011 teaches the step of mapping the sequence reads in claim 10. Pet. 57. In support of this contention, Petitioner cites to the teaching in Kinde 2011 where the sequences are aligned using the Eland pipeline, as well as the testimony of Dr. Gabriel. Pet. 57 (citing Ex. 1023, SI2; Ex. 1002 ¶ 183). Dr. Gabriel also testified that it would have been obvious to map the sequences against a human reference. Ex. 1002 ¶¶ 188–189.

Petitioner contends that Kinde 2011 teaches the step of determining unique sequence reads corresponding to the extracellular polynucleotides from among the sequence reads. Pet. 58. In support of this contention Petitioner cites to the portion of Kinde 2011 that teaches that the DNA molecules to be analyzed are assigned unique identifiers that are used to

group similar DNA fragments after amplification. Pet. 85; Ex. 1023, 9530, 9535. Petitioner contends that Kinde 2011 combined with Forshew teaches the step of claim 10 calling for "identifying a subset of mapped unique sequence reads that include a variant as compared to the reference sequence at each mappable base positon." Pet. 59.

Petitioner contends that Kinde 2011 teaches identifying a daughter fragment that includes a variant. Pet 59; Ex. 1023, 9530. Petitioner also contends that Forshew teaches a method that includes identifying subsets of reads that include a variant for each of multiple mappable base positions. Pet. 59. Petitioner contends that Forshew teaches the step for each mappable base position, calculating a ratio of (a) a number of mapped unique sequence reads that include a variant as compared to the reference sequence, to (b) a number of total unique sequence reads for each mappable base position. In support of this contention, Petitioner cites to Forshew where it teaches calculating an allele frequency for each specific genomic location. Pet. 60; Ex. 1035, 1, Fig 2; Ex. 1002 ¶ 197. Forshew teaches that an allele frequency is "calculated as the fraction of reads with the mutant." Ex. 1035, Fig 2. Dr. Gabriel testifies that this calculation would require knowing the number of reads with the mutation and the total number of reads. Ex. 1002 ¶ 202. Petitioner contends that one skilled in the art would have understood that by teaching allele frequency, Forshew was also teaching "calculating a ratio of mapped sequence reads that include a variant as compared to the reference sequence to a number of total sequence reads." Pet. 61.

Petitioner contends that Forshew teaches the step in claim 10 calling for "processing the ratio with a similarly derived number from a reference sample." Pet. 63. Petitioner contends that the processing step simply means "comparing a ratio from a test sample to a ratio from a reference sample."

Pet. 64; Ex. 1002 ¶ 208. Petitioner contends that the term "reference sample" is not defined in the Specification, but that one skilled in the art would understand the term to mean a sample that is "similar to the test sample and can be readily compared to the test sample." Pet. 64; Ex. 1002 ¶ 209. Petitioner contends that Forshew teaches comparing the ratios determined to the test samples with ratios of reference samples, particularly samples with normal DNA. Pet. 65; Ex. 1035, 8, Fig. 4A; Ex. 1002 ¶¶ 211–215.

Petitioner contends that one skilled in the art would have been motivated to combine the teachings of Kinde 2011 and Forshew because they both relate to using similar technology to detect rare mutations. Pet. 54. Petitioner also contends that one skilled in the art would have been motivated to combine the references, because Forshew teaches the desirability to increase the accuracy of the method and reduce false positives and Kinde 2011 teaches an improved method to achieve that goal. Pet. 54; Ex. 1002 ¶ 170. Petitioner contends that one skilled in the art would have had a reasonable expectation of success because "the method of Kinde 2011 was capable of detecting mutations at precisely the mutation frequencies that were likely to occur in cell-free DNA." Pet. 55.

In its Preliminary Response, Patent Owner contended that the references do not teach or suggest all elements of the claimed method. Prelim Resp. 42. Specifically, Patent Owner contended that Forshew fails to teach the analysis of unique sequence reads as required by claim 10. *Id.* at 44. Patent Owner contended that although Forshew refers to mapped sequence reads, nothing in Forshew teaches or suggests the use of unique sequence reads as required by claim 10. *Id.*

Patent Owner contended that Petitioner has not shown why one skilled in the art would have been motivated to combine Kinde 2011 and Forshew.

Id. at 44–49. Patent Owner contended that, other than a general allegation that the references are in the same field, Petitioner did not provide any specific reason why one skilled in the art would apply the teachings of Kinde 2001 to Forshew. *Id.* at 445. Patent Owner contended this is particularity important because Petitioner has not shown that Kinde 2011 is relevant to the cell-free DNA analysis described in Forshew. *Id.* Patent Owner also contended that Petitioner fails to describe with particularity how the references are to be combined. *Id.* at 48–49.

In our Decision to Institute, we concluded that Petitioner had failed to demonstrate a reasonable likelihood of showing that claims 10–13, 15–19 and 21 were unpatentable over Kinde 2011 combined with Forshew. Inst. Dec. 31. In the Decision to Institute, we found that Petitioner had failed to show that either Kinde 2011 or Forshew disclosed the use of unique sequence reads, i.e., a sequence read that contains unique sequences, such that the sequence reads can be identified without the need for a barcode or other added identifier. *Id.* In its Response, Patent Owner reiterates that the references fail to teach or suggest using a unique sequence read and that Petitioner failed to articulate a motivation to combine the references. PO Resp. 46–49.

As discussed above, our construction of the term "unique sequence reads" has not changed from the construction we gave the term in our Decision to Institute. As we found in the Decision to Institute, there is nothing in the teachings of either Kinde 2011 or Forshew that teach the use of unique sequence reads. Inst. Dec. 31.

In its Reply, Petitioner contends that the construction we gave to the term is in error and that Kinde 2011 specifically references unique identifiers to determine unique reads. Pet. Reply 26.

As discussed above, we do not find our original construction of the term "unique sequence reads" to be in error. Thus, the term excludes the use of barcodes or other added identifier.

While Kinde 2011 refers to unique identifiers, that term refers to the barcodes attached to the isolated sequences to aid in sorting the sequences. Ex. 1021, 9532, Figure 1. Thus, Kinde 2011 teaches the use of barcodes, which are excluded from the definition of unique sequence reads.

Claims 11–13, 15–19, and 21 depend, either directly or indirectly, on claim 10 and thus also require the use of "unique sequence reads," which Petitioner fails to establish is taught by the prior art. Thus, based on the foregoing we conclude that Petitioner has failed to establish by a preponderance of the evidence that claims 10–13, 15–19, and 21 are unpatentable over Kinde 2011 combined with Forshew.

G. Ground 4 – Obviousness based on Kinde 2011, Forshew, and Kinde 2012.

Petitioner contends that the subject matter of claim 14 would have been obvious over Kinde 2011 combined with Kinde 2012 and Forshew. Pet. 69–70.

Claim 14 depends indirectly from claim 10, and adds the step of determining a percent of sequences having copy number variation or rare mutation or variant in the bodily sample by calculating a percent of predefined regions with an amount of polynucleotides above or below a threshold. Ex. 1001, col. 64, ll. 6–11.

Petitioner contends that Kinde 2012 teaches this additional step. Pet. 70. Petitioner supports this contention by pointing to the teaching in Kinde 2012 that when the z-score for the samples were calculated, "no autosome in any sample has a z-score outside the range of -3.0 and 3.0." Pet. 70; Ex.

1041, Fig. 2. Petitioner contends that this passage teaches identifying the amount of nucleotides above and below a certain threshold.

Patent Owner contends that Petitioner has failed to show that claim 14 would have been obvious for the same reasons as discussed for claim 10 above. PO Resp. 51. Patent Owner also contends that Petitioner has failed to show how the cited portion of Kinde 2012 teaches the limitation of claim 14. *Id.*

In its Reply, Patent Owner contends claim 14 is not unpatentable for the reasons argued with respect to Ground 3 above. PO Resp. 51. Patent Owner also contends that none of the references teaches the limitation calling for "calculating a percentage of predefined regions with an amount of polynucleotides above or below a threshold." *Id.*

For the reasons stated above with respect to Ground 3, we conclude that Petitioner has failed to demonstrate by a preponderance of the evidence that claim 14 is unpatentable over Kinde 2011 combined with Forshew and Kinde 2012.

III. PATENT OWNER'S MOTION TO AMEND

Patent Owner's motion to amend is contingent on the Board's finding of unpatentability of the challenged claims 1–9, 20, and 22–26. MTA 1. Because we conclude that Petitioner has demonstrated that the challenged claims are unpatentable, we proceed to consider Patent Owner's motion to substitute claims 27–41 for claims 1–9, 20, and 22–26. For the reasons discussed below, Patent Owner's motion to amend is denied.

A. Threshold Requirements

In an *inter partes* review, claims may be added as part of a proposed motion to amend. 35 U.S.C. § 316(d). The Board must assess the

TEMPUS EX1002.171

patentability of the proposed substitute claims "without placing the burden of persuasion on the patent owner." *Aqua Prods., Inc. v. Matal*, 872 F.3d 1290, 1328 (Fed. Cir. 2017) (en banc); *see also Lectrosonics, Inc. v. Zaxcom, Inc.*, IPR2018-001129, Paper 15 at 3–4 (PTAB Feb. 25, 2019) (precedential). Subsequent to the issuance of *Aqua Products*, the Federal Circuit issued a decision in *Bosch Automotive Service Solutions, LLC v. Matal*, 878 F.3d 1027 (Fed. Cir. 2017) ("*Bosch*"), as well as a follow-up Order amending that decision on rehearing. *See Bosch Auto. Serv. Sols., LLC v. Iancu*, No. 2015-1928 (Fed. Cir. Mar. 15, 2018) (Order on Petition for Panel Rehearing).

In accordance with *Aqua Products*, *Bosch*, and *Lectrosonics*, a patent owner does not bear the burden of persuasion to demonstrate the patentability of the substitute claims presented in the motion to amend. Rather, ordinarily, "the petitioner bears the burden of proving that the proposed amended claims are unpatentable by a preponderance of the evidence." *Bosch*, 878 F.3d at 1040 (as amended on rehearing); *see Lectrosonics*, Paper 15 at 3–4. In determining whether a petitioner has proven unpatentability of the substitute claims, the Board focuses on "arguments and theories raised by the petitioner in its petition or opposition to the motion to amend." *Nike, Inc. v. Adidas AG*, 955 F.3d 45, 51 (Fed. Cir. 2020).

Patent Owner's proposed substitute claims, however, must still meet the statutory requirements of 35 U.S.C. § 316(d) and the procedural requirements of 37 C.F.R. § 42.121 as a threshold matter. *Lectrosonics,* Paper 15 at 4–8.. Accordingly, Patent Owner must demonstrate: (1) the amendment proposes a reasonable number of substitute claims; (2) the amendment does not seek to enlarge the scope of the claims of the patent or

introduce new subject matter; (3) the amendment responds to a ground of unpatentability involved in the trial; and (4) the original disclosure sets forth written description support for each proposed claim. *See* 35 U.S.C. § 316(d)(1)(B), (3); 37 C.F.R. § 42.121;

B. Proposed Substitute Claims

Proposed substitute claims 27–41 are reproduced below with markings showing proposed changes from claims 1–9, 20, and 22–26 respectively. Underlining indicates inserted text and double brackets indicates deletions.

> 27. (Substitute for claim 1, if found unpatentable) A method for detecting <u>sub-chromosomal</u> copy number variation, comprising:
>
> a) sequencing extracellular polynucleotides from a bodily sample from a subject, wherein each of the extracellular polynucleotides generates a plurality of sequence reads;
>
> b) filtering out reads that fail to meet a set accuracy, quality score, or mapping score threshold;
>
> c) mapping the plurality of sequence reads to a reference sequence;
>
> d) quantifying mapped reads or unique sequence reads in a plurality of
>
> predefined regions of the reference sequence; and
>
> e) determining sub-chromosomal copy number variation in one or more of the plurality of predefined regions by:
>
> i) normalizing a number of reads in the plurality of predefined regions to each other, or a number of unique sequence reads in the plurality of predefined regions to each other; and/or
>
> ii) processing a number of reads in the plurality of predefined regions or a number of unique sequence reads in the plurality of predefined regions with numbers obtained from a control sample,
>
> <u>wherein each of the predefined regions is up to 100 kb.</u>

28. (Substitute for claim 2, if found unpatentable) The method of claim [[1]] 27, further comprising isolating extracellular polynucleotides from the bodily sample.

29. (Substitute for claim 3, if found unpatentable) The method of claim [[1]] 27, further comprising generating copies of the extracellular polynucleotides prior to sequencing.

30. (Substitute for claim 4, if found unpatentable) The method of claim [[1]] 27, further comprising determining a percent of sequences having copy number variation or rare mutation or variant in the bodily sample.

31. (Substitute for claim 5, if found unpatentable) The method of claim [[1]] 27, further comprising attaching one or more barcodes to the extracellular polynucleotides or fragments thereof prior to sequencing.

32. (Substitute for claim 6, if found unpatentable) The method of claim [[5]] 31, wherein each barcode attached to the extracellular polynucleotides or fragments thereof prior to sequencing is not unique.

33. (Substitute for claim 7, if found unpatentable) The method of claim [[5]] 31, wherein each barcode comprises a fixed or semi-random oligonucleotide sequence that in combination with a diversity of molecules sequenced from a selected region enables identification of unique molecules.

34. (Substitute for claim 8, if found unpatentable) The method of claim [[1]] 27, further comprising selectively enriching regions from a genome or transcriptome of the subject prior to sequencing.

35. (Substitute for claim 9, if found unpatentable) The method of claim [[1]] 27, further comprising attaching one or more barcodes to the extracellular polynucleotides or fragments thereof prior to an amplification or enrichment step.

> 36. (Substitute for claim 20, if found unpatentable) The method of claim [[1]] 27, wherein each of the plurality of predefined regions is a single base.
>
> 37. (Substitute for claim 22, if found unpatentable) The method of claim [[1]] 27, wherein e) comprises i) and ii).
>
> 38. (Substitute for claim 23, if found unpatentable) The method of claim [[1]] 27, wherein e) comprises i) normalizing the number of reads in the plurality of predefined regions to each other.
>
> 39. (Substitute for claim 24, if found unpatentable) The method of claim [[1]] 27, wherein e) comprises i) normalizing the number of unique sequence reads in the plurality of predefined regions to each other.
>
> 40. (Substitute for claim 25, if found unpatentable) The method of claim [[1]] 27, wherein e) comprises ii) processing the number of reads in the plurality of predefined regions with numbers obtained from the control sample.
>
> 41. (Substitute for claim 26, if found unpatentable) The method of claim [[1]] 27, wherein e) comprises ii) processing the number of unique sequence reads in the plurality of predefined regions with numbers obtained from the control sample.

MTA, App'x A.

C. Statutory Requirements for a Motion to Amend

We construe only those terms that are in controversy, and only to the extent necessary to resolve the controversy. *See Vivid Techs.*, 200 F.3d at 803; *see also Nidec Motor Corp.*, 868 F.3dat, 1017 (applying *Vivid Techs.* in the context of an *inter partes* review). None of the newly added claim terms are in controversy, so no claim construction is required.

The number of substitute claims is also deemed to be reasonable. Patent Owner proposed 15 substitute claims to replace an equal number of original claims. *See* MTA, App'x. A.

Moreover, for the reasons discussed below, we determine that the proposed substitute claims do not broaden the invention and that substitute claims 27–41 are definite and have adequate written description support and that a reasonable number of claims have been proposed..

1. The Proposed Substitute Claims Do Not Broaden the Scope of the Claims

Proposed substitute claim 27 adds limitations limiting copy number variations to sub-chromosomal variations and limits the predefined regions to 100 kb or less. MTA 4. The amendments narrow the types of mutations detected by the claimed method and narrow the predefined regions analyzed by the method. The amendments do not broaden the scope of the claims.

2. The Proposed Substitute Claims Are Supported by the Specification

Patent Owner contends that the added limitation limiting the copy number variations to sub-chromosomal variations is supported by the Specification and by the prior applications. MTA 5–11. Patent Owner contends that one skilled in the art reading the present Specification and the prior applications would understand that the Specification teaches detecting copy number variation at the sub-chromosomal level. *Id.* Patent Owner points to the teaching in the Specification that detection is performed with "sub-chromosomal resolution: e.g., at least 100 megabase resolution, at least 10 megabase resolution, at least 1 megabase resolution, at least 100 kilobase resolution, at least 10 kilobase resolution or at least 1 kilobase resolution." MTA 6–7 (emphasis omitted); Ex. 1001, col. 7, ll. 38–45.

Patent Owner contends that the Specification also supports the limitation limiting the predefined regions to up to 100 kb. MTA 10–11. Patent Owner supports this contention by pointing to the teaching in the Specification where it states "In some embodiments, predefined regions are uniform or substantially uniform in size, about 10 kb, 20 kb, 30 kb 40 kb, 50 kb, 60 kb, 70 kb, 80 kb, 90 kb, or 100 kb in size." *Id.*; Ex. 1001, col. 3, ll. 61–63.

Petitioner contends that the term "sub-chromosomal copy number variation" is not supported by the Specification. MTA Opp. 2–4. Petitioner argues that while the term sub-chromosomal is used in the Specification, the term is used in connection with the resolution achieved by the claimed method and not the size of the copy number variation. *Id.* at 3. Petitioner also contends that the discussion of genetic aberrations in the Specification does not support the term sub-chromosomal copy number variations. *Id.* at 4.

Petitioner does not appear to contend that the limitation restricting the predefined region to up to 100 kb is not supported by the Specification. *See* MTA Opp. 4.

We have considered the arguments presented by the parties and the evidence of record and conclude that the proposed substitute claims are supported by the present disclosure. While we agree with Petitioner that the Specification does not specifically use the term "sub-chromosomal copy number variation" that does not mean that the term is unsupported by the Specification. *Lockwood v. Am. Airlines, Inc.*, 107 F.3d 1565, 1572 (Fed. Cir. 1997) (finding support "'need not be used *in haec verba*'"). As Dr. Bustamante testifies, and we agree, one skilled in the art, reading the present specification, would have understood that the Specification discloses a

method for detecting sub-chromosomal copy number variations. Ex. 2024 ¶¶ 35–44.

Specifically, the Specification refers to performing the claimed method at a sub-chromosomal resolution of "at least 100 megabase resolution, at least 10 megabase resolution, at least 1 megabase resolution, at least 100 kilobase resolution, at least 10 kilobase resolution or at least 1 kilobase resolution." Ex. 1001, col. 7, ll. 38–45. While this passage refers to sub-chromosomal resolution and not sub-chromosomal copy number variation, we are persuaded by Dr. Quackenbush's testimony that the level of resolution speaks to the size of the aberration detected. Ex. 1051, 104–107. We agree with Dr. Bustamante that one skilled in the art would have understood that this passage teaches detecting copy number variation at the sub-chromosomal level. Ex. 2024 ¶¶ 34–45.

3. The Proposed Substitute Claims Address a Ground of Unpatentability Raised in the Petition

The proposed substitute claims address two issues of patentability raised in the Petition, namely anticipation by Chiu and obviousness based on Kinde 2012 combined with Kinde 2011. In those grounds, we found that the references disclosed detecting copy number variation in that the references disclosed detection of aneuploidy, which is a type of copy number variation involving entire chromosomes. *See* Section II.E.3, *supra.* We also found that predefined regions included whole chromosomes and that the references disclosed or taught predefined regions that include whole chromosomes.

The amendments introduced in the proposed substitute claims address both of these issues by limiting copy number variations to sub-chromosomal copy number variations and by limiting predefined regions to a maximum of 100 kb, which is smaller than a whole chromosome.

D. Unpatentability of the Proposed Substitute Claims

Petitioner asserts that the proposed substitute claims are unpatentable under 35 U.S.C. §§ 102 or 103. MTA Opp. 1. Petitioner contends that claims 27 and 37–41 are anticipated by Rava.[7] *Id.* at 7–15. Petitioner also contends that the subject matter of claims 28–36 would have been obvious over Rava. *Id.* at 15–21. Petitioner also contends that the subject matter of claims 27 and 38–41 would have been obvious over Hyland.[8] *Id.* at 21–25. Petitioner supports it contentions with the declaration of Dr. Gabriel. Ex. 1060 Patent Owner disagrees. MTA Reply 1–12. Patent Owner relies on the declaration of Dr. Bustamante. Ex. 2037.

E. Anticipation by Rava

Petitioner contends that proposed substitute claims 27 and 37–41 are anticipated by Rava. MTA Opp. 7–15.

1. Rava

U.S. Patent No. 9,323,888 ("Rava"), entitled "Detecting and Classifying Copy Number Variation," was filed as U.S. Patent Application No. 13/600,043 on August 30, 2012. Ex. 1055, codes (10), (54), (21), (22).). Rava describes methods "for determining copy number variations (CNV) of a sequence of interest in a test sample that comprises a mixture of nucleic acids that are known or are suspected to differ in the amount of one or more sequence [sic] of interest." *Id.* at col. 2, ll. 31–35; Ex. 1060 ¶¶ 19–23. "Sequences of interest include genomic segment sequences ranging from, e.g., kilobases (kb) to megabases (Mb) to entire chromosomes that are known or are suspected to be associated with a genetic or a disease

[7] Rava et al., US 9,323,888 B2, issued April 26, 2016 (Ex. 1055) ("Rava").
[8] Hyland et al., US 2012/0046877 A1, published February 23, 2012 (Ex. 1061) ("Hyland").

condition." Ex. 1055, col. 27, l.65–col. 28, l.1; *see also id.* at col. 2, ll. 40–45; Ex. 1060 ¶ 21.

Rava applies its methods to analyze nucleic acid molecules in test samples, such as "maternal sample[s]," which include "a mixture of fetal and maternal cell-free DNA molecules." Ex. 1055, col. 9, ll. 8–10; Ex. 1060 ¶ 22. Rava's methods "may employ next generation sequencing technology (NGS)" to obtain quantitative information, e.g., "each sequence read is a countable 'sequence tag' representing an individual clonal DNA template or a single DNA molecule." Ex. 1055, col. 38, ll, 62–64, col. 39, ll. 2–16; Ex. 1002 ¶ 23. Rava uses generated sequencing data, *inter alia*, to:

> align the reads to portions of a reference sequence/genome and determine the amount of DNA (e.g., the number of reads) that map to defined portions [of] the reference sequence (e.g., to defined chromosomes or chromosome segments); calculate a dose of one or more of the defined portions by normalizing the amount of DNA mapping to the defined portions with an amount of DNA mapping to one or more normalizing chromosomes or chromosome segments selected for the defined portion; determining whether the dose indicates that the defined portion is "affected[.]"

Ex. 1055, col. 39, ll. 23–33; Ex. 1002 ¶ 23.

2. Analysis of Proposed Substitute Claim 27

Proposed substitute claim 27 is similar to claim 1 discussed above with the added limitations that the copy number variation is sub-chromosomal and that the predefined region contains up to 100 kb. MTA 4–5.

a) A method for detecting sub-chromosomal copy number variation comprising

Petitioner contends that Rava discloses this limitation. MTA Opp. 7. Specifically, Petitioner contends that Example 20 of Rava discloses a

method for noninvasively detecting fetal sub-chromosomal abnormalities including sub-chromosomal copy number variations. *Id.*

Patent Owner contends that Rava is directed to the detection of aneuploidy and that aneuploidy is outside the scope of the proposed substitute claims. MTA Reply 5.

We have considered the arguments presented by the parties and the evidence of record and conclude that Rava discloses a method for detecting sub-chromosomal copy number variation. Example 20 of Rava specifically discloses a method for detecting "mosaic sub-chromosomal copy number change . . . and [] sub-chromosomal duplication." Ex. 1055, col. 218, ll. 50–61; Ex. 1060 ¶ 32.

Patent Owner argues that Petitioner relies on examples in Rava that are directed to aneuploidy, which is not sub-chromosomal copy number variation. MTA Reply 5. While Patent Owner may be correct with respect to some examples, Petitioner specifically references Example 20 of Rava with respect to the limitation calling for detection of sub-chromosomal copy number variation. MTA Opp. 7. , Example 20 of Rava clearly relates to sub-chromosomal copy number variations. Ex. 1055, col. 218, ll. 50–61; Ex. 1060 ¶ 32.

Patent Owner also argues that a related Rava reference[9] ("Rava 2") was cited in prosecution as an anticipatory reference and that the rejection based on Rava 2 was overcome based on the lack of the filtering step. MTA Reply 6. We are not persuaded by this argument.

[9] Rava et al, US 2013/0029852 A1, published January 31, 2013 (Ex. 3002) ("Rava 2"). While both parties reference Rava 2 in their arguments, neither party made the reference part of the record.

We begin by noting that that 35 U.S.C. § 325(d) does not apply to motions to amend: "*In determining whether to institute or order a proceeding* under this chapter, chapter 30, or chapter 31, the Director may take into account whether, and reject the petition or request because, the same or substantially the same prior art or arguments previously were presented to the Office." *Id.* (emphasis added). In particular, § 325(d) does not apply to motions to amend because the Office has not previously passed on a proposed substitute claim before it was proposed.

Moreover, even if we were to consider whether the same or substantially the same prior art or arguments were previously presented to the OIffice, we find that there are significant differences between the two Rava references. The Rava reference cited by Petitioner is a continuation-in-part of U.S. Appl. No. 13/555,037. Ex. 1001, code (63). Rava 2 is the publication that corresponds to that application. Ex. 3002 (21). As Petitioner points out, Rava 2 does not contain Example 20, which serves as the basis for most, if not all of Petitioner's arguments opposing Patent Owner's motion to amend. MTA Sur-Reply 6; *compare* Ex. 1055, col. 218, l. 43 – col. 226, l. 50 *with* Ex. 3002 ¶¶ 916–922 (Example 19 is the last example in Ex. 3002). As discussed above, Example 20 relates specifically to detecting sub-chromosomal aberrations including sub-chromosomal copy number variations. Thus, for purpose of this decision, we find that Rava 2 is not substantially the same as Rava, and irrelevant to Petitioner's arguments in its opposition to Patent Owner's motion to amend.

b) Sequencing extracellular polynucleotides from a bodily sample from a subject, wherein each of the extracellular polynucleotides generates a plurality of sequence reads

Petitioner contends that Rava discloses this limitation. MTA Opp. 7–8. Petitioner contends that Rava discloses "sequencing cell free DNA from the maternal test sample to provide the sequence reads." *Id.* (quoting Ex. 1055, col. 12, ll. 5–34). Petitioner also contends that Example 20 of Rava discloses obtaining polynucleotides from maternal plasma samples, which are then sequenced to produce sequence reads. *Id.* (citing Ex, 1055, col. 219, ll. 32–40).

Patent Owner does not present a specific argument with respect to this limitation. *See* MTA Reply 5–9.

We find that Rava discloses the recited limitation. Rava discloses using sequences that originated from maternal blood samples. Ex. 1055, col. 219, ll. 10–31. The samples were then sequenced to produce a set of sequence reads. *Id.* at col. 219, ll. 31–44.

c) Filtering out reads that fail to meet a set accuracy, quality score, or mapping score threshold

Petitioner contends that Rava discloses this limitation. MTA Opp. 8–9. Petitioner contends that, in Example 20, Rava only uses reads that can be unambiguously mapped to a single genomic location and that up to two mismatches during alignment were allowed to be used in the analysis. *Id.* (citing Ex. 1055, col. 220, ll. 31–38).

Patent Owner contends that Rava does not disclose a filtering step. MTA Reply 6–7. Rather Rava discloses a threshold of unique tags in a sample, which is not the same as a mapping score or mapping threshold. *Id.* Patent Owner also contends that Rava allows a certain amount of mismatches in its analysis. *Id.*

We have considered the arguments presented by the parties and the evidence of record and conclude that Rava discloses the recited filtering step. Example 20 of Rava discloses

> Sequence reads were aligned to the human genome assembly hg19 obtained from the UCSC database (hgdownload.cse.csc.edu/goldenPath/hgl 9/bigZips/). Aligmnents were carried out utilizing the Bowtie short read aligner (version 0.12.5), allowing for up to two base mismatches during alignment. *Only reads that unambiguously mapped to a single genomic location were included.*

Ex. 1055, col. 220, ll. 31–38 (emphasis added). We agree with Dr. Gabriel that this discloses the recited filtering step of substitute claim 27. Ex. 1060 ¶¶ 40–42.

d) Mapping the plurality of sequence reads to a reference sequence

Petitioner contends that Rava discloses this limitation. MTA Opp. 9–10. Petitioner contends that Rava discloses "aligning the sequence reads to one or more chromosome reference sequences using a computing apparatus and thereby providing sequence tags corresponding to the sequence reads." *Id.* (quoting Ex. 1055, col. 10, ll. 49–52). Petitioner also contends that Example 20 of Rava discloses "[s]equence reads were aligned to the human genome assembly hg19 obtained from the UCSC database." *Id.* at 10 (quoting Ex. 1055, col. 220, ll. 31–43).

Patent Owner does not provide a specific argument regarding this limitation. *See* MTA Reply 5–9.

We find that Rava discloses the mapping step recited in substitute claim 27. Rava discloses aligning the sequences to the human genome assembly hg19. Ex. 1055, col. 220, ll. 31–42. Dr. Gabriel testifies, and we

agree, that this disclosure in Rava meets the requirement that the sequences are mapped against a reference sequence. Ex. 1060 ¶¶ 44–45.

e) Quantifying mapped reads or unique sequence reads in a plurality of predefined regions of the reference sequence

Petitioner contends that Rava discloses this limitation. Petitioner contends that Rava discloses "determin[ing] the amount of DNA (e.g., the number of reads) that map to defined portions the reference sequence (e.g., to defined chromosomes or chromosome segments)." MTA Opp. 10 (quoting Ex. 1055, col. 39, ll. 17–36). Petitioner also contends that this step is exemplified in Example 20 of Rava where it states "reads that unambiguously mapped to a single genomic location were included," and "[g]enomic sites at which [those] reads mapped were counted as tags." *Id.* at 11 (quoting Ex. 1055, col. 220, ll. 36–38).

Patent Owner contends that Rava does not disclose the quantification step as the portions of Rava relied upon by Petitioner are directed to detection of aneuploidy and not sub-chromosomal copy number variation. MTA Reply 5. Patent Owner also contends that Rava does not disclose predefined regions as the term is used in substitute claim 27. *Id.* at 7. Patent Owner contends that the term predefined means that the regions are defined in advance and that Rava does not define the regions until after the mapping and quantification steps are performed. *Id.*

We have considered the arguments presented by the parties and the evidence of record and find that Rava discloses the quantification step of substitute claim 27.

Rava discloses the step of "determin[ing] the amount of DNA (e.g., the number of reads) that map to defined portions the reference sequence (e.g., to defined chromosomes or chromosome segments)." Ex. 1055, col.

39, ll. 17–36. In Example 20, Rava discloses that "reads that unambiguously mapped to a single genomic location were included," and "[g]enomic sites at which [those] reads mapped were counted as tags." *Id.*, col. 220, ll. 36–38. Thus, we agree with Dr. Gabriel that Rava discloses the quantification step of substitute claim 27. Ex. 1060 ¶ 49.

We are not persuaded by Patent Owner's argument regarding Rava being limited to detecting aneuploidy. As discussed above, Example 20 of Rava specifically discloses detecting sub-chromosomal copy number variation. Ex. 1055, col. 218, ll. 50–61; Ex. 1060 ¶ 32.

We are also not persuaded by Patent Owner's argument that Rava does not disclose predefined regions. While Rava discloses that the sequences can be sorted into bins after the tags are counted, the bins themselves are defined before the tags are sorted. For example, in Example 20, the bins are defined by size. Ex. 1055, col. 220, ll. 44–47. Moreover, Rava discloses that the bin boundaries can be defined prior to alignment. *Id.* at col. 143, ll. 10–11; *see also id.* at col. 142, ll. 4–17; col. 143, ll. 10–27 and col. 145, ll. 46–59.

f) Determining sub-chromosomal copy number variation in one or more of the plurality of predefined regions by i) normalizing a number of reads in the plurality of predefined regions to each other, or a number of unique sequence reads in the plurality of predefined regions to each other; and/or ii) processing a number of reads in the plurality of predefined regions or a number of unique sequence reads in the plurality of predefined regions with numbers obtained from a control sample,

Petitioner contends that Rava discloses these limitations. MTA Opp. 11. Petitioner contends that Rava discloses "[d]etection of a sub-chromosomal abnormality was a multi-step process for classifying specific regions as having a copy number variant." *Id.* (quoting Ex. 1055, col. 221, ll.

41–54). Petitioner contends that Rava discloses both the normalization step and the processing step of substitute claim 27. *Id.* at 11–14.

Patent Owner presents the same arguments regarding this claim element as it did with respect to the quantification step discussed above. *See* MTA Reply 5–7.

We have considered the arguments presented by the parties and the evidence of record and find that Rava discloses determining sub-chromosomal copy number variation using normalization and processing steps. In Example 20, Rava discloses "[d]etection of a sub-chromosomal abnormality was a multi-step process for classifying specific regions as having a copy number variant." Ex. 1055, col. 221, ll. 41–43. Rava teaches that the data for each bin is first normalized. *Id.* at col. 220, ll. 44–57. Rava then discloses that the results of the normalization step were then examined for deviations from the median values measures across multiple samples to detect sub-chromosomal differences including copy number variations. *Id.* at col. 220, l. 63 – col. 221, l. 12.

As discussed above, we are not persuaded by Patent Owner's arguments that Rava is limited to detecting aneuploidy or that Rava does not disclose predefined regions.

g) Wherein each of the predefined regions is up to 100kb.

Petitioner contends that Rava discloses this limitation in that Rava teaches that the bins can be up to 100 kb in size. MTA Opp. 14.

Patent Owner does not present a separate argument with respect to this limitation. *See* MTA Reply 5–9.

We find that Rava discloses this limitation. In Example 20, Rava discloses dividing the sequences into 1Mb and 100 kb bins. Ex. 1055, col. 220, ll. 44–45. Dr. Gabriel has testified that the bins of Rava are the

equivalent of the predefined regions recited in proposed substitute claim 27 and that Rava discloses bins of up to 100 kb. Ex. 1060 ¶¶ 52, 61–62.

h) Conclusion

Based on the foregoing we conclude that Petitioner has demonstrated by a preponderance of the evidence that proposed substitute claim 27 is unpatentable as anticipated by Rava.

3. Proposed substitute claims 37–41

Proposed substitute claims 37–41 depend from proposed substitute claim 27. Proposed substitute claim 37 recites that the method comprises both steps e)i) and e)ii). Proposed substitute claims 38 and 39 recite that step e) comprises i) normalizing the number of reads or unique sequence reads, respectively, in the plurality of predefined regions to each other; and proposed substitute claims 40 and 41 recite that step e) comprises ii) processing the number of reads or unique sequence reads, respectively, in the plurality of predefined regions with numbers obtained from the control sample. MTA 25.

Petitioner contends that proposed substitute claims 37–41 are anticipated by Rava because Rava discloses the elements recited in the claims. MTA Opp. 12. Specifically, Petitioner contends that Rava discloses using both a normalization and a processing step and that Rava discloses the use of unique reads which also satisfies the read limitation. *Id.*

Patent Owner contends that the proposed substitute claims are not anticipated for the reasons proposed substitute claim 27 is not anticipated. MTA Reply 8. Patent Owner also argues that Petitioner has not established that Rava discloses "unique sequence reads." *Id.* Patent Owner supports this argument by pointing out that Dr. Gabriel has been inconsistent in the use of the term when analyzing Rava. *Id.*

TEMPUS EX1002.188

We have considered the arguments presented by the parties and the evidence of records and find that proposed substitute claims 37–41 are anticipated by Rava.

Proposed substitute claim 37 calls for the use of both the normalization step and the processing step of claim 27. MTA 25. As discussed above, in Example 20, Rava discloses the use of both a normalization step and a processing step.

Proposed substitute claim 38 calls for performing the normalization step of proposed substitute claim 27 using reads. MTA 25. As discussed above, Rava discloses normalizing sequences, which "are unambiguously mapped to a single genomic location." Ex. 1055, col. 220, ll. 32–43. Dr. Gabriel testifies, and we agree, that the sequences constitute reads as the term is used in the substitute claims. *See* Ex. 1060 ¶¶ 46, 48–49.

Proposed substitute claim 39 calls for performing the normalization step of proposed substitute claim 27 using unique sequence reads. MTA 25. As discussed above, Rava discloses normalizing sequences, which "are unambiguously mapped to a single genomic location." Ex. 1055, col. 220, ll. 32–43. Dr. Gabriel testifies, and we agree, that sequences that can be mapped to only one location are unique sequence reads. Ex. 1060 ¶¶ 48–49. This conclusion is consistent with our construction of unique sequence reads. Section II.C.3 *supra*.

Proposed substitute claim 40 calls for performing the processing step of proposed substitute claim 27 using reads. MTA 25. As discussed above, Rava discloses processing sequences, which "are unambiguously mapped to a single genomic location." Ex. 1055, col. 220, l. 63–col. 221, l. 12. Dr. Gabriel testifies, and we agree, that the sequences constitute reads as the term is used in the substitute claims. *See* Ex. 1060 ¶¶ 46, 48–49.

TEMPUS EX1002.189

Proposed substitute claim 41 calls for performing the processing step of proposed substitute claim 27 using unique sequence reads. MTA 25. As discussed above, Rava discloses normalizing sequences, which are "unambiguously mapped to a single genomic location." Ex. 1055, col. 220, l. 63–col. 221, l. 12. Dr. Gabriel testifies, and we agree, that sequences that can be mapped to only one location are unique sequence reads. Ex. 1060 ¶¶ 48–49. This conclusion is consistent with our construction of unique sequence reads. Section II.C.3 *supra*.

We are unpersuaded by Patent Owner's arguments. As discussed above, we find that proposed substitute claim 27 is anticipated by Rava. Section III.E.2 *supra.*

With respect to the term unique sequence reads, while Dr. Gabriel may have used two different definitions for the term, as discussed above, Rava's disclosure of using sequences that are "unambiguously mapped to a single genomic location" meets the definition of the term as we have construed it. Moreover, in the method disclosed in Example 20, Rava does not resort to the use of bar codes or other tags to sort or identify the sequences. *See* Ex. 1055 col. 220, ll. 32–43.

Based on the foregoing we conclude that Petitioner has demonstrated by a preponderance of the evidence that proposed substitute claims 37–41 are unpatentable as anticipated by Rava.

F. Obviousness based on Rava

Petitioner contends that the subject matter of proposed substitute claims 28–36 would have been obvious to one skilled in the art at the time the invention was made over Rava. MTA Opp. 15.

TEMPUS EX1002.190

1. Proposed Substitute Claim 28

Proposed substitute claim 28 depends from proposed substitute claim 27 and adds the step of "isolating extracellular polynucleotides from the bodily sample." MTA 23. Petitioner contends that Rava teaches this step. MTA Opp. 15.

Patent Owner argues that proposed substitute claim 28 is not rendered obvious by Rava, because Rava does not teach all the limitations of proposed substitute claim 27. MTA Reply 9.

We find that Rava teaches the step of isolating extracellular polynucleotides from a bodily sample. Rava teaches that "[t]o separate cell-free DNA from cells in a sample, various methods including, but not limited to fractionation, centrifugation (e.g., density gradient centrifugation), DNA-specific precipitation, or high-throughput cell sorting and/or other separation methods can be used." Ex. 1055, col. 54, ll. 2–7. Dr. Gabriel testifies, and we agree, that a person of ordinary skill in the art would have understood that, applying Rava's teachings "[t]o separate cell-free DNA from cells in a sample" to maternal plasma samples, which include cells and cell-free DNA, would isolate cell-free DNA from a bodily sample. Ex. 1060 ¶ 67.

Patent Owner's argument to the contrary is unpersuasive. As discussed above, we find that Rava discloses all of the elements of proposed substitute claim 27.

2. Proposed Substitute Claim 29

Proposed substitute claim 29 depends from proposed substitute claim 27 and adds the step of generating copies of the extracellular polynucleotides prior to sequencing. MTA 23. Petitioner contends that Rava teaches this step. MTA Opp. 16.

Patent Owner argues that proposed substitute claim 29 is not rendered obvious by Rava, as Rava does not teach all the limitations of proposed substitute claim 27. MTA Reply 9.

We find that Rava teaches the step of generating copies of the extracellular polynucleotides prior to sequencing. Rava teaches that "[i]n various embodiments the cfDNA present in the sample can be enriched specifically or non-specifically prior to use (e.g., prior to preparing a sequencing library)." Ex. 1055, 54:15–17. Rava explains that nonspecific enrichment of sample DNA can refer, e.g., "to the whole genome amplification of the genomic DNA fragments of the sample that can be used to increase the level of the sample DNA prior to preparing a cfDNA sequencing library." *Id.* at 54:17–42. Dr. Gabriel testifies, and we agree, that one skilled in the art would understand that Rava teaches generating copies of the extracellular polynucleotides prior to sequencing. Ex. 1060 ¶¶ 69–71.

Again we are unpersuaded by Patent Owner's argument to the contrary for the reasons set forth above.

3. Proposed Substitute Claim 30

Proposed substitute claim 30 depends from proposed substitute claim 27 and adds the step of "determining a percent of sequences having copy number variation or rare mutation or variant in the bodily sample." MTA 23. Petitioner contends that Rava teaches this step. MTA Opp. 16–17.

Patent Owner argues that proposed substitute claim 30 is not rendered obvious by Rava, as Rava does not teach all the limitations of proposed substitute claim 27. MTA Reply 9.

We find that Rava teaches the step of determining a percent of sequences having copy number variation or rare mutation or variant in the bodily sample. Rava teaches

> The fraction of the minor allele for any given informative marker is calculated by dividing the peak height of the minor component by the sum of the peak height for the major component, and the fraction is expressed as a percent for each informative locus as
>
> $$(\%\ \text{fetal fraction}) = \left[\frac{\text{peak height of minor allele(s)}}{\sum(\text{peak height of major allele(s)})}\right] \times 100 \qquad \text{Equation 3}$$

Ex. 1055, col. 134, ll. 40–50. Dr. Gabriel testifies, and we agree, that this method teaches one skilled in the art how to calculate the percentage of sequences having copy number variation or rare mutation or variant in the bodily sample. Ex. 1060 ¶¶ 72–74.

Again we are unpersuaded by Patent Owner's argument to the contrary for the reasons set forth above.

4. Proposed Substitute Claim 31

Proposed substitute claim 31 recites that the method of proposed substitute claim 27 further comprises "attaching one or more barcodes to the extracellular polynucleotides or fragments thereof prior to sequencing." MTA 23.

Petitioner contends that Rava teaches each of the limitations recited in substitute claim 31. MTA Opp. 17–18.

Patent Owner argues that proposed substitute claim 31 is not rendered obvious by Rava, as Rava does not teach all the limitations of proposed substitute claim 27. MTA Reply 9.

We find that Rava teaches the limitation of proposed substitute claim 31.

Rava teaches "Parallelizing sequencing for multiple samples i.e. multiplex sequencing, requires the incorporation of sample-specific index sequences, also known as barcodes, during the preparation of sequencing

libraries." Ex. 1055, col. 77, ll. 14–17. Rava teaches that one application of the index sequences is to verify the integrity of the multiple samples, which is "accomplished by first grouping sequence tags associated with identical index sequences to associate the genomic and marker sequences and distinguish sequences belonging to each of the libraries made from genomic molecules of a plurality of samples." *Id.* at col. 78, ll. 9–14. Dr. Gabriel testifies, and we agree, that because Rava discloses adding "sample-specific" barcodes prior to sequencing in order to later distinguish sequences from each sample, a person of ordinary skill in the art would have understood that extracellular polynucleotides in a sample would have a common (i.e., non-unique) index (i.e., barcode). Ex. 1060 ¶¶ 76-77.

In addition, Rava describes that the sample-specific barcodes can be incorporated into PCR primers used for amplification. Ex. 1055, col. 77, ll. 14–25. Dr. Gabriel testified that a person of ordinary skill in the art would have understood that a "sample specific" barcode sequence would be known or "fixed" prior to adding the barcode to polynucleotides. Ex. 1060 ¶ 78. Dr. Gabriel also testifies, and we agree, that because unique molecules in a multiplex sample can be identified by (1) mapping sequence reads to determine their start/end coordinates (*see, e.g.*, Ex. 1055 166:58–167:5), and (2) attributing different reads with the same start/end coordinates to particular samples using sample-specific barcodes (*id.*, 77:14–17), a person of ordinary skill in the art would have understood that Rava's fixed sample-specific barcodes in combination with the sequenced molecules enable identification of unique molecules. Ex. 1060 ¶¶ 78–79; *see also* Ex. 1001, col. 37, ll. 50–59.

Again we are unpersuaded by Patent Owner's argument for the reasons set forth above.

5. Proposed Substitute Claim 32

Proposed substitute claim 32 adds the limitation to proposed substitute claim 31 that the barcode attached to the polynucleotides or fragments thereof are not unique. MTA, 24.

Petitioner contends that proposed substitute claim 32 is unpatentable as obvious over Rava. MTA Opp. 17–18. Petitioner contends that Rava teaches that the index sequences, i.e. barcodes, are sample specific and that one skilled in the art would understand this to mean that the barcodes are not unique. *Id.* (citing Ex. 1055, col. 77, ll. 14–17; col. 78, ll. 9–14; Ex. 1060 ¶¶ 76–77.

Patent Owner argues that proposed substitute claim 32 is not rendered obvious by Rava, as Rava does not teach all the limitations of proposed substitute claim 27. MTA Reply 9.

We find that Rava teaches the limitation of proposed substitute claim 32.

Rava teaches

> Parallelizing sequencing for multiple samples i.e. multiplex sequencing, requires the incorporation of sample-specific index sequences, also known as barcodes, during the preparation of sequencing libraries. Sequencing indexes are distinct base sequences of about 5, about 10, about 15, about 20 about 25, or more bases that are added at the 3' end of the genomic and marker nucleic acid.

Ex. 1055, col. 77, ll. 14–20. Rava also teaches

> Verification of the integrity of 10 each of the plurality of source samples is accomplished by first grouping sequence tags associated with identical index sequences to associate the genomic and marker sequences and distinguish sequences belonging to each of the libraries made from genomic molecules of a plurality of samples.

Id. at col. 78, ll. 9–14. We agree with Dr. Gabriel that one skilled in the art would understand that Rava teaches that use of barcodes that are not unique. Ex. 1060 ¶¶ 76–77.

Again we are unpersuaded by Patent Owner's argument for the reasons set forth above.

6. Proposed Substitute Claim 33

Proposed substitute claim 33 adds the limitation that each barcode comprises a fixed on semi-random oligonucleotide sequence that in combination with a diversity of molecules sequenced from a selected region enables identification of unique molecules.

Petitioner contends that Rava teaches this limitation. MTA Opp. 18. Petitioner contends that Rava teaches that the barcodes sequences are known or fixed prior to adding them to the polynucleotide sequence and that the fixed sequences in combination with the sequences molecules enables identification of unique molecules. *Id.* (citing Ex. 1055, col. 77, ll. 14–25; col. 166, l. 58–col. 167, l. 5; Ex. 1060 ¶¶ 78–79).

Patent Owner argues that proposed substitute claim 33 is not rendered obvious by Rava, as Rava does not teach all the limitations of proposed substitute claim 27. MTA Reply 9.

We find that Rava teaches the limitation of proposed substitute claim 33.

As discussed above, Rava teaches the incorporation of sample specific barcodes into PCR primers used for amplification. Ex. 1055, col. 77, ll. 14–25. Rava also teaches "Sequencing indexes are distinct base sequences of about 5, about 10, about 15, about 20 about 25, or more bases." *Id.* at col. 77, ll. 17–19. Rava goes on to teach that the combination of the fixed or known barcodes in combination with the sequenced molecule allows for the

identification of unique molecules. *Id.* at col. 77, ll. 5–43; col. 166, l. 58–co. 167, l. 5. We agree with Dr. Gabriel that claim 33 would have been obvious over the teachings of Rava. Ex. 1060 ¶ 79.

Again we are unpersuaded by Patent Owner's argument for the reasons set forth above.

7. Proposed Substitute Claim 34

Proposed substitute claim 34 depends from proposed substitute claim 27 and adds the step of "selectively enriching regions from a genome or transcriptome of the subject prior to sequencing." MTA 24. Petitioner contends that Rava teaches this step. MTA Opp. 18–19.

Patent Owner argues that proposed substitute claim 34 is not rendered obvious by Rava, as Rava does not teach all the limitations of proposed substitute claim 27. MTA Reply 9.

We find that Rava teaches the step recited in proposed substitute claim 34. Rava discloses that "[i]n various embodiments the cfDNA present in the sample can be enriched specifically or non-specifically prior to use (e.g., prior to preparing a sequencing library)." Ex. 1055, col. 54, ll. 15–17. For example, Rava describes a method comprising "enriching the mixture of nucleic acids for polymorphic target nucleic acids in operation" and "sequencing the enriched mixture of nucleic acids in operation." *Id.* at col. 128, ll. 35–42; *see also, e.g., id.* at col. 131, ll. 7–12 ("[i]n other embodiments, the cfDNA in the sample is enriched specifically. Specific enrichment refers to the enrichment of a genomic sample for specific sequences, e.g. polymorphic target sequence … ."); col. 54, ll. 21–27; col. 128, ll. 57–67; col. 29, ll. 1–25; col. 129, ll. 32–37. We agree with Dr.

Gabriel that Rava teaches the selective enrichment step of substitute claim 34. Ex. 1060 ¶¶ 80–82.

Again we are unpersuaded by Patent Owner's argument for the reasons set forth above.

8. Proposed Substitute Claim 35

Proposed substitute claim 35 depends from proposed substitute claim 27 and adds the limitation of "attaching one or more barcodes to the extracellular polynucleotides or fragments thereof prior to an amplification or enrichment step." MTA 24. Petitioner contends that Rava teaches this limitation. MTA Opp. 19. Petitioner contends that Rava teaches that "the index sequence is incorporated into the adaptor which is ligated to the cfDNA prior to PCR amplification." *Id.* (quoting Ex. 1055, col. 77, ll. 25–28.)

Patent Owner argues that proposed substitute claim 35 is not rendered obvious by Rava, as Rava does not teach all the limitations of proposed substitute claim 27. MTA Reply 9.

We find that Rava teaches the step recited in proposed substitute claim 35. Rava teaches "the index sequence can be incorporated into the adaptor, which is ligated to the cfDNA prior to the PCR amplification." Ex. 1055, col. 77, ll. 25–28. As discussed above, the index sequence of Rava is the same as a barcode. Ex. 1055, col. 77, ll. 15–16. Dr. Gabriel also testifies, and we agree that this passage of Rava teaches attaching one or more barcodes to the extracellular polynucleotide prior to amplification. Ex. 1060 ¶ 84.

Again we are unpersuaded by Patent Owner's argument for the reasons set forth above.

9. Proposed Substitute Claim 36

Proposed substitute claim 36 depends from proposed substitute claim 27 and adds the limitation that "each of the plurality of predefined regions is a single base." MTA 24. Petitioner contends that Rava teaches this limitation. MTA Opp. 19–21.

Patent Owner argues that proposed substitute claim 36 is not rendered obvious by Rava, as Rava does not teach all the limitations of proposed substitute claim 27. MTA Reply 9.

We find that Rava teaches the limitation calling for the predefined region to be a single base. Rava teaches identifying single nucleotide polymorphisms (SNPs) in fetal and maternal DNA. *See, e.g.*, Ex. 1055, col. 22, ll. 54–59, Fig. 12. Rava also teaches that "the smaller the bin length, the greater the resolution that is obtained to localize the CNV of the segment in the chromosome of interest." *Id.* at col. 49, ll. 65–67. Dr. Gabriel testifies, and we agree, that a person of ordinary skill in the art "reading Rava would have thus understood that a smaller predefined region should be used in order to resolve or detect a small mutation, and that predefined regions of a single base should be used in order to resolve or detect a single base mutation, such as a SNP." Ex. 1060 ¶ 86.

Again, we are unpersuaded by Patent Owner's argument for the reasons set forth above.

10. Conclusion

Based on the foregoing we conclude that Petitioner has demonstrated by a preponderance of the evidence that proposed substitute claims 28–36 are unpatentable as obvious over Rava.

TEMPUS EX1002.199

G. Obviousness based on Hyland

Petitioner contends that the subject matter of proposed substitute claims 27 and 38–41 would have been obvious to a person of ordinary skill in the art at the time the invention was made over Hyland. MTA Opp. 21–25.

Patent Owner contends that Hyland does not render proposed substitute claim 27 obvious in that Hyland fails to teach certain steps of the substitute claim including the filtering step. MTA Reply 9–12.

As discussed more fully below, after considering the arguments presented by the parties and the evidence of record, we conclude that Hyland does not render proposed substitute claims 27 and 38–41 unpatentable for obviousness.

1. Hyland

Hyland describes, *inter alia*, methods "for copy number variation determination from analyzing biomolecule-related sequence reads." Ex. 1061 ¶ 7. Hyland states

> According to the present teachings, nucleic acid sequencing technologies can be utilized for genome-wide interrogation of CNV s. In contrast to conventional approaches (e.g., array-based methods, etc.), with sequencing, genomic coverage data is available at single base resolution which allows for high levels of fidelity when researchers and clinicians search for genomic variants such as CNVs in a genome.

Id. ¶ 27.

a) Analysis of Proposed Substitute Claim 27

Like claim 1 discussed above, proposed substitute claim 27 includes the limitation "filtering out reads that fail to meet a set accuracy, quality score, or mapping score threshold." MTA 22–23. Petitioner contends that

TEMPUS EX1002.200

Hyland teaches this step of substitute claim 27. MTA Opp. 22. Petitioner contends that Hyland teaches filtering out reads that do not meet predetermined length requirements or mappability thresholds. *Id.*

Patent Owner contends that Hyland does not teach filtering out reads but teaches thresholds for merging window regions, which refer to the reference sequence and not the sequence reads. MTA Reply 11–12.

We have considered the arguments presented by the parties and the evidence of record and conclude that Petitioner has not demonstrated that Hyland teaches the filtering step of proposed substitute claim 27. While Hyland teaches filtering windows, the filtering is done as part of the data analysis to determine if copy number variation exists and not as part of eliminating sequence reads. Hyland teaches the filtering step as one embodiment of step 514, which relates to using the copy number states of each window to identify copy number variation. Ex. 1061 ¶ 99. As shown in Figure 5 of Hyland, step 514 is the final step of the method coming after the normalization step.

TEMPUS EX1002.201

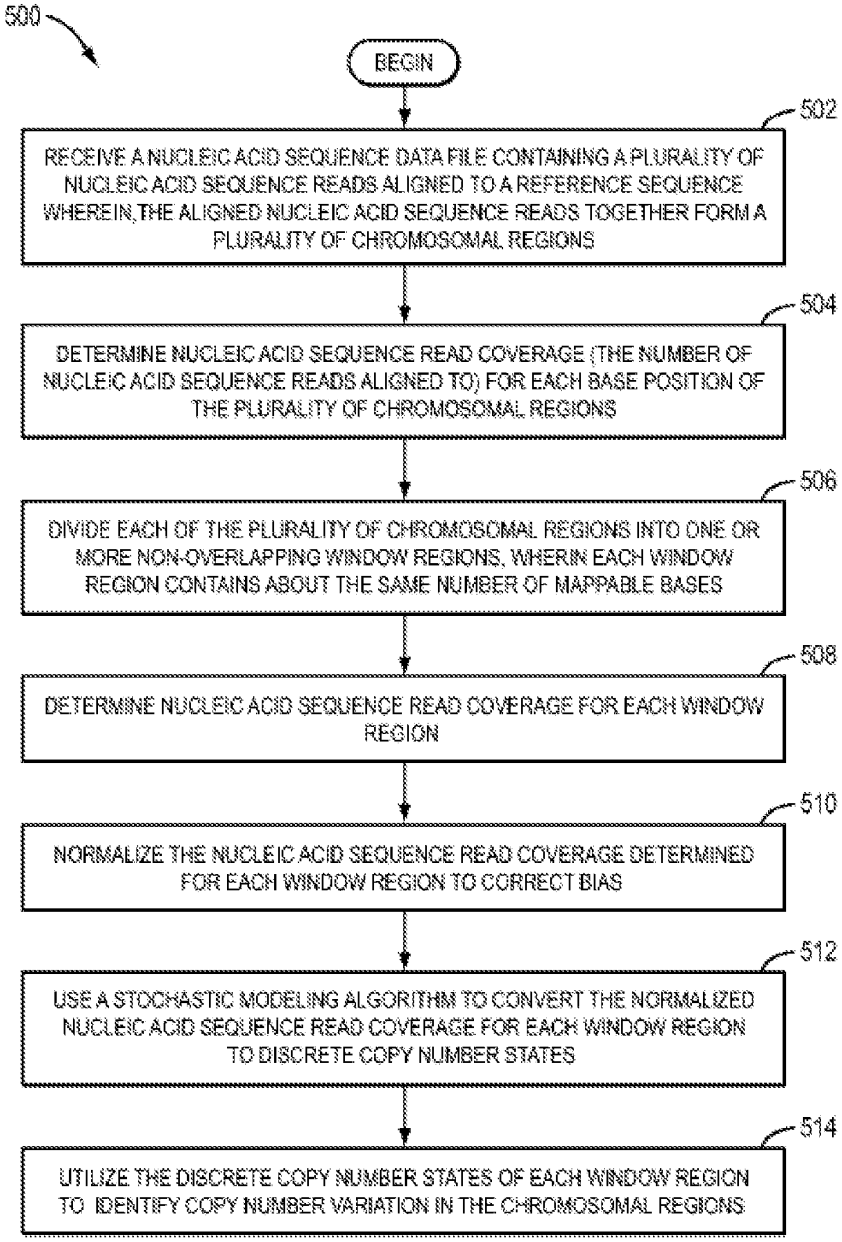


FIG. 5

Figure 5 of Hyland showing a flow chart of a method for determining copy number variation.

As Patent Owner points out, the filtering operation is done to determine if the windows can be merged. MTA Reply 11; Ex. 1061 ¶ 99. We discern nothing in the cited paragraph nor has Petitioner pointed to any teaching in Hyland that teaches filtering out reads.

Petitioner contends that the disclosure in Hyland is strikingly similar to the disclosure in the present Specification and that this supports the contention that Hyland teaches filtering out reads. MTA Sur-Reply 11. We are not persuaded. Both the passage in Hyland and the passage in the present Specification relate to merging window regions as part of the reporting process. *Compare* Ex. 1001, col. 46, ll. 40–47, *with* Ex. 1061 ¶ 99.[10]

b) Proposed Substitute Claims 38–41

Proposed substitute claims 38–41 depend from proposed substitute claim 27 and include all the limitations of proposed substitute claim 27. 35 U.S.C. § 112(d). As discussed above, Hyland does not render proposed substitute claim 27 obvious as Hyland does not teach or suggest the filtering step of proposed substitute claim 27. For the same reasons, Hyland does not render proposed substitute claims 38–41 obvious.

c) Conclusion

We conclude that Petitioner has failed to demonstrate by a preponderance of the evidence that proposed substitute claims 27 and 38–41 are unpatentable as obvious over Hyland.

[10] Petitioner presents a comparison of Ex. 1061 paragraph 99 with paragraph 64 of Ex. 2043. MTA Sur-Reply 11. Ex. 2034 is a copy of the Provisional Application, which led to the '743 Patent. *Id.* The paragraph is identical to the cited paragraph in the present Specification.

H. Patent Owner's Motion to Exclude

Patent Owner has moved to exclude exhibits 1002, 1004, 1007, 1008, 1013, 1015, 1017, 1019, 1025, 1032, 1043, 1048, 1052–1054, 1057–160, and 1063. Paper 42 ("PO Mot. Excl."). Petitioner filed an Opposition Paper 45 ("Pet Opp. PO Mot Excl.") and Patent Owner filed a reply. Paper 46 ("Reply PO Mot. Excl.").

With respect to Exhibits 1004, 1007, 1008, 1013, 1015, 1017, 1019, 1025, 1032, 1043, 1052–105, 1057–59, and 1063, we dismiss Patent Owner's Motion as moot. The panel did not rely on those exhibits in reaching its decision in this proceeding.

We also dismiss as moot the motion with respect to Exhibit 1058. Exhibit 1058 was cited only to define the acronym COSMIC used in the present Specification.

Patent Owner seeks to exclude Exhibits 1002 and 1060, which are the declarations of Dr. Gabriel. PO Mot. Excl. 1–4, 6–8. Patent Owner contends that they are irrelevant and misleading. *Id.* Petitioner contends that the declarations of Dr. Gabriel are highly relevant and are not misleading. Opp. PO Mot. Excl. 2–7, 9–11.

We have considered the arguments presented by the parties and we are not persuaded that Exhibits 1002 and 1060 should be excluded. Petitioner has shown and Patent Owner does not dispute, that Dr. Gabriel is an expert in the relevant art. Opp. PO Mot. Excl. 2–3. As an expert in the relevant art, Dr. Gabriel's testimony is highly relevant regarding how one skilled in the art would interpret the references as well as to the general knowledge in the field at the time the invention was made.

Patent Owner contends that Ex. 1002 should be excluded because Dr. Gabriel applied an inappropriate definition of a person of ordinary skill in

TEMPUS EX1002.204

the art. PO Mot. Excl. 1–3. We are not persuaded that this constitutes sufficient grounds to strike Exhibit 1002. As Petitioner points out, Dr. Gabriel applied the same definition of a person of ordinary skill in the art as the one we adopted in our Decision to Institute. Opp. PO Mot Excl. 4.

With respect to Exhibit 1060, Patent Owner contends that Dr. Gabriel failed to apply the plain and ordinary meaning to certain terms in her analysis of the art asserted against the substitute claims. PO Mot. Excl. 6–9. Even if we were to accept Patent Owner's contentions as true, we find that any such errors go to the weight given Dr. Gabriel's testimony and not its admissibility.

For the reasons set forth above, Patent Owner's Motion to Exclude is denied.

I. Petitioner's Motion to Exclude

Petitioner seeks to exclude Exhibits 2017, 2023, 2024, 2026, 2037, 2035, 2037, and 2038 in whole or in part. Paper 43 ("Pet. Mot Excl.") 1. Patent Owner filed an Opposition, Paper 46("Opp. Pet. Mot. Excl.") and Petitioner filed a Reply. Paper 47 ("Reply Pet. Mot. Excl.").

With respect to Exhibits 2017, 2035 and 2038, we deny Petitioner's motion as moot. These exhibits do not form any basis in reaching our decision in this proceeding.

The remaining exhibits are the declarations of Patent Owner's experts, Drs. Bustamante and Quackenbush. Petitioner contends that portions of these declarations should be excluded as the testimony "fails to rely on sufficient facts and is not the product of reliable principles and methods." Pet. Mot. Excl. 2.

As with the Declarations of Dr. Gabriel discussed above, we find that these alleged deficiencies go to the weight that should be afforded the

TEMPUS EX1002.205

experts' testimony and does not support a motion to exclude. The testimony of Drs. Bustamante and Quackenbush is highly relevant to the issues before us and the risk of any possible prejudice is very low. "Where, as here, the decision is by an administrative agency, rather than a jury, there is a diminished concern that [] exhibits would be prejudicial." *Eli Lilly and Co., v. Teva Pharms. Int'l GMBH*, IPR2018-01710, Paper 68 at 161 (PTAB Mar. 31, 2020) (citing *Corning Inc. v. DSM IP Assets B.V.*, IPR2013-00053, Paper 66 at 19 (PTAB May 1, 2014)).

For the reasons set forth above, Petitioner's Motion to Exclude is denied.

IV. CONCLUSION[11]

We conclude that Petitioner has satisfied its burden of demonstrating by a preponderance of the evidence that original claims 1–9, 20, and 22–26 are unpatentable.

We conclude that Petitioner has not satisfied its burden of demonstrating by a preponderance of the evidence that claims 10–19 and 21 are unpatentable.

We conclude that Petitioner has shown by a preponderance of the evidence that proposed substitute claims 27–41 are unpatentable.

---

[11] Should Patent Owner wish to pursue amendment of the challenged claims in a reissue or reexamination proceeding subsequent to the issuance of this decision, we draw Patent Owner's attention to the April 2019 *Notice Regarding Options for Amendments by Patent Owner Through Reissue or Reexamination During a Pending AIA Trial Proceeding. See* 84 Fed. Reg. 16,654 (Apr. 22, 2019). If Patent Owner chooses to file a reissue application or a request for reexamination of the challenged patent, we remind Patent Owner of its continuing obligation to notify the Board of any such related matters in updated mandatory notices. *See* 37 C.F.R. § 42.8(a)(3), (b)(2).

TEMPUS EX1002.206

V. ORDER

In consideration of the foregoing, it is hereby:

ORDERED that claims 1–9, 20, and 22–26 of U.S. Patent 9,840,743 B2 have been shown to be unpatentable;

FURTHER ORDERED that claims 10–19 and 21 have not been shown to be unpatentable;

FURTHER ORDERED that Patent Owner's Contingent Motion to Amend is denied;

FURTHER ORDERED that Patent Owner's Motion to Exclude is denied;

FURTHER ORDERED that Petitioner's Motion to Exclude is denied; and

FURTHER ORDERED that, because this is a Final Written Decision, parties to the proceeding seeking judicial review of the decision must comply with the notice and service requirements of 37 C.F.R. § 90.2.

In summary:

Claims	35 U.S.C. §	Reference(s)/Basis	Claims Shown Unpatentable	Claims Not shown Unpatentable
1–3, 22–26	102(a)	Chiu	1–3, 22–26	
1–9, 20, 22–26	103	Kinde 2011, Kinde 2012	1–9, 20, 22, 23, 25	24, 26
10-13, 15-19, 21	103	Kinde 2011, Forshew		10–13, 15–19, 21
14	103	Kinde 201, Kinde 2012, Forshew		14
Overall Outcome			1–9, 20, 22–26	10–19, 21

Motion to Amend Outcome	Claims(s)
Substitute Claims Proposed in the Amendment	26–41
Substitute Claims: Motion to Amend Granted	
Substitute Claims: Motion to Amend Denied	26–41
Substitute Claims: Not Reached	

For PETITIONER:

Rolando Medina
Eric Marandett
Stephanie Schonewald
CHOATE, HALL & STEWART LLP
rmedina@choate.com
emarandett@choate.com
sschoenwald@choate.com

For PATENT OWNER:

Michael Rosato
Steven Parmelee
Sonja Gerrard
WILSON SONSINI GOODRICH & ROSATI
mrosato@wsgr.com
sparmelee@wsgr.com
sgerrard@wsgr.com

Trials@uspto.gov
571-272-7822

Paper 47
Date: August 18, 2020

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

FOUNDATION MEDICINE, INC.,
Petitioner,

v.

GUARDANT HEALTH, INC.,
Patent Owner.

IPR2019-00652
Patent 9,834,822 B2

Before SUSAN L. C. MITCHELL, TINA E. HULSE, and KRISTI L. R. SAWERT, *Administrative Patent Judges.*

SAWERT, *Administrative Patent Judge*.

JUDGMENT
Final Written Decision
Determining Some Claims Unpatentable
Dismissing in Part and Denying in Part Petitioner's Motion to Exclude
Dismissing in Part and Denying in Part Patent Owner's Motion to Exclude
35 U.S.C. § 318(a)

I. INTRODUCTION

This is a Final Written Decision in an *inter partes* review challenging the patentability of claims 1–13 and 17–20 ("the challenged claims") of U.S. Patent No. 9,834,822 B2 (Ex. 1001, "the '822 patent"). We have jurisdiction under 35 U.S.C. § 6 and enter this Decision pursuant to 35 U.S.C. § 318(a) and 37 C.F.R. § 42.73. For the reasons set forth below, we determine that Petitioner has shown, by a preponderance of the evidence, that claims 1–11, 13, and 17–20 are unpatentable. We determine that Petitioner has not shown, by a preponderance of the evidence, that claim 12 is unpatentable. *See* 35 U.S.C. § 316(e) (2012).

A. Procedural History

Foundation Medicine, Inc. ("Petitioner") filed a Petition for an *inter partes* review under 35 U.S.C. § 311. Paper 2 ("Pet."). Petitioner supported its Petition with the Declaration of Stacey Gabriel, Ph.D. Ex. 1002. Guardant Health, Inc. ("Patent Owner") filed a Preliminary Response. Paper 6. On our authorization (Paper 9), Petitioner filed a Reply to Patent Owner's Preliminary Response (Paper 11).

On August 19, 2019, pursuant to 35 U.S.C. § 314(a), we instituted trial to determine whether any challenged claim of the '822 patent is unpatentable based on the grounds raised in the Petition:

TEMPUS EX1002.211

Claims Challenged	35 U.S.C. §	Reference(s)/Basis
1–13, 17–20	103(a)[1]	Schmitt,[2] Schmitt 2012,[3] and Fan[4] or Forshew[5]

Paper 12, 7, 36 ("Institution Decision" or "Inst. Dec.").

Patent Owner filed a Response. Paper 26 ("PO Resp."). Patent Owner supported its Response with the Declaration of Jay Shendure, M.D., Ph.D., Ex. 2023, and the Declaration of John Quackenbush, Ph.D., Ex. 2025. Petitioner filed a Reply to Patent Owner's Response. Paper 32 ("Pet. Reply"). Petitioner supported its Reply with a Reply Declaration of Dr. Gabriel. Ex. 1104. Patent Owner filed a Sur-Reply. Paper 34 ("PO Sur-Reply"). Patent Owner supported its Sur-Reply with a Supplemental Declaration of Dr. Quackenbush. Ex. 2042.

---

[1] The Leahy-Smith America Invents Act ("AIA"), Pub. L. No. 112-29, 125 Stat. 284, 287–88 (2011), amended 35 U.S.C. § 103, effective March 16, 2013. Because the challenged claims have an effective filing date before this date, the pre-AIA version of § 103 applies.

[2] Michael Schmitt et al., U.S. Patent No. 9,752,188 B2, issued Sept. 5, 2017 (Ex. 1011, "Schmitt").

[3] Michael W. Schmitt et al., *Detection of Ultra-rare Mutations by Next-generation Sequencing*, 109(36) PROC. NATL. ACAD. SCI. 14508–513 (2012) (Ex. 1047, "Schmitt 2012").

[4] Christina Fan et al., *Noninvasive diagnosis of fetal aneuploidy by shotgun sequencing DNA from maternal blood*, 105(42) PROC. NATL. ACAD. SCI. 16266–271 (2008) (Ex. 1048, "Fan")

[5] Tim Forshew et al., *Noninvasive Identification and Monitoring of Cancer Mutations by Targeted Deep Sequencing of Plasma DNA*, 4(136) SCI. TRANSL. MED. 1–34 (2012) (Ex. 1004, "Forshew").

TEMPUS EX1002.212

Petitioner and Patent Owner each filed respective Motions to Exclude Evidence. *See* Paper 38 ("Pet. Mot."); Paper 39 ("PO Mot."). Petitioner filed an Opposition to Patent Owner's Motion, Paper 40 ("Pet. Opp."), to which Patent Owner filed a Reply, Paper 42 ("PO Reply"). Patent Owner filed an Opposition to Petitioner's Motion, Paper 41 ("PO Opp."), to which Petitioner filed a Reply, Paper 43 ("Pet. Reply Opp.").

An oral hearing was held on May 13, 2020. A transcript of the hearing is included in the record. Paper 46 ("Tr.").

B. Real Parties in Interest

Petitioner identifies Foundation Medicine, Inc., Roche Holdings, Inc., Roche Finance Ltd., and Roche Holding Ltd. as the real parties-in-interest. Pet. 73. Patent Owner identifies Guardant Health, Inc., as the real party-in-interest. Paper 4, 2.

C. Related Matters

Patent Owner has asserted the '822 patent against Petitioner in *Guardant Health, Inc. v. Foundation Medicine, Inc.*, Case No. 17-cv-1616 (D. Del.) ("the co-pending litigation"). Pet. 74; Paper 4, 2. Patent Owner has also asserted the '822 patent against Personal Genome Diagnostics, Inc. ("PGDx") in *Guardant Health, Inc. v. Personal Genome Diagnostics, Inc.*, Case No. 17-cv-1623 (D. Del.). Pet. 74; Paper 4, 2.

Petitioner filed a second petition seeking *inter partes* review of the '822 patent, designated IPR2019-00653. Paper 4, 2. A Decision denying institution in that case was issued on August 19, 2019 (Paper 12), and a Decision denying Petitioner's request for rehearing issued on January 22, 2020 (Paper 14).

Petitioner also filed several petitions seeking *inter partes* review of patents related to the '822 patent, including: IPR2017-01170, IPR2017-

01447, and IPR2017-01448 (challenging U.S. Patent No. 9,340,830); IPR2019-00130 (challenging U.S. Patent No. 9,598,731); IPR2019-00634 (challenging U.S. Patent No. 9,840,743); and IPR2019-00636 and IPR2019-00637 (challenging U.S. Patent No. 9,902,992). Of these cases, only IPR2019-00634 is pending.

PGDx also filed petitions seeking post-grant review of the '822 patent, designated PGR2018-00058, and of U.S. Patent No. 9,840,743, designated PGR2018-00057. *Id.* at 2. Both petitions were dismissed before a decision on institution.

D. Summary of the '822 Patent

The '822 patent relates to methods for detecting rare mutations and copy number variations in cell free polynucleotides. Ex. 1001, code (57). The '822 patent states that cell-free DNA ("cfDNA"), found in different types of bodily fluids, may be used to detect and monitor disease. *Id.* at 1:29–45. For instance, cfDNA may contain genetic aberrations—like a change in copy number variations and/or single or multiple sequence variations associated with a particular disease—that can be used to detect or monitor such disease. *Id.* at 1:29–41, 30:8–14. The '822 patent states that "there is a need in the art for improved methods and systems for using cell free DNA to detect and monitor disease." *Id.* at 1:41–45.

The '822 patent states that the disclosed methods generally "comprise sample preparation, or the extraction and isolation of cell free polynucleotide sequence[s] from a bodily fluid; subsequent sequencing of cell free polynucleotides by techniques known in the art; and application of bioinformatics tools to detect rare mutations and copy number variations as compared to a reference." *Id.* at 30:4–14. The '822 patent states that "[s]ample preparation typically involves converting polynucleotides in a

sample into a form compatible with the sequencing platform used." *Id.* at 32:58–61. "This conversion [c]an involve tagging polynucleotides" with "polynucleotide sequence[s]." *Id.* at 32:61–63. The '822 patent refers to these polynucleotide sequences as "identifiers." *Id.* at 38:3–6. The identifier may be a molecular barcode. *Id.* at 38:6–7.

The '822 patent explains that "the efficient conversion of individual polynucleotides in a sample of initial genetic material into sequence-ready tagged parent polynucleotides" is an important tool "for detecting with high sensitivity genetic variation in a sample of initial genetic material." *Id.* at 32:33–39. According to the '822 patent, efficient conversion "increase[s] the probability that individual polynucleotides in a sample of initial genetic material will be represented in a sequence-ready sample" and "can produce sequence information about more polynucleotides in the initial sample." *Id.* at 32:39–42.

The '822 patent states that the parent polynucleotides may be tagged with either unique or non-unique identifiers. *Id.* at 37:44–49; *see also id.* at 3:10–15 (stating that, in some embodiments, the barcodes are unique, but in other embodiments, the barcodes are not unique); *see also id.* at 6:26–28 (stating that, in some embodiments, "each tagged parent polynucleotide in the set is uniquely tagged," whereas in other embodiments, "the tags are non-unique"). In the case of non-uniquely tagged parent polynucleotides, the '822 patent explains that "the use of non[-]unique barcodes, in combination with sequence data at the beginning (start) and end (stop) portions of individual sequencing reads and sequencing read length may allow for the assignment of a unique identity to individual sequences." *Id.* at 37:43–48.

E. Illustrative Claim

Claim 1 is the only independent claim. Claims 2–13 and 17–20 depend directly from claim 1. *See* Ex. 1001, 62:51–64:22. Claim 1 is illustrative and reproduced below:

1. A method, comprising:

a) providing a population of cell-free DNA ("cfDNA") molecules obtained from a bodily sample from a subject;

b) converting the population of cfDNA molecules into a population of non-uniquely tagged parent polynucleotides, wherein each of the non-uniquely tagged parent polynucleotides comprises (i) a sequence from a cfDNA molecule of the population of cfDNA molecules, and (ii) an identifier sequence comprising one or more polynucleotide barcodes;

c) amplifying the population of non-uniquely tagged parent polynucleotides to produce a corresponding population of amplified progeny polynucleotides;

d) sequencing the population of amplified progeny polynucleotides to produce a set of sequence reads;

e) mapping sequence reads of the set of sequence reads to one or more reference sequences from a human genome;

f) grouping the sequence reads into families, each of the families comprising sequence reads comprising the same identifier sequence and having the same start and stop positions, whereby each of the families comprises sequence reads amplified from the same tagged parent polynucleotide;

g) at each genetic locus of a plurality of genetic loci in the one or more reference sequences, collapsing sequence reads in each family to yield a base call for each family at the genetic locus;

h) determining a frequency of one or more bases called at the locus from among the families.

Ex. 1001, 62:18–48.

II. PETITIONER'S MOTION TO EXCLUDE

Petitioner moves to exclude Exhibits 2002, 2032, 2036, 2037, 2038, 2039, 2040, and 2041 in their entirety and Exhibits 2023 and 2025 in whole or in part. Pet. Mot. 1.

We *dismiss as moot* Petitioner' Motion to Exclude as it relates to Exhibits 2002, 2032, 2036, 2037, 2038, 2039, 2040, and 2041, for the following reasons. *First*, we do not rely on or cite to Exhibits 2032, 2036, 2037, or 2038 in this Decision. *Second*, Patent Owner cites to Exhibits 2002, 2039, 2040, and 2041 as support for its arguments about reasonable expectation of success. To the extent we refer to these exhibits herein, we determine that the record as a whole supports Petitioner's position as to reasonable expectation of success. *Infra* § IV.E.2.b. *Third*, Exhibit 2023 is the Declaration of Dr. Shendure that is limited in scope to the applicability of Schmitt to cfDNA. Ex. 2023 ¶¶ 16–37. Again, to the extent we refer to Exhibit 2023, we determine that the record as a whole supports Petitioner's position as to motivation to combine the Schmitt references with Fan or Forshew and as to reasonable expectation of success. *Infra* § IV.E.2.a–b. Thus, as to Exhibits 2002, 2023, 2032, 2036, 2037, 2038, 2039, 2040, and 2041, on which we do not rely to support our decision, Petitioner's Motion to Exclude is *dismissed as moot*.

Exhibit 2025 is the Declaration of Dr. Quackenbush. Petitioner contends that Dr. Quackenbush's Declaration should be excluded under Federal Rules of Evidence 702, 703, 705, and 401–403. Pet. Mot. 5–7.

TEMPUS EX1002.217

Specifically, Petitioner contends that we should exclude at least certain paragraphs of Dr. Quackenbush's Declaration because "they lack a disclosed basis of sufficient facts or data," "are not based on sufficient facts or data, the product of reliable principles and methods, and/or a reliable application of the principles and methods to the facts," and "are misleading, confusing, and/or needlessly cumulative." *Id.* at 5.

We have considered Petitioner's arguments but are not persuaded that Exhibit 2025 should be excluded in whole or in part. Patent Owner has shown, and Petitioner does not dispute, that Dr. Quackenbush is qualified to opine as to the perspective of an ordinarily skilled artisan at the time of the invention. PO. Opp. 5–93; *see also infra* § IV.B. As such, Dr. Quackenbush's testimony is highly relevant about how an ordinarily skilled artisan would have interpreted the prior-art references, as well as to the general knowledge in the field at the time the invention was made. Any deficiencies in Dr. Quackenbush's Declaration go to the weight that we should afford his testimony and do not support a motion to exclude. *See, e.g., Yorkey v. Diab*, 601 F.3d 1279, 1284 (Fed. Cir. 2010) (holding that the Board has discretion to give more weight to one item of evidence over another "unless no reasonable trier of fact could have done so"); *In re Am. Acad. of Sci. Tech Ctr.*, 367 F.3d 1359, 1368 (Fed. Cir. 2004) ("[T]he Board is entitled to weigh the declarations and conclude that the lack of factual corroboration warrants discounting the opinions expressed in the declarations."). Thus, as to Exhibit 2025, Petitioner's Motion is Exclude is *denied*.

III. PATENT OWNER'S MOTION TO EXCLUDE

Patent Owner moves to exclude Exhibits 1002, 1007, 1013, 1016–1020, 1022, 1036, 1037, 1055, 1080, 1082, 1084–1093, 1100, 1101, 1104,

TEMPUS EX1002.218

1110, and 1111. PO Mot. 1. We do not rely on Exhibits 1007, 1013, 1016–1020, 1022, 1036, 1037, 1055, 1080, 1082, 1084–1093, 1100, 1110, or 1111 in this Decision. Thus, we *dismiss as moot* Patent Owner's Motion to Exclude as it relates to Exhibits 1007, 1013, 1016–1020, 1022, 1036, 1037, 1055, 1080, 1082, 1084–1093, 1100, 1110, and 1111.

Exhibits 1002 and 1104 are the first and second Declarations of Dr. Gabriel, respectively. PO Mot. 1–8. Patent Owner argues that Exhibit 1002 should be excluded in its entirety because the "entire declaration is premised on an impossible standard to be met by a person of ordinary skill in the art" and "Dr. Gabriel uses this impossible perspective to support a hindsight-based obviousness analysis." *Id.* at 1–3. We agree with Petitioner, however, that Patent Owner's arguments constitute a disagreement about a question of fact (i.e., the proper definition of an ordinarily skilled artisan), and "[a] motion to exclude is not the proper mechanism to direct [its] attention to differences in the evidence." Pet. Opp. 2 (quoting *Google Inc. v. Visual Real Estate, Inc.*, IPR2014-01339, Paper 39 at 37 (PTAB Jan. 25, 2016)). Moreover, we note that Dr. Gabriel applies in her Declarations the same definition of an ordinarily skilled artisan as that we adopted in our Institution Decision and reaffirm here. Inst. Dec. 7–8; *infra* § IV.B. Thus, we are not persuaded that we should exclude Dr. Gabriel's Declarations Exhibits 1002 and 1104 for this reason.

Patent Owner also argues that Dr. Gabriel's second Declaration (Exhibit 1104) should be excluded because it attempts to "fill[] a gap in [Petitioner's] *prima facie* case," relies on "new theories regarding the teachings of Schmitt," and "presents an entirely different theory of obviousness with respect to the 'non-uniquely tagged' limitation." PO Mot. 3–8. We are not persuaded, however, that Exhibits 1002 and 1104

TEMPUS EX1002.219

should be excluded. "A motion to exclude is not a mechanism to argue that a reply contains new arguments or relies on evidence necessary to make out a prima facie case." *Vibrant Media, Inc. v. General Elec. Co.*, IPR2013-00170, Paper 56 at 31 (PTAB June 26, 2014). Thus, Patent Owner's arguments in this regard are improper.

In any event, Petitioner has shown, and Patent Owner does not dispute, that Dr. Gabriel is qualified to opine as to the perspective of an ordinarily skilled artisan at the time of the invention. Pet. Opp. 2–3; *see also infra* § IV.B. As with Dr. Quackenbush's testimony above, Dr. Gabriel's testimony is highly relevant about how an ordinarily skilled artisan would have interpreted the prior-art references, as well as to the general knowledge in the field at the time the invention was made. And any deficiencies in Dr. Gabriel's Declaration go to the weight that we should afford her testimony and do not support a motion to exclude.

For these reasons, as to Exhibits 1002 and 1104, Patent Owner's Motion to Exclude is *denied*.

IV. ANALYSIS

We have reviewed the parties' respective briefs as well as the relevant evidence discussed in those papers. For the reasons discussed in detail below, we determine that Petitioner has shown by a preponderance of the evidence that claims 1–11, 13, and 17–20 of the '822 patent are unpatentable under 35 U.S.C. § 103 as having been obvious, but not that claim 12 would have been unpatentable as obvious.

A. Principles of Law

To prevail in its challenges to the patentability of all claims of the '822 patent, Petitioner must demonstrate by a preponderance of the evidence that the claims are unpatentable. 35 U.S.C. § 316(e) (2012); 37 C.F.R.

TEMPUS EX1002.220

§ 42.1(d) (2018). "In an [*inter partes* review], the petitioner has the burden from the onset to show with particularity why the patent it challenges is unpatentable." *Harmonic Inc. v. Avid. Tech., Inc.*, 815 F.3d 1356, 1363 (Fed. Cir. 2016); *see also* 35 U.S.C. § 312(a)(3) (requiring *inter partes* review petitions to identify "with particularity . . . the evidence that supports the grounds for the challenge to each claim"). That burden of persuasion never shifts to Patent Owner. *Dynamic Drinkware, LLC v. Nat'l Graphics, Inc.*, 800 F.3d 1375, 1378 (Fed. Cir. 2015); *see also In re Magnum Oil Tools Int'l, Ltd.*, 829 F.3d 1364, 1375–78 (Fed. Cir. 2016) (discussing the burden of proof in *inter partes* review).

A claim is unpatentable for obviousness if, to one of ordinary skill in the pertinent art, "the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made." 35 U.S.C. § 103(a) (2006); *see also KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007). The question of obviousness is resolved on the basis of underlying factual determinations including the scope and content of the prior art, any differences between the claimed subject matter and the prior art, the level of ordinary skill in the art, and objective evidence of nonobviousness. *Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966). A petitioner cannot satisfy its burden of proving obviousness by employing "mere conclusory statements." *Magnum Oil*, 829 F.3d at 1380. Moreover, a decision on the ground of obviousness must include "articulated reasoning with some rational underpinning to support the legal conclusion of obviousness." *KSR*, 550 U.S. at 418 (citing *In re Kahn*, 441 F.3d 977, 988 (Fed. Cir. 2006)).

We analyze Petitioner's asserted grounds of unpatentability in accordance with the above-stated principles.

B. Level of Ordinary Skill in the Art

We consider the asserted grounds of unpatentability in view of the understanding of a person of ordinary skill in the art and, thus, begin with the level of ordinary skill in the art. The level of ordinary skill in the art is "a prism or lens through which . . . the Board views the prior art and the claimed invention" to prevent hindsight bias. *Okajima v. Bourdeau*, 261 F.3d 1350, 1355 (Fed. Cir. 2001).

Relying on the declaration testimony of its declarant, Dr. Gabriel, Petitioner contends that a person of ordinary skill in the art for the '822 patent "would have had a Ph.D. in genetics, molecular biology, bioinformatics or a related field, and at least five years of research in an academic or industry setting, including at least two to three years of research experience in the field of cancer genomics." Pet. 20 (citing Ex. 1002 ¶ 72). In response, Patent Owner does not appear to dispute the type of experience Petitioner proposes for the level of ordinary skill in the art. *See generally* PO Resp. We observe, however, that Patent Owner's declarants, Dr. Shendure and Dr. Quackenbush, contend that Petitioner's definition of an ordinarily skilled artisan relates more to an artisan with "extraordinary," rather than "ordinary," skill. *See* Ex. 2023 ¶ 15, Ex. 2025 ¶ 23.

In our Institution Decision, we preliminarily adopted Petitioner's proposed level of ordinary skill. Inst. Dec. 7–8. We also determined that the prior art itself was sufficient to demonstrate the level of ordinary skill in the art at the time of the invention. Inst. Dec. 8. For this Decision, we maintain that the prior art demonstrates the appropriate level of ordinary skill in the art. *See Okajima*, 261 F.3d at 1355 (the prior art, itself, can reflect appropriate level of ordinary skill in art).

TEMPUS EX1002.222

Nevertheless, for further clarity, we set forth the definition of an ordinarily skilled artisan as follows. As to level of education, because of the nature of the subject matter of the '822 patent, we agree with Petitioner that an ordinarily skilled artisan would have had a doctorate degree (Ph.D.) in genetics, molecular biology, bioinformatics, or a related field. Ex. 2023 ¶¶ 28–29; Ex. 1002 ¶ 25. We also agree with Petitioner that an ordinarily skilled artisan would have had five years' research experience in industry or academia, including at least two to three years of research experience in the field of cancer genomics.

We acknowledge Petitioner's contention that an ordinarily skilled artisan "would have had knowledge of DNA sequencing, including NGS [next-generation sequencing] and related sequencing methods, and related sample preparation techniques, bioinformatics methods for grouping and comparing sequence reads and mapping sequence reads onto genomes, and methods for identifying genetic variants in a sample." Pet. 20. But we conclude that these statements more aptly apply to the scope and content of the prior art under *Graham* and, thus, are best addressed in relation to Petitioner's asserted grounds of unpatentability based on obviousness.

Finally, we have considered the qualifications of Dr. Gabriel, Dr. Shendure, and Dr. Quackenbush and find that each is qualified to opine as to the perspective of an ordinarily skilled artisan at the time of the invention. *See* Ex. 1003 (Dr. Gabriel's *curriculum vitae*); Ex. 2024 (Dr. Shendure's *curriculum vitae*); Ex. 2026 (Dr. Quackenbush's *curriculum vitae*).

C. Claim Construction

The instant Petition was filed on February 2, 2019. Thus, the new rules amending the claim construction standard apply here because the

TEMPUS EX1002.223

Petition was filed after the November 13, 2018, effective date of the amendment. *See* Changes to the Claim Construction Standard for Interpreting Claims in Trial Proceedings Before the Patent Trial and Appeal Board, 83 Fed. Reg. 51,340, 51,358 (Oct. 11, 2018) (amending 37 C.F.R. § 42.100(b) effective November 13, 2018) (now codified at 37 C.F.R. § 42.100(b) (2019)). Accordingly, for this *inter partes* review, the Board applies the same claim construction standard as that applied in federal courts.

Under this standard, we construe the claim "in accordance with the ordinary and customary meaning of such claim as understood by one of ordinary skill in the art and the prosecution history pertaining to the patent." *Id.*; *see also Phillips v. AWH Corp.*, 415 F.3d 1303, 1312–13 (Fed. Cir. 2005) (en banc) (stating that claim terms "are generally given their ordinary and customary meaning" as understood by a person of ordinary skill in the art in question at the time of the invention). Only terms that are in controversy need to be construed, and then only to the extent necessary to resolve the controversy. *Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868 F.3d 1013, 1017 (Fed. Cir. 2017).

1. Overview

At the close of trial, the only claim-construction dispute remaining in this proceeding concerns the meaning of "non-uniquely tagged" in the context of a population of parent polynucleotides, as recited in claim 1. *See* Ex. 1001, 62:22–24 (reciting "converting the population of cfDNA molecules into a population of non-uniquely tagged parent polynucleotides"). Petitioner contends that "non-uniquely tagged" "means that the number of different identifiers attached to the polynucleotides is at least 2 and fewer than the number of polynucleotides." Pet. 21. Patent

TEMPUS EX1002.224

Owner argues that "non-uniquely tagged" means "the number of different identifiers is at least 2 and fewer than the number of polynucleotides *that map to the mappable base position*."[6] PO Resp. 16.

As an initial matter, we acknowledge that, in our Institution Decision, we referred to "non-uniquely tagged" as both "the number of different identifiers attached to the polynucleotides is at least 2 and fewer than the number of polynucleotides" and "the number of different identifiers attached to the polynucleotides is at least 2 and fewer than the number of polynucleotides that map to the mappable base position." *Compare* Inst. Dec. 9 ("Petitioner contends that '[t]he term "non-uniquely tagged" means that the number of different identifiers attached to the polynucleotides is at least 2 and fewer than the number of polynucleotides.' We agree." (quoting Pet. 21)), *with id.* (stating, "In the context of 'non-unique' identifiers, the '822 patent follows the words 'non-uniquely tagged' with '*that is*, the number of different identifiers can be at least 2 and fewer than the number of polynucleotides that map to the mappable base position.'" (quoting Ex. 1001, 41:42–46)). As a result, both parties claim that the Board adopted their respective construction of "non-uniquely tagged." *See* PO Resp. 15–16 (stating that "[t]here is no dispute as to the express definition of 'non-uniquely tagged' in the '822 patent"); Pet. Reply 3 (stating that "[t]he Board should reject Patent Owner's arguments and reaffirm its construction, which is identical to the District Court's").

[6] The Declarants for both parties agree that an ordinarily skilled artisan would understand that "mappable base position" means a position in the reference sequence to which polynucleotide molecules can be confidently mapped. *See* Ex. 2027 ¶ 67; Ex. 2022, 23:22–24:13, 30:4–18, 35:5–7; Ex. 2025 ¶ 65. We apply that understanding for this Decision.

To be clear, we have considered the totality of the arguments and evidence anew at the close of trial, as well as the reasoning set forth in our Institution Decision, and determine, for the reasons discussed below, that "non-uniquely tagged" should be construed to mean that "the number of different identifiers is at least 2 and fewer than the number of polynucleotides" in a sample, without the additional phrase "that map to a mappable position."

At the heart of the parties' dispute is the following passage from the '822 patent:

> Accordingly, this invention also provides compositions of tagged polynucleotides. The polynucleotides can comprise fragmented DNA, e.g. cfDNA. *A set of polynucleotides in the composition that map to a mappable base position in a genome can be non-uniquely tagged, that is, the number of different identifiers can be [] at least 2 and fewer than the number of polynucleotides that map to the mappable base position.*

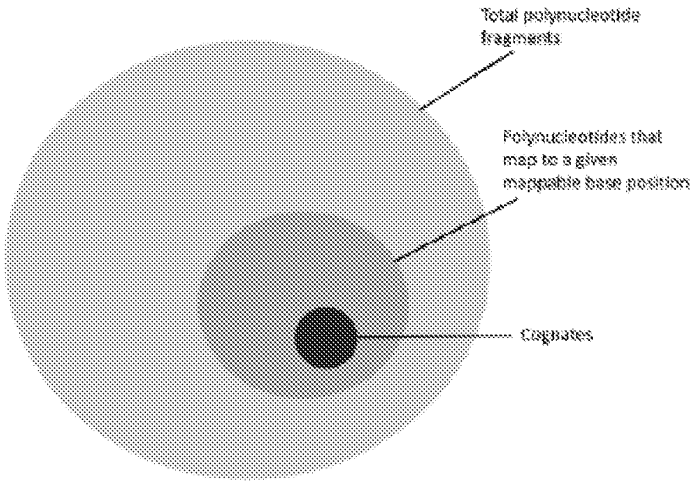
Ex. 1001, 41:42–47 (emphasis added).

The parties agree that "non-uniquely tagged" means that "the number of different identifiers [i.e., the tag count[7]] can be at least 2 and fewer than the number of polynucleotides," but disagree whether the phrase "that map to the mappable base position" should be included in the construction of "non-uniquely tagged." Put differently, the parties agree that, to be "non-uniquely tagged," the *lower limit* for the tag count is two, but disagree on the *upper limit* for the tag count.

[7] As relevant here, the '822 patent refers to the "number of unique identifiers" as the "tag count." Ex. 1001, 40:63. For clarity, we use these terms interchangeably in this Decision.

2. Patent Owner's Arguments

Patent Owner argues that the upper limit on the tag count is "fewer than the number of polynucleotides that map to the mappable base position." PO Resp. 15; *see also* PO Sur-Reply 6 (arguing that "the '822 patent expressly defines 'non-uniquely tagged' as limited by the number of polynucleotides that map to the mappable base position"). Relying on Dr. Quackenbush's Declaration, Patent Owner offers the following schematic distinguishing between (a) the total polynucleotide fragments in a sample, (b) the number of polynucleotides that map to a given mappable base position, and (c) cognates.[8] PO Resp. 13.



Schematic submitted by Patent Owner distinguishing between the polynucleotides in a sample. PO Resp. 13.

According to Patent Owner, the upper limit on the number of different identifiers does *not* depend on the total polynucleotide fragments in a sample (shown by the light blue circle). *Id.* at 14. Instead, the tag count "depends

[8] The '822 patent defines "cognates" (or "duplicates") as "more than one polynucleotide from different genomes [that] have the same start and stop positions." Ex. 1001, 40:4–8.

. . . on a subpopulation of fragments that map to a given position in the reference genome" (shown by the medium blue circle) "and the expected number of cognates" (shown by the dark blue circle). *Id.* (citing Ex. 1001, 41:6–13). Patent Owner argues that the '822 patent expressly defines "non-uniquely tagged" as limited by the number of polynucleotides that map to the mappable base position, and that any argument otherwise "strains credulity." PO Sur-Reply 5–6; *see also* PO Reply 15.

3. Petitioner's Arguments

Petitioner contends that the upper limit on the number of different identifiers is "fewer than the number of polynucleotides" in a sample. Pet. 21. Specifically, Petitioner contends that, when "fewer barcodes than original DNA fragments" are used, "not every fragment has a unique barcode," and thus the population of polynucleotides are "non-uniquely tagged." *Id.* at 11–12. As an example, Petitioner contends that "if there are 1000 DNA fragments and only 500 different barcodes, [then] each fragment does not have its own unique barcode," and thus the population of polynucleotides is "non-uniquely tagged" in the context of claim 1. *Id.* at 12 (citing Ex. 1002 ¶ 51). Conversely, "if there are 1000 original DNA fragments and at least 1000 different barcodes, [then] each fragment will have its own unique barcode." *Id.* at 11 (citing Ex. 1002 ¶ 50). Petitioner contends that we should not include "that map to the mappable base position" in the construction of "non-uniquely tagged" because that phrase represents "a single embodiment" of the '822 patent and because that construction "is inconsistent" with Patent Owner's arguments in District Court about the meaning of "non-uniquely tagged." Pet. Reply 3–10.

TEMPUS EX1002.228

4. Analysis

We begin with the words of claim 1. *See Amgen Inc. v. Hoechst Marion Roussel, Inc.*, 457 F.3d 1293, 1301 (Fed. Cir. 2006) ("claim construction must begin with the words of the claims themselves"). The first step of claim 1 recites "providing *a population* of cell free DNA ('cfDNA') molecules," and the second step recites "converting *the population* of cfDNA molecules into a population of non-uniquely tagged parent polynucleotides." Ex. 1001, 62:19–24 (emphases added). We observe that nothing in the plain language of claim 1 requires the population of cfDNA molecules that undergoes tagging (i.e., the population that is converted into a population of "non-uniquely tagged parent polynucleotides") to be limited to a population where each polynucleotide maps to a mappable base position. Instead, claim 1 recites that *the* population (i.e., not necessarily the mappable population) is converted into non-uniquely tagged parent polynucleotides. *Id.* For these reasons, we determine that the words of claim 1 do not support Patent Owner's argument that the tag count is limited by the number of polynucleotides that map to a mappable base position.

We now turn to the intrinsic record. Starting with the passage reproduced above that we identified as the heart of the parties' dispute (Ex. 1001, 41:42–47), we acknowledge that the use of "that is" (or "i.e.") typically "signals an intent to define the word to which it refers." *Edwards Lifesciences LLC v. Cook Inc.*, 582 F.3d 1322, 1334 (Fed. Cir. 2009). Here, however, we agree with Petitioner that the use of "non-uniquely tagged" in this passage of the written description refers to a specific embodiment of the '822 patient—that is, an embodiment where *the set* of polynucleotides map to a mappable base position. *See* Pet. 4–5. The '822 patent describes various "sets" of polynucleotides and states that, "[*i*]*n some embodiments*

TEMPUS EX1002.229

each polynucleotide in a set is mappable to a reference sequence." Ex. 1001, 6:1–2 (emphasis added). The '822 patent also states that, "*[i]n some embodiments*, each set of parent polynucleotides is mappable to a position in a reference sequence, and the polynucleotides in each set are not uniquely tagged." *Id.* at 18:39–42 (emphasis added). That each polynucleotide in a set is mappable *in some embodiments* necessarily implies that, in other embodiments, not every polynucleotide is mappable. Moreover, that each parent polynucleotide is mappable and non-uniquely tagged *in some embodiments* suggests that mappability and non-unique tagging are not necessarily concomitant. *See* W*esternGeco LLC v. ION Geophysical Corp.*, 889 F.3d 1308, 1323–24 (Fed. Cir. 2018) ("It is well established that claims are not limited to preferred embodiments, unless the specification clearly indicates otherwise.").

Next, we determine that other portions of the '822 patent do not support Patent Owner's argument that the upper limit of the tag count depends on the number of polynucleotides that map to a mappable base position. PO Resp. 13–14. The '822 patent explains that, when polynucleotides have a tag count equal to 1 ("equivalent to having no unique tags or not tagging"), "it may not be possible to determine which sequence reads are derived from which parent molecules." *Id.* at 40:34–61. The '822 patent further explains that "[t]his problem can be diminished by tagging parent molecules with a sufficient number of unique identifiers." *Id.* at 40:61–63. According to the '822 patent, the methods of the prior art "uniquely tag[ged] every, or nearly every, different parent molecule in the sample." *Id.* at 40:67–41:2. But, these prior art methods were "cumbersome and expensive," because they required "billions of different unique identifiers." *Id.* at 41:2–6.

TEMPUS EX1002.230

The '822 patent explains that "[t]his invention" solves the problem of distinguishing polynucleotides by tagging a population of DNA "with n different unique identifiers, wherein n is at least 2 and no more than 100,000*z, wherein z is a measure of the central tendency (e.g., mean, median, mode) of an expected number of duplicate molecules having the same start and stop positions" (i.e., cognates). *Id.* at 41:6–13. The '822 patent describes embodiments where the lower limit of the tag count is between 2*z to 20*z, and other embodiments where the upper limit of the tag count is between 100*z to 100,000*z. *Id.* at 41:13–18.[9] The '822 patent then expressly states that "n can range from any combination of these upper and lower limits." *Id.* at 41:18–19.

Thus, the '822 patent clearly teaches the ordinarily skilled artisan that the number of different identifiers—that is, the tag count "n"—is based on the number of expected cognates in a sample ("z"), which in turn "is a function of the number of haploid genome equivalents in a sample and the distribution of fragment sizes." Ex. 1001, 40:8–11; 41:6–39; *see also* PO Resp. 12 ("The inventor of the '822 patent determined empirically that the number of cognates in a sample is proportional to the size distribution of the polynucleotide fragments and the number of genomic equivalents—that is, the number of haploid genome copies present in the sample.").

Patent Owner does not persuasively point us to any teaching in this description—which Dr. Quackenbush admits describes a "non-uniquely

[9] The '822 patent's reference to "1000,000*z" at column 41, line 17, appears to be a typographical error. *See* Ex. 2001 (Patent Owner's claim-construction brief quoting the same passage from the '822 patent as "2 and no more than 100,000*z"). For this Decision, we read that term as "100,000*z."

tagged" embodiment of the '822 patent—that instead correlates the upper limit of the tag count to a *subpopulation* of polynucleotides in a sample that map to a mappable base position. *See* Ex. 2025 ¶ 58 (citing Ex. 1001, 41:6–13); Ex. 1103, 36:18–21, 37:7–11, 14–16.[10] Rather, according to the '822 patent—in an embodiment described as "[t]his invention"—the tag count can encompass up to 100,000*z different identifiers. Ex. 1001, 41:6–13. The '822 patent also teaches that, "in a sample of about 10,000 haploid genome equivalents of human DNA, there are about 3 duplicate polynucleotides beginning at any given position." *Id.* at 40:24–27, 41:10–13. Thus, for a sample comprising 10,000 haploid copies of polynucleotides, z=3, the sample can be tagged with up to 100,000*3 (300,000) identifiers and still be considered "non-uniquely tagged." *Id.* at 41:6–13. This number of unique identifiers (300,000) is far greater than the number of polynucleotides that would map to a given mappable base position (10,000) in a sample of 10,000 haploid genome equivalents under Patent Owner's construction. *See* PO Resp. 17–19 (arguing that for "for non-uniquely tagged," a single copy of the genome would produce 1 polynucleotide that maps to a mappable base position, and five haploid genome equivalents would produce 5 polynucleotides that map to a mappable base position).

[10] Indeed, claim 12 contrasts the expected number of duplicate molecules with the population of polynucleotides *in the sample*. *See* Ex. 1001, 63:15–19 (Claim 12, reciting "The method of claim 1, wherein *the population of polynucleotides* is tagged with n different unique identifiers, wherein n is no more than 100*z, wherein z is a mean of an expected number of duplicate molecules having the same start and stop positions *in the sample*." (emphases added)).

Patent Owner appears to acknowledge that this embodiment would be outside the scope of claim 1, but nevertheless argues that this "is of no moment" because it "is not required to claim all that it discloses." PO Sur-Reply 6–7. Even so, we find that Patent Owner has not provided an adequately persuasive reason, based on either the claim language or the '822 patent, for doing so here. *See SynQor, Inc. v. Artesyn Techs., Inc.*, 709 F.3d 1365, 1378–79 (Fed. Cir. 2013) ("A claim construction that excludes the preferred embodiment is rarely, if ever, correct and would require highly persuasive evidentiary support." (quotation marks and citation omitted)); *see also In re Katz Interactive Call Processing Patent Litig.*, 639 F.3d 1303, 1324 (Fed. Cir. 2011) ("there is a strong presumption against a claim construction that excludes a disclosed embodiment").

Turning to the extrinsic evidence, we note that Patent Owner argued in its Opening Brief on Claim Construction to the District Court in the co-pending litigation that, as to the construction of "non-uniquely tagged," "[t]he dispute between the parties is whether the number of 'non-unique' barcodes can approach the total number of polynucleotides." Ex. 2001, 7. This accords with the dispute between the parties in this proceeding as well. But before the District Court, unlike here, Patent Owner argued that "the total number of barcodes must plainly be less than the *total number of polynucleotides*." *Id.* at 9 (emphasis added). Patent Owner also argued that "in describing the embodiments that use 'non-unique' barcoding, the specification explains that they use between '2 and no more than 100,000*z' barcodes 'wherein z is a measure of central tendency (e.g., mean, median, mode) of an expected number of duplicate molecules.'" *Id.* at 8 (quoting Ex. 1001, 41:10–13). Thus, unlike here, Patent Owner did not rely on the

TEMPUS EX1002.233

"number of polynucleotides that map to the mappable base position" as setting forth the upper limit for the number of different identifiers. *Id.*

Patent Owner argues that its previous statements to the District Court were made simply to distinguish the claims over the prior art at issue in the co-pending litigation and are not inconsistent with its construction of "non-uniquely tagged" here. PO Sur-Reply 6. We disagree. In its claim-construction brief, Patent Owner criticized the opposing parties' proposed constructions as "uniformly at odds with the *intrinsic* record," Ex. 2001, 1 (emphasis added), and argued that its own construction of "non-uniquely tagged" was most consistent with the specification and the plain meaning of "non-uniquely tagged," *id.* at 7–9. Although we do not consider Patent Owner's arguments in the co-pending litigation dispositive, we find that they nonetheless further support our construction of "non-uniquely tagged" here.

For all the above reasons, we determine that "non-uniquely tagged" means that "the number of different identifiers can be at least 2 and fewer than the number of polynucleotides" in a sample. Thus, a population of parent polynucleotides is "non-uniquely tagged" in the context of claim 1 when the number of different identifiers (the tag count) is at least 2 but fewer than the number of parent polynucleotides in the sample.

D. Asserted References

Before turning to Petitioner's asserted grounds of unpatentability, we provide a brief summary of the asserted references.

1. Schmitt (Ex. 1011)

Schmitt relates to a method for sequencing target DNA fragments called Duplex Consensus Sequencing ("DCS"). Ex. 1011, code (57). According to Schmitt, the DCS method greatly reduces sequencing errors by independently tagging and sequencing each of the two strands of a target

TEMPUS EX1002.234

DNA fragment. *Id.* Because the two strands of DNA are complementary, true mutations can be found at the same position on both strands, as opposed to on a single strand if PCR or sequencing errors occur. *Id.*

Schmitt's DCS methods utilize "a single molecule identifier (SMI) adaptor molecule for use in sequencing a double-stranded target nucleic acid molecule." Ex. 1011, 2:66–3:1. Schmitt teaches that the SMI adaptor molecule comprises an SMI sequence or "tag" (which is a degenerate or semi-degenerate n-mer sequence[11]) and an SMI ligation adaptor (which allows the SMI adaptor molecule to be ligated to the target DNA fragment). *Id.* at 3:1–9; *see also id.* at 5:57–59 (providing that the "SMI adaptor molecule is double stranded, and may include a single molecule identifier (SMI) sequence, and an SMI ligation adaptor"); *id.* at 6:46–47 (referring to the SMI sequence as a "tag"). The DCS method includes the steps of:

> ligating a double-stranded target nucleic acid molecule to at least one SMI adaptor molecule to form a double-stranded SMI-target nucleic acid complex; amplifying the double-stranded SMI-target nucleic acid complex, resulting in a set of amplified SMI-target nucleic acid products; and sequencing the amplified SMI-target nucleic acid products.

Id. at 3:12–20.

In one embodiment, the SMI sequence is "a unique, double-stranded, complementary n-mer random tag," such that every DNA fragment becomes labeled with two distinct SMI sequences. *Id.* at 3:47–53. In this embodiment, the "nucleotide n-mer sequences may be any suitable length to produce a sufficiently large number of unique tags to label a set of sheared

[11] It is well known in the art that an "n-mer" is a short sequence comprising "n" number of nucleotides. For example, a "4-mer" polynucleotide is a short sequence of four nucleotides in a DNA sequence.

DNA fragments from a segment of DNA." *Id.* at 6:59–63. Schmitt provides as an example a "nucleotide n-mer sequence which is 12 nucleotides in length" that, once ligated to each end of the target DNA fragment, "results in the generation of up to 4^{24} (i.e., 2.8×10^{14}) distinct tag sequences." *Id.* at 6:66–7:5.

In another embodiment, referred to as the "hybrid method," the SMI sequence comprises "a shorter n-mer tag (such as 1 or 2 or 3 or 4 or more degenerate or semi-degenerate bases)." *Id.* at 9:9–13. In this embodiment, the SMI adaptor molecules "serve as unique molecular identifiers" by combining information from the sheared ends of the target DNA fragment and the short n-mer tag. *Id.*

Schmitt teaches that, for error correction through DCS, sequence reads sharing a unique set of SMI tags are grouped into paired families, each pair reflecting one double-stranded DNA fragment. *Id.* at 4:4–10. Mutations present in only one or a few family members, or mutations occurring in only one of the two strands, represent sequencing mistakes or PCR-introduced errors. *Id.* at 4:10–18. True mutations are present on both strands and appear in all members of a family pair. *Id.* at 4:18–20.

2. Schmitt 2012 (Ex. 1047)

Schmitt 2012 also relates to the DCS method and uses the same library-preparation techniques and sequencing methods disclosed in Schmitt. Ex. 1047, Abstract, 14509 (Fig. 1). Schmitt 2012 discloses a "data processing" workflow comprising the steps of filtering sequence reads, aligning reads to a reference genome, grouping reads containing identical tag sequences to form single-strand consensus sequence reads ("SSCS"), re-aligning reads to the reference genome, and identifying and comparing

partner strands among SSCS reads compared to form DCS reads. *Id.* at 14513, SI1–SI2; *see also* Ex. 1002 ¶ 91.

3. Fan (Ex. 1048)

Fan relates to a method for diagnosing fetal aneuploidy by directly sequencing cfDNA from the plasma of pregnant women with high-throughput shotgun sequencing technology. Ex. 1048, Abstract. In doing so, Fan was able to measure the over- and under-representation of chromosomes from an aneuploidy fetus. *Id.*

4. Forshew (Ex. 1004)

Forshew relates to a method for identifying cancer mutations present in cfDNA using tagged-amplicon deep sequencing ("TAm-Seq"). Ex. 1004, Abstract. Tagged-amplicon deep sequencing allows amplification and deep sequencing of genomic regions spanning thousands of bases. *Id.* at 1. Forshew applied the technique to both abundant and rare mutations in circulating DNA from blood plasma of ovarian and breast cancer patients. *Id.* at 1–2.

E. Obviousness over Schmitt in View of Schmitt 2012, and Further in View of Fan or Forshew

Petitioner contends that claims 1–13 and 17–20 of the '822 patent are unpatentable as obvious over Schmitt, Schmitt 2012, and Fan or Forshew. Pet. 30–70. Patent Owner opposes. PO Resp. 19–62; PO Sur-Reply 3–26. Having considered the totality of the arguments and evidence, we find that Petitioner has shown by a preponderance of the evidence that claims 1–11, 13, and 17–20 are unpatentable as having been obvious over Schmitt, Schmitt 2012, and Fan or Forshew, but has not shown that claim 12 would have been obvious.

TEMPUS EX1002.237

1. Limitations of the challenged claims

Petitioner contends that the combination of Schmitt, Schmitt 2012, and Fan or Forshew discloses or suggests each element of the challenged claims. Petitioner presents arguments mapping the language of claims 1–13 and 17–20 to the disclosures of each reference. Pet. 41–71.

a. *Independent Claim 1*

Step (a) of claim 1 recites "providing a population of cell-free DNA ('cfDNA') molecules obtained from a bodily sample from a subject." Ex. 1001, 62:19–21. We agree with Petitioner that Fan and Forshew teach this limitation. Pet. 41–42. Fan describes extracting cfDNA from the plasma of pregnant women to detect fetal aneuploidy. *See* Ex. 1048, 16266 ("We directly sequenced cell-free DNA with high-throughput shotgun sequencing technology form plasma of pregnant women . . . ."), 16270 (stating that "DNA was extracted from cell-free plasma"). Specifically, Fan extracted between 1.2 and 8 ng of cfDNA from plasma samples ranging from 1.3 to 3.2 ml. *Id.* at SI7 (Table S1); *see also* Ex. 1002 ¶ 122. Similarly, Forshew describes extracting cfDNA from plasma samples to identify mutations in cancer patients. *See* Ex. 1004, 1 (stating that "[p]lasma of cancer patients contains cell-free tumor DNA"), 10 (stating that "[c]irculating DNA was extracted from between 0.85 and 2.2 ml of plasma"). As shown in Table S6, Forshew obtained amounts ranging from 0.9 to 19.7 ng of cfDNA from plasma samples of cancer patients. *Id.* at 32 (Table S6); *see also* Ex. 1002 ¶ 23.

Turning to step (b) of claim 1, the first portion of this step recites "converting the population of cfDNA molecules into a population of non-uniquely tagged parent polynucleotides." Ex. 1001, 62:22–24. We agree with Petitioner that Schmitt in combination with Fan or Forshew teaches this

TEMPUS EX1002.238

limitation. Pet. 42–46. Schmitt teaches that, in the DCS method, SMI adaptor molecules comprising an "SMI sequence (or 'tag') of nucleotides" and an SMI ligation adaptor are ligated to the ends of target DNA polynucleotides. *See* Ex. 1011, 3:1–9, 5:57–59, 6:46–51. Schmitt refers to the SMI sequences as a "nucleotide n-mer" sequence. *Id.* at 6:46–66.

As explained above, we construe "non-uniquely tagged" to mean that the number of different identifiers attached to the parent polynucleotides is at least 2 but fewer than the number of polynucleotides in the sample. *Supra* § IV.C.4. Schmitt discloses a DCS "hybrid method" tagging target DNA polynucleotides that uses "a combination of sheared ends and *a shorter n-mer tag* (such as 1 or 2 or 3 or 4 or more degenerate or semi-degenerate bases) in the adaptor." Ex. 1011, 9:9–11 (emphasis added); *see also* Pet. 43–46; Ex. 1002 ¶¶ 126–127. Schmitt also expressly discloses examples of 4-mer tags. *See* Ex. 1011, 4:30–54; *id.* at 9:9–13. Dr. Gabriel testifies without rebuttal that using a 4-mer tag yields 4^4, or 256, different tag sequences—which when ligated at both ends of a parent polynucleotide—yields 256^2, or 65,536 different or "unique" identifiers. *See* Ex. 1002 ¶ 127 & n.7; Pet. 44–45; *see also* Ex. 2025 ¶ 69 (testimony of Dr. Quackenbush as to "[t]he 65,536 possible unique tags from a 4-mer sequence"); PO Resp. 18 (accord).

We agree with Petitioner that the use of n-mer tags in the hybrid DCS method results in a population of "non-uniquely tagged parent polynucleotides," because the n-mer tags are not "unique" for each parent polynucleotide. That is, the number of different n-mer tags is fewer than the number of polynucleotides in the sample because Schmitt explicitly distinguishes the hybrid method from embodiments in which every polynucleotide in the sample is tagged with a unique SMI identifier. *See* Ex. 1011, 9:1–4 (stating that the DCS method "does *not strictly require the*

use of an SMI tag, as the sheared ends can be used as identifiers to differentiate unique individuals molecules from PCR duplicates" (emphasis added)); *see also id.* at 6:59–63 (stating that, in the main embodiment, the "nucleotide n-mer sequences may be any suitable length to produce a sufficiently large number of unique tags to label a set of sheared DNA fragments from a segment of DNA"). Finally, although Schmitt does not expressly teach that the target polynucleotide is cfDNA, both Fan and Forshew teach the extraction and analysis of cfDNA as described immediately above.

The second portion of step (b) provides that "each of the non-uniquely tagged parent polynucleotides comprises (i) a sequence from a cfDNA molecule of the population of cfDNA molecules, and (ii) an identifier sequence comprising one or more polynucleotide barcodes." Ex. 1001, 62:24–28. We again agree with Petitioner that Schmitt teaches this limitation. Pet. 45. Schmitt describes embodiments (a–c) in Figure 3, in which sequence reads are derived from parent polynucleotides comprising both a "strand identifier" and a sequence derived from the parent DNA molecule. Ex. 1011, 4:4–54; *see also* Ex. 1002 ¶ 129.

Turning to the remaining limitations of claim 1, we agree with Petitioner's contentions and supporting evidence about steps (c) through (h) of the claimed method. *See* Pet. 47–56. In particular, we agree with Petitioner's contentions that Schmitt, either alone or in combination with Schmitt 2012, discloses:

- amplifying the tagged polynucleotides and sequencing the amplified progeny (steps (c) and (d) of claim 1), *see* Pet. 47; Ex. 1011, 3:10–20 (disclosing "amplifying the double-stranded SMI-target nucleic acid complex, resulting in a set of amplified SMI-target nucleic acid

products; and sequencing the amplified SMI-target nucleic acid products"); *see also id.* at 21:55–57; Ex. 1002 ¶¶ 132–133;

- mapping the sequence reads to a reference sequence (step (e) of claim 1), *see* Pet. 47–49; Ex. 1011, 20:39–64 (teaching aligning sequence reads "to the human genome with the Burrows Wheeler Aligner (BWA)"); *id.* at 23:10–14, 24:33–37; Ex. 1047, SI1 ("Reads were then aligned to the reference genome with the Burrows Wheeler aligner (BWA) and nonmapping reads were discarded. . . . Reads sharing identical tag sequences were then grouped together and collapsed to consensus reads."); *see also* Ex. 1002 ¶¶ 134–137, 139;
- grouping the sequence reads into families (step (f) of claim1), *see* Pet. 40–51; Ex. 1011, 4:9–10, 9:9–14, 17:61–18:2, 20:39–64, 21:55–61; *see also* Ex. 1002 ¶¶ 140–144;
- collapsing sequence reads in each family to yield a "base call" (step (g) of claim 1), *see* Pet. 51–55; Ex. 1011, 4:20–29, 4:55–66, 18:44–61, 20:50–56; 22:50–53, 23:19–22, 23:61–24:1, 25:20–30, Figures 3 and 5C; *see also* Ex. 1002 ¶¶ 145–152; and
- determining the frequency of bases called at the locus from among the families (step (h) of claim 1), *see* Pet. 55–56; Ex. 1011, 4:25–29, Figure 3; *see also* Ex. 1002 ¶¶ 154–155.

b. *Analysis of Patent Owner's Arguments*

We are not persuaded by Patent Owner's arguments that the combination of Schmitt, Schmitt 2012, and Fan or Forshew fails to teach or suggest all the limitations of claim 1. PO Resp. 19–56; PO Sur-Reply 3–26. Patent Owner first argues that Schmitt fails to disclose "non-uniquely tagged parent polynucleotides" as this phrase should be correctly construed, and that the Petition fails because Petitioner never applied the correct

TEMPUS EX1002.241

construction of "non-uniquely tagged" to the prior art in the Petition. PO Resp. 20–23; PO Sur-Reply 3–4. But, as explained in detail above, we do not read "non-uniquely tagged" as limiting the tag count to a subpopulation of polynucleotides that map to a mappable base position. *Supra* § IV.C.4. Thus, we disagree with Patent Owner that Petitioner failed to apply the correct construction of "non-uniquely tagged parent polynucleotides" in its Petition.

Patent Owner's next argument, that "Petitioner's contrived scenario does not support its assertion that Schmitt teaches non-uniquely tagged parent polynucleotides," PO Resp. 23–27, also relies on its unduly narrow construction of "non-uniquely tagged parent polynucleotides." In this regard, Patent Owner argues that "Schmitt is silent regarding a sub-population of fragments in a sample that map to a given mappable base position," *id.* at 20–21, and that an ordinarily skilled artisan would not have been able to determine the number of polynucleotides that map to a mappable base position from Schmitt's disclosure, *id.* at 23 (citing Ex. 2025 ¶ 72). But, again, we do not read "non-uniquely tagged parent polynucleotides" as limiting the tag count to a subpopulation of polynucleotides that map to a mappable base position. *Supra* § IV.C.4. Patent Owner's argument therefore fails, and its follow-on arguments about Dr. Quackenbush's testimony and Dr. Gabriel's cross-examination testimony as to the tagging and sequencing of one haploid human genome equivalent, *see* PO Resp. 24–26, are irrelevant.[12]

[12] Similarly, Patent Owner's argument that "Schmitt does not disclose sequencing targets as large as the entire 3 billion base pair haploid human genome," PO Resp. 26, is unpersuasive because nothing in claim 1 limits the size of the sequencing target. *See* Ex. 1001, 62:18–48; *see also In re Hiniker*

TEMPUS EX1002.242

We also disagree with Patent Owner's argument that Schmitt fails to disclose "non-uniquely tagged parent polynucleotides." PO Resp. 27–33 (arguing that "Schmitt exclusively teaches unique tagging methods"). As explained above, *supra* § IV.D.1., Schmitt discloses methods for sequencing target parent polynucleotides that utilize SMI sequences as tags. Ex. 1011, 2:66–3:1. In the main embodiment, the SMI sequence is "a *unique*, double-stranded, complementary n-mer random tag," and every target polynucleotide is labeled with two distinct SMI sequences. *Id.* at 3:47–53 (emphasis added). In this embodiment, Schmitt teaches that the "nucleotide n-mer sequences may be any suitable length to produce a sufficiently large number of unique tags to label a set of sheared DNA fragments from a segment of DNA." *Id.* at 6:59–63. Schmitt provides as an example a "nucleotide n-mer sequence which is 12 nucleotides in length," that, once ligated to each end of the target DNA fragment, "results in the generation of up to 4^{24} (i.e., 2.8×10^{14}) distinct tag sequences." *Id.* at 6:66–7:5. There is no dispute that this embodiment of Schmitt is directed to a *unique* tagging method because the number of SMI tags exceeds the number of parent polynucleotides in the sample. *See* Ex. 1002 ¶ 126 (testimony of Dr. Gabriel that "[i]n some embodiments, Schmitt describes the use of a large number of *unique* barcode sequences such that there is a high probability that a different barcode sequence is ligated to each parental DNA template").

But, contrary to Patent Owner's arguments otherwise, Schmitt also teaches an alternative "hybrid" method that we conclude results in the generation of "non-uniquely tagged parent polynucleotides" as claimed.

Co., 150 F.3d 1362, 1369 (Fed. Cir. 1998) ("the name of the game is the claim").

TEMPUS EX1002.243

Specifically, Schmitt teaches that the DCS method "does not strictly require the use of an SMI tag" because the sheared ends of the target parent DNA polynucleotides "can be used as identifiers to differentiate unique individual molecules from PCR duplicates." Ex. 1011, 9:1–4. Schmitt acknowledges, however, that "there are a limited number of shear points flanking any given genomic position and thus the power to sequence deeply is increased via inclusion of the SMI tag." *Id.* at 9:6–9. Schmitt thus proposes a solution that utilizes both the sequence information from the sheared ends and an n-mer: the "hybrid" method. *Id.* at 9:9–13. Specifically, Schmitt teaches that the hybrid method uses "a shorter n-mer tag (such as 1 or 2 or 3 or 4 or more degenerate or semi-degenerate bases)" in combination with the sheared ends of the target DNA fragment to "serve as unique molecular identifiers." *Id.*

We find credible, and supported by preponderant record evidence, Dr. Gabriel's testimony that the *combination* of the sheared ends and a *non-unique* short n-mer tag results in a "*unique* molecular identifier[]." Ex. 1002 ¶ 126 (emphasis added); *see also* Ex. 1011, 9:1–13. Specifically, we are persuaded by Dr. Gabriel's testimony that the short n-mer tag used in Schmitt's hybrid method is itself *not unique* for each parent polynucleotide in the sample. Ex. 1002 ¶ 127. Schmitt expressly teaches that the "shorter n-mer tag" may comprise 4 or fewer nucleotide bases. Ex. 1011, 9:10–12. From this disclosure—as Dr. Gabriel persuasively testifies and we agree—an ordinarily skilled artisan would have readily recognized that a 4-mer, for example, yields 4^4, or 256, different tag sequences, which when ligated at both ends of a parent polynucleotide, would yield 256^2, or 65,536 different or "unique" identifiers. Ex. 1002 ¶ 127 & n.7; *see also* Pet. 44–45. Indeed, Dr. Quackenbush acknowledges that "65,536 possible unique tags [result] from a 4-mer tag sequence." Ex. 2025 ¶ 69.

TEMPUS EX1002.244

An ordinarily skilled artisan would have readily recognized that this number of unique identifiers is far less than the millions or more parent polynucleotides that would result from Schmitt's size selection for DNA fragments in the range of 200 to 500 base pairs. Ex. 1002 ¶ 127 (citing Ex. 1011, 22:43–46 (stating that "DNA for sequencing was sheared and end-repaired by standard methods, with size-selection for fragments in the range of ~200–500 bp by size-selecting binding")). Thus, as Dr. Gabriel testifies and we agree, more than one parent polynucleotide would necessarily share the same short n-mer tag in Schmitt's hybrid method, because the number of unique identifiers (i.e., a tag count of 65,536) is at least 2 but fewer than the number of parent polynucleotide fragments in the sample (i.e., millions or more), as we have construed "non-uniquely tagged parent polynucleotides." *Id.*

Patent Owner argues that Petitioner was wrong to "assume a combination of more fragments in a sample (i.e., millions of fragments) than SMI tags (i.e., 65,536 tags) in order to conclude that the number of fragments is greater than the number of tags." PO Resp. 28. In this regard, Patent Owner argues that "Schmitt discloses applying DCS to much smaller targets—'ranging from ~300 bp to ~20 kb in size'" and that shearing those targets into 200 bp fragments would result in, at most, between 1 and 100 fragments, "far fewer fragments than 'millions of fragments'" Petitioner asserts would be in a sample. *Id.* at 29 (citing Ex. 1011, 19:65–66, 24:7–11, Ex. 2025 ¶ 85). Although we have carefully considered Patent Owner's argument, we find that it lacks evidentiary support in the disclosure of Schmitt and, therefore, is not persuasive.

Schmitt clearly discloses shearing 3 micrograms of DNA into fragments, ligating the SMI adaptor molecules to 750 ng of T-tailed DNA,

TEMPUS EX1002.245

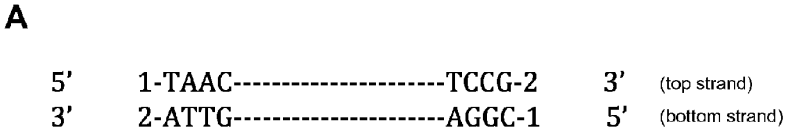
and amplifying (by PCR) 375 ng of the resulting adaptor-ligated (or tagged) DNA. Ex. 1011, 19:58–20:23. Then, for sequence analysis, Schmitt discloses that "an arbitrary 758 kb region of the genome consisting of both coding and noncoding sequences" was "targeted" and "capture[d]" using capture baits. *Id.* at 20:27–30. The much smaller fragments to which Patent Owner refers—i.e., DNA molecules ranging from ~300 bp to ~20 kb in size—represent DNA targets that would have been similarly captured for the purpose of sequence analysis *after* the process of shearing and tagging. *See id.* at 20:16–30 (describing the process of pre-capture amplification and DNA capture); 24:7–9 (describing the ~300 bp to ~20 kb DNA molecules as "vastly smaller *target* DNA molecules" (emphasis added)); Ex. 1104 ¶ 18 (explaining that "Schmitt discloses tagging DNA fragments *before* target capture, not after"). An ordinarily skilled artisan would have understood from Schmitt's disclosure, then, that these ~300 bp to ~20 kb DNA molecules would not have been subjected to shearing and ligation reactions, because those molecules would have been already tagged with the SMI sequence identifiers. Ex. 1104 ¶ 18. For these reasons, Patent Owner's argument (and Dr. Quackenbush's corresponding testimony) about the number of polynucleotide fragments that would result from *further shearing* of 300 bp to ~20 kb DNA molecules reflects a misreading of Schmitt's disclosure and is not persuasive. *See* Ex. 2025 ¶ 85 (Dr. Quackenbush's testimony "[a]ssuming" shearing of Schmitt's ~300 bp to ~20 kb DNA targets).

Moreover, we find that an ordinarily skilled artisan would have understood that the total number of polynucleotides in Schmitt's sample of T-tailed DNA would have far exceeded the 65,536 possible unique tags resulting from a 4-mer sequence, and thus reads on a "a population of non-

uniquely tagged parent polynucleotides." Dr. Quackenbush admits that, for "a sample comprised of five haploid copies of the human nuclear genome that are fragmented into 200 base pair polynucleotides," "[t]he entire sample comprises about *75 million* polynucleotide fragments," which is much more than "[t]he 65,536 possible unique tags from a 4-mer tag sequence." Ex. 2025 ¶ 69 (emphasis added). There can be no dispute that, if a sample of five haploid copies of DNA produces about 75 million polynucleotide fragments, then a sample of 750 ng T-tailed DNA in Schmitt would represent a much greater amount of polynucleotide fragments. Pet. Reply 11–12; Pet. 66; Ex. 1104 ¶ 19 (stating that a sample of 20 ng of DNA contains approximately 6,060 haploid genome equivalents, correlating to billions of polynucleotide fragments); Ex. 1001, 40:15–19 (stating that a sample of 30 ng of DNA contains approximately 10,000 haploid genome equivalents, while a sample of 100 ng of DNA contains approximately 30,000 haploid genome equivalents).

Patent Owner argues that the DCS method is different from the claimed subject matter in the '822 patent because "[t]he method of Schmitt expressly requires flooding a sample with 'SMI tags' or 'n-mers' to ensure that every DNA fragment in the sample is uniquely labeled." PO Resp. 5; *see also, e.g., id.* at 29–30 (arguing that "where Schmitt does relate the number of barcodes to polynucleotide fragments in a sample it instructs that the barcodes should be in excess of the sample fragments" (citing Ex. 1011, 3:47–53, 6:59–62; Ex. 2025 ¶ 86)). Although we agree with Patent Owner that the disclosure of Schmitt is directed, in the main, to the use of an excess number of *unique* SMI sequences for tagging, Schmitt's "hybrid" method utilizes shorter *non-unique* n-mer SMI sequences that produce, in substance, a population of "non-uniquely tagged" DNA fragments. *See* Pet. 44–45; *see*

also Ex. 1002 ¶¶ 126–127. Patent Owner's argument that "[l]imiting n-mer tags as suggested in the petition materials would be contrary to the express instruction of Schmitt," PO Resp. 31, is unpersuasive because it ignores Schmitt's express disclosure and examples of 4-mer tags. *See* Ex. 1011, 4:30–54; *id.* at 9:9–13. Figure 4a of Schmitt, reproduced below, "illustrates an example of how a SMI sequence with n-mers of 4 nucleotides in length (4-mers) are read by Duplex Consensus Sequencing." *Id.* at 4:30–32.



> Figure 4a of Schmitt provides an example of an SMI sequence comprising an n-mer of 4 nucleotides in length (i.e., a 4-mer). Ex. 1011, 4:30–54.

And although Schmitt expressly labels the hybrid approach as providing a "*unique* molecular identifier," Ex. 1011, 9:12–13 (emphasis added), this is so only because the *combination* with the sequence information from the sheared ends with the short n-mer tag produces the "*unique* molecular identifier." The '822 patent similarly describes a "unique sequence" identifier that results from the combination of a non-unique barcode and sequence information at the ligation site:

> [A] plurality of barcodes may be used such that *barcodes are not necessarily unique to one another* in the plurality. In this example, the barcodes may be ligated to individual molecules such that the *combination of the bar code and the sequence it may be ligated to creates a unique sequence that may be individually tracked.* As described herein, detecting of non unique barcodes in combination with sequence data of beginning (start) and end (stop) portions of sequence reads may *allow assignment of a unique identity to a particular molecule*.

TEMPUS EX1002.248

Ex. 1001, 39:13–22 (emphases added); *see also* Pet. 45–46; Ex. 1002 ¶ 130. Moreover, although claim 1 recites that the parent polynucleotides are "non-uniquely tagged," the parent polynucleotides are nevertheless uniquely identifiable because their amplified copies are grouped into families, "whereby each of the families comprises sequence reads amplified from the same tagged parent polynucleotide." Ex. 1001, 62:40–42.

Finally, we disagree with Patent Owner that Schmitt's shearing process distinguishes over the method recited in claim 1 of the '822 patent. PO Resp. 33–35. Specifically, Patent Owner argues that "the hybrid method critically relies on sheared cellular DNA fragments" and that "shearing is incompatible with cfDNA samples." *Id.* at 33–34 (citing Ex. 1011, 9:9–13; Ex. 2020 ¶ 196; Ex. 2025 ¶¶ 93–95). Patent Owner's argument is unpersuasive because it treats the ordinarily skilled artisan like an automaton. "A person of ordinary skill is also a person of ordinary creativity . . . ." *KSR*, 550 U.S. at 421. And here, the record evidence supports Petitioner's contention that an ordinarily skilled artisan would have understood that cfDNA is already fragmented by nature and, thus, would have simply omitted the shearing step when tagging cfDNA with Schmitt's short 4-mer tags. Ex 1002 ¶¶ 118, 143; Ex. 1048, 16270–271 ("Library preparation was carried out according to the manufacturer's protocol with slight modifications. *Because cell-free plasma DNA was fragmented in nature, no further fragmentation by nebulization or sonication was done on plasma DNA samples.*" (emphasis added)); Ex. 1052,[13] 94 (describing plasma DNA library preparation and stating that "plasma DNA molecules

[13] Gary J.W. Liao et al., *Targeted Massively Parallel Sequencing of Maternal Plasma DNA Permits Efficient and Unbiased Detection of Fetal Alleles*, 57(1) CLIN. CHEM. 92–101 (2011) (Ex. 1052).

TEMPUS EX1002.249

[are] already fragmented in nature, so no additional fragmentation step [is] required").

Patent Owner's insinuation that an ordinarily skilled artisan would have found the shearing step to be an insurmountable obstacle to the combination of Schmitt and Fan or Forshew is therefore unpersuasive. *See ClassCo, Inc. v. Apple, Inc.*, 838 F.3d 1214, 1219 (Fed. Cir. 2016) ("The rationale of *KSR* does not support [the] theory that a person of ordinary skill can only perform combinations of a puzzle element A with a perfectly fitting puzzle element B."). Indeed, library generation, shearing (or not shearing), and ligating adaptors to DNA fragments were well known and routine steps in the art before the earliest priority date for the '822 patent. *See, e.g.*, Ex. 1011, 19:5–20:15 (Example 1); Ex. 1004, 16–34 (describing technical methods and design); Ex. 1048, 16270–271 (describing library preparation and adaptor ligation).

For all the above reasons, we agree with Petitioner and Dr. Gabriel that claim 1 does not distinguish over Schmitt's "hybrid" method. Pet. 45; Ex. 1002 ¶ 130. Although the '822 patent touts that "a sample comprising about 10,000 haploid human genome equivalents of cfDNA can be tagged with about 36 unique identifiers," Ex. 1001, 41:31–33, the tag count of claim 1 is not so limited. Indeed, dependent claim 11 recites tagging "with from 10 to 100,000 different identifiers," Ex. 1001, 63:11–13, thus suggesting that the number of non-unique tags in claim 1, from which claim 11 depends, is broader. *See AK Steel Corp. v. Sollac & Ugine*, 344 F.3d 1234, 1242 (Fed. Cir. 2003) ("Under the doctrine of claim differentiation, dependent claims are presumed to be of narrower scope than the independent claims from which they depend."). Thus, we are satisfied that Petitioner establishes by a preponderance of the evidence that the combination of

Schmitt, Schmitt 2012, and Fan or Forshew teaches or suggests each and every limitation of claim 1.

c. *Dependent Claims 3, 4, 6–8, 10, 11, and 17–20*

Having decided that the combination of Schmitt, Schmitt 2012, and Fan or Forshew teaches or suggests each and every limitation of claim 1, we turn to the remaining challenged claims of the '822 patent, which all directly depend from claim 1 (i.e., claims 1–13 and 17–20). Patent Owner presents separate arguments for dependent claims 2, 5, 9, 12, and 13. *See* PO Resp. 53–57. We address those claims individually below. *See infra* § IV.E.1.d–h.

As to dependent claims 3, 4, 6–8, 10, 11, and 17–20, we find that Petitioner also shows by a preponderance of the evidence that Schmitt, Schmitt 2012, and Fan or Forshew account for the limitations in these claims, or that certain limitations were well-known in the art as admitted in the '822 patent. Pet. 59–65, 67–70. We have also reviewed Dr. Gabriel's testimony and find that a preponderance of the evidence supports her contention that the cited references collectively disclose or suggest each and every limitation of claims 3, 4, 6–8, 10, 11, and 17–20. *See* Ex. 1002 ¶¶ 162–163 (claim 3) (citing Ex. 1011, 3:1–6, 3:44–62, 7:58–62, 15:21–38); *id.* ¶¶ 164–166 (claim 4) (citing Ex. 1011, 23:10–14; Ex. 1047, SI1); *id.* ¶¶ 168–169 (claim 6) (Ex. 1011, 4:25–29, 17:9–11, 29:57–30:10, Figure 3); *id.* ¶ 170 (claim 7) (citing Ex. 1011, 21:10–19, 23:10–14); *id.* ¶¶ 171–172 (claim 8) (citing Ex. 1004, Abstract, 3, SI1, Table S1 (as evidenced by Ex. 1081, 84, 87)); *id.* ¶¶ 174–176 (claim 10) (citing Ex. 1001, 32:16–19 (stating that "[m]ethods of reducing noise and/or distortion from a sequencing process are known" and such methods "include, for example, filtering sequences, e.g., requiring them to meet a quality threshold, or

TEMPUS EX1002.251

reducing GC bias" (evidenced by Ex. 1079, 893; Ex. 1014, 446))); *id.* ¶ 177 (claim 11) (citing Ex. 1011, 9:1–13); *id.* ¶¶ 186–188 (claim 17) (citing Ex. 1011, 20:23–29, 22:50–67); *id.* ¶¶ 189–190 (claim 18) (citing Ex. 1004, 1, 3, SI1, Table S1); *id.* ¶¶ 191–192 (claim 19) (citing Ex. 1011, 20:65–21:19; Ex. 1047, SI2; Ex. 1001, 32:16–19 (as evidenced by Ex. 1079, 893; Ex. 1014, 446)); *id.* ¶¶ 193–194 (claim 20) (citing Ex. 1011, 9:1–13).

Again, Patent Owner does not present separate and specific arguments for any of these dependent claims. *See* PO Resp. 53–57 (presenting arguments only as to dependent claims 2, 5, 9, 12, and 13). We, therefore, adopt the teachings set forth in the Petition and in Dr. Gabriel's Declaration as mapped to the limitations of the challenged claims as our own findings. *See In re NuVasive, Inc.*, 841 F.3d 966, 974 (Fed. Cir. 2016) (explaining that the Board need not make specific findings about claim limitations that a patent owner does not dispute are disclosed in the prior art).

d. *Dependent Claim 2*

Claim 2 depends from claim 1 and recites "further comprising detecting, at one or more loci, at least one single nucleotide variant, at least one gene fusion and at least one copy number variant." Ex. 1001, 62:49–51. We agree with Petitioner that Schmitt teaches or suggests detecting each of these types of genetic aberrations. Pet. 57–59.

Specifically, Schmitt teaches that the DCS method allows for the confirmation of "the presence of a true mutation (as opposed to a PCR error or other artifactual mutation) in a target nucleic acid sequence." Ex. 1011, 16:43–47. Schmitt also teaches that the DCS method has the "ability to indirectly infer that damage is present on the DNA," which is a useful biomarker for, e.g., cancer risk, cancer metabolic state, mutator phenotype related to defective damage repair, carcinogen exposure, chronic

inflammation exposure, individual-specific aging, and neurodegenerative diseases. *Id.* at 15:47–53. As an example, Schmitt teaches using DCS to detect "single base substitutions" in a series of M13mp2 variants. Ex. 1011, 29:57–30:10; *see also id.* at 5:3–6 (stating that "DCS analysis removes . . . strand bias and reveals the true mtDNA mutational spectrum to be characterized by an excess of transitions"). As another example, Schmitt teaches using DCS for the "accurate detection of altered genomic copy number (e.g., for sensitive diagnosis of genetic conditions such as trisomy 21)." Ex. 1011, 18:3–21 (footnote omitted); *see also* Ex. 1002 ¶¶ 158–159.

Patent Owner argues that Schmitt does not expressly disclose detecting genomic copy variation (as opposed to genomic copy number). PO Sur-Reply 23–24. But we agree with Petitioner that an ordinarily skilled artisan would have understood that Schmitt's reference to "true mutations" includes each type of genetic aberration recited in claim 2. Ex. 1002 ¶ 157; *see also Arendi S.A.R.L. v. Apple Inc.*, 832 F.3d 1355, 1361 (Fed. Cir. 2016) (stating that we must "consider common sense, common wisdom, and common knowledge in analyzing obviousness").

e. *Dependent Claim 5*

Claim 5 depends from claim 1 and recites "wherein making the base call comprises voting, averaging, maximum a posteriori or maximum likelihood detection, dynamic programming, Bayesian methods, hidden Markov methods or support vector machine methods." Ex. 1001, 62:59–63. As Petitioner points out, the '822 patent states that "'voting, averaging, statistical, maximum a posteriori or maximum likelihood detection, dynamic programming, Bayesian, hidden Markov or support vector machine methods' for generating consensus sequences were 'known in the art.'" Pet. 60 (quoting Ex. 1001, 45:64–46:5). Patent Owner argues that Petitioner

fails to sufficiently point to a teaching *in Schmitt* for these base-calling methods. PO Resp. 564. But this argument merely attacks the references individually and ignores what Patent Owner's own written description provides as common knowledge in the art. Thus, Patent Owner's argument is not persuasive. *See Randall Mfg. v. Rea*, 733 F.3d 1355, 1362 (Fed. Cir. 2013) ("rejecting a blinkered focus on individual documents" in an obviousness analysis).

f. *Dependent Claim 9*

Claim 9 depends from claim 1 and recites "wherein mapping the sequence reads comprises using information about a length of each of the sequence reads." Ex. 1001, 63:7–9. We agree with Petitioner that the references teach or suggest mapping sequence reads using information about the length of each sequence read. Pet. 64. Both Schmitt and Schmitt 2012 teach aligning (or mapping) sequence reads to a reference genome using the Burrows-Wheeler aligner (BWA). Ex. 1011, 20:57–58; Ex. 1047, SI1. Dr. Gabriel testifies that an ordinarily skilled artisan would have understood that the BWA uses information about a length of each of the sequence reads for mapping. Ex. 1002 ¶ 173. We find Dr. Gabriel's testimony credible as supported by preponderant record evidence. Specifically, Li and Durbin[14] provide an overview of the BWA and teach that "[d]etermining the allowed maximum number of differences" is based on "a read of length *m*." Ex. 1077, 1757. Patent Owner's argument to the contrary lacks specificity and is unpersuasive. *See* PO Resp. 55 (arguing that "[t]he petition does not

[14] Heng Li and Richard Durbin, *Fast and accurate short read alignment with Burrows–Wheeler transform*, 25(14) BIOINFORMATICS 1754–1760 (2009) (Ex. 1077, "Li and Durbin").

TEMPUS EX1002.254

explain what aspect of the BWA algorithm that allegedly provides this teaching").

g. *Dependent Claim 12*

Claim 12 depends from claim 1 and recites "wherein the population of polynucleotides is tagged with n different unique identifiers, wherein n is no more than 100*z, wherein z is a mean of an expected number of duplicate molecules having the same start and stop positions in the sample." Ex. 1001, 63:15–19. Petitioner contends that Fan and Forshew disclose cfDNA amounts between 1 and 20 ng. Pet. 66 (citing Ex. 1048, 16270, Table S1; Ex. 1004, Table S6). Relying on Dr. Gabriel's testimony, Petitioner contends that "[a] sample of 20 ng DNA contains about 6,060 haploid genome equivalents, which would have approximately 1–2 duplicate polynucleotides beginning at any given position." *Id.* (citing Ex. 1002 ¶ 179). Petitioner contends that, "when there are 6,060 haploid genome equivalents, $z \approx 1–2$ and $100*z \approx 10–200$," and thus "claim 12 requires tagging with no more than 100–200 different unique identifiers." *Id.* (citing Ex. 1002 ¶ 179). Petitioner contends that Schmitt teaches "the use of a small number of different tags," such as 1-mer tags, and "when considering 1-mer tags at both ends of a fragment, the number of different combinations is 16." *Id.* at 67 (citing Ex. 1011, 9:1–13; Ex. 1002 ¶ 180, n.12). "Accordingly," Petitioner contends, "Schmitt teaches the additional limitation of claim 12." *Id.* (citing Ex. 1002 ¶ 181).

Patent Owner argues that claim 12 is not unpatentable because "[n]one of the cited references disclose[s] the expected number of duplicate molecules in a sample much less a tagging scheme based on such a number." PO Resp. 55. Patent Owner argues that Petitioner's analysis of claim 12 is based on hindsight because Petitioner relies on the '822 patent's disclosure

of the empirical relationship between "z" and the number of tags, and that Petitioner fails to provide any explanation "for selecting a 20 ng sample size from the 1–20 ng range." *Id.* at 55–56 (citing Pet. 66).

In its Reply, Petitioner contends that "[c]laim 12 is directed to a number of identifiers for non-unique tagging, *i.e.*, no more than 100*z," and that the combination of prior-art references teaches this number. Pet. Reply 16.

Upon review of the parties' respective arguments and evidence, we determine that Patent Owner has the better position. As Patent Owner points out and Petitioner does not dispute, the relationship between "z" and the number of tags "n" in a sample was unknown prior to the '822 patent. PO Resp. 55–56; *see also* Ex. 1001, 41:6–13 (expressly defining "z" as the "measure of central tendency (e.g., mean, median, mode) of an expected number of duplicate molecules having the same start and stop positions"). This is because the relationship was determined empirically and disclosed in the '822 patent. *See* Ex. 1001, 40:24–27 ("It has been empirically determined that in a sample of about 10,000 haploid genome equivalents of human DNA, there are about 3 duplicate polynucleotides beginning at any given position."). Petitioner has not explained adequately, nor provided preponderant evidence, that the relationship between "100*z" and "n," as recited in claim 12, was known in the art. *See In re Omeprazole Patent Litigation*, 536 F.3d 1361, 1379–80 (Fed. Cir. 2008) (holding that even where a general method that could have been applied to make the claimed product was known and within the level of skill of the ordinary artisan, the claim may nevertheless be nonobvious if the problem which had suggested use of the method had been previously unknown). For example, Petitioner

does not rely on an inherency theory for a teaching or suggestion of the claimed subject matter in the prior art. Pet. 66–67; Pet. Reply 16.

We find that Petitioner's arguments and evidence as to claim 12 are conclusory and not persuasive. *See Magnum Oil*, 829 F.3d at 1380 (stating that a petitioner cannot satisfy its burden of proving obviousness by employing "mere conclusory statements"). In particular, Petitioner fails to specifically address the previously unknown relationship between "n" and "z" recited in the "wherein" clause of claim 12. Thus, we determine that Petitioner fails to show by a preponderance of the evidence that the prior art teaches or suggests the subject matter of claim 12.

h. *Dependent Claim 13*

Claim 13 depends from claim 1 and recites "wherein no more than 100 nanograms of polynucleotides from the bodily sample are converted in b)." Ex. 1001, 63:20–22. We agree with Petitioner that both Fan and Forshew explicitly disclose providing no more than 100 ng of cfDNA molecules from the plasma of human subjects for sequencing. Pet. 67; *see also* Ex. 1002 ¶¶ 182–185. Specifically, Forshew extracted between 0.9 ng and 19.7 ng of DNA from cell-free plasma samples for sequence library preparation. Ex. 1004, Table S6. Similarly, Fan extracted between 1.2 ng and 8 ng of DNA from cell-free plasma for sequencing library construction. Ex. 1048, 16270, Table S1. These amounts fall within "no more than 100 nanograms of polynucleotides from the bodily sample" as recited in claim 13. Patent Owner argues that claim 13 "is not unpatentable" because "Schmitt discloses using 3 μg of human colonic mucosa DNA." PO Resp. 56–57. Again, this argument is unpersuasive because it attacks the references individually. *See In re Merck & Co.*, 800 F.2d 1091, 1097 (Fed. Cir. 1986) (attacking references individually is improper).

TEMPUS EX1002.257

2. Motivation to Combine/Reasonable Expectation of Success

Even "[i]f all elements of the claims are found in a combination of prior art references," "the factfinder should further consider whether a person of ordinary skill in the art would [have been] motivated to combine those references, and whether in making that combination, a person of ordinary skill would have [had] a reasonable expectation of success." *Merck & Cie v. Gnosis S.P.A.*, 808 F.3d 829, 833 (Fed. Cir. 2015). The "motivation to combine" and "reasonable expectation of success" factors are subsidiary requirements for obviousness subsumed within the *Graham* factors. *Pfizer, Inc. v. Apotex, Inc.*, 480 F.3d 1348, 1361 (Fed. Cir. 2007). We address motivation to combine and reasonable expectation of success in turn below.

a. *Motivation to Combine*

Petitioner contends that an ordinarily skilled artisan, at the time of the '822 patent, would have been prompted to combine "the teachings of Schmitt and Schmitt 2012 regarding the DCS method" with "the teachings of either Fan or Forshew regarding sequencing and analysis of cell free DNA," with a reasonable expectation of success, based on the teachings of the art, the knowledge of an ordinarily skilled artisan, and the teachings of the Schmitt references. Pet. 32–41. In our Institution Decision, we determined that Petitioner had shown sufficiently for institution that an ordinarily skilled artisan would have been motivated to use Schmitt's DCS method to detect low-frequency genetic mutations in the cfDNA samples disclosed in Fan or Forshew. Inst. Dec. 24–25. Upon review of the complete record and the parties' respective arguments and evidence, we affirm our initial determination.

Specifically, we find that a preponderance of record evidence supports Petitioner's argument that Schmitt "provides explicit motivation to combine" its teachings with those of Fan, because Schmitt references Fan when discussing the state of the prior art and limitations thereof. Pet. 32 (citing Ex. 1011, 1:28–55; Ex. 1002 ¶ 99). When discussing the prior art, Schmitt states that "deep sequencing" technologies have "ushered in a new era of genomic exploration," because they "offer the unique ability to detect minor variants within heterogeneous mixtures." Ex. 1011, 1:28–37. Schmitt states that those technologies have been implemented in a variety of fields and that "[c]linical applications, such [as] prenatal screening for fetal aneuploidy . . . are rapidly being developed." Ex. 1011, 1:38–45. As an example of prenatal screening, Schmitt directly cites to Fan. *Id.* at 1:42–43 (citing reference "[9]"). Schmitt then discusses the limitations of the prior-art deep-sequencing techniques, such as limits on sensitivity and accuracy. *See id.* at 1:56–2:55 ("Although, in theory, DNA subpopulations of any size should be detectable when deep sequencing a sufficient number of molecules, a practical limit of detection is imposed by errors introduced during sample preparation and sequencing."). Schmitt states that the method disclosed in the '822 patent is "an approach for tag-based error correction, which reduces or eliminates artifactual mutations arising from DNA damage, PCR errors, and sequencing errors." *Id.* at 2:56–59. Schmitt states that the disclosed method is an improvement over the prior art because it "allows rare variants in heterogeneous populations to be detected with unprecedented sensitivity." *Id.* at 2:56–62.

We agree with Petitioner that an ordinarily skilled artisan would have had a reason to combine the disclosures of the Schmitt references with Fan: that is, to use Schmitt's DCS method to accurately detect prenatal genetic

mutations in cfDNA as taught by Fan. Pet. 36–37. Schmitt specifically refers to "prenatal screening for fetal aneuploidy" and references Fan in discussing the prior art and limitations thereof. Ex. 1011, 1:42–43; *see also* Pet. 36. An ordinarily skilled artisan screening for fetal aneuploidy, as described by Fan, Ex. 1048, Abstract, 16266, would have been prompted to look to Schmitt's improved DCS method to "reduce[] or eliminate[] artifactual mutations arising from DNA damage, PCR errors, and sequencing errors." Ex. 1011, at 2:56–59; *see also* Pet. 36–37; Ex. 1002 ¶¶ 103–108.

We further agree with Petitioner that an ordinarily skilled artisan screening cfDNA for cancer mutations, as described by Forshew, Ex. 1004, Abstract, 1–2, would have also looked to Schmitt's DCS method for the same reasons. Pet. 37; *see also* Ex. 1011, Abstract, 1:28–5; Ex. 1002 ¶¶ 109–110. Specifically, in addition to discussing the use of deep-sequencing techniques for prenatal screening for fetal aneuploidy, Schmitt also refers to "early detection of cancer." Ex. 1011, 1:42–45. Thus, we agree with Petitioner that an ordinarily skilled artisan screening for cancer mutations in cell-free plasma DNA, as described by Forshew, Ex. 1004, Abstract, would have been prompted to look to Schmitt's DCS method to reduce amplification and sequencing errors and to improve detection of rare genetic mutations. Ex. 1004, 1 (stating that "[s]ensitive methods for detecting cancer mutations in plasma may find use in early detection screening, prognosis, monitoring tumor dynamics over time, or detection of minimal residual disease"); *see also* Pet. 36–37; Ex. 1002 ¶ 101.

We also find that the record supports, by a preponderance of the evidence, Petitioner's contentions about the knowledge the "ordinarily skilled artisan would have brought to bear when considering combinations or modifications" of the prior-art references. *Randall Mfg. v. Rea*, 733 F.3d

1355, 1362 (Fed. Cir. 2013); *see* Pet. 32–35. Specifically, the record evidence shows that the ordinarily skilled artisan would have understood that genetic mutations in cfDNA occur at low frequencies compared to wild-type DNA and, thus, represent a small portion of the heterogeneous mixture of DNA in the plasma. Ex. 1002 ¶ 100; Ex. 1051, Abstract (discussing the low frequency of mutant adenomatous polyposis coli (APC) DNA molecules in the plasma of cancer patients compared to normal APC); Ex. 1052, 92 (stating that "the coexistence in maternal plasma of a minor population of fetal DNA within a major background of maternal DNA" poses challenges for non-invasive prenatal diagnosis of fetal genetic and chromosomal diseases); *see also* Pet. 32–33. Indeed, Fan expressly states that "measuring aneuploidy remains challenging because of the high background of maternal DNA; fetal DNA often constitutes <10% of total DNA in maternal cell-free plasma." Ex. 1048, 16266.

We agree with Petitioner that an ordinarily skilled artisan, wishing to detect cfDNA genetic mutations as discussed in Fan or Forshew with high accuracy and sensitivity, would have been motivated to use Schmitt's improved DCS method, because the DCS method "reduces or eliminates artifactual mutations arising from DNA damage, PCR errors, and sequencing errors," and "allows rare variants in heterogeneous populations to be detected with unprecedented sensitivity." Ex. 1011, 2:56–60; *see also KSR*, 550 at 417 ("[I]f a technique has been used to improve one device, and a person of ordinary skill in the art would recognize that it would improve similar devices in the same way, using the technique is obvious unless its actual application is beyond his or her skill.").

We have carefully considered Patent Owner's arguments and evidence to the contrary, but remain persuaded that Petitioner has provided a

TEMPUS EX1002.261

sufficient reason with rational underpinning for combining Schmitt with Fan or Forshew. *See* PO Resp. 35–38. Citing to the testimonies of Dr. Shendure and Dr. Quackenbush, Patent Owner argues that an ordinarily skilled artisan would not have understood "the background section [of Schmitt] to be contemplating application of DCS to fetal aneuploidy detection." PO Resp. 35 (citing Ex. 2025 ¶¶ 102–110; Ex. 2023 ¶ 20). We disagree. As explained above, Schmitt states that deep-sequencing techniques have been implemented in a variety of fields, including prenatal screening for fetal aneuploidy and for early detection and monitoring of cancer. Ex. 1011, 1:28–55. Schmitt then discusses the limitations of the prior-art techniques, such as limits on sensitivity and accuracy, and states that the method disclosed is an improvement over the prior art because it "reduces or eliminates artifactual mutations arising from DNA damage, PCR errors, and sequencing errors," and "allows rare variants in heterogeneous populations to be detected with unprecedented sensitivity." *Id.* at 1:56–2:62. Taken together, we find that an ordinarily skilled artisan would have viewed these teachings as an express suggestion to use the DCS method for the prior-art applications described (e.g., prenatal and cancer-mutation screening). Ex. 1104 ¶¶ 33–34.[15]

[15] We are not persuaded otherwise by the testimonies of Dr. Shendure and Dr. Quackenbush. Dr. Quackenbush states that "Schmitt is merely disclosing that next-generation sequencing ('NGS') has been used by others to detect fetal aneuploidy." Ex. 2025 ¶ 103. Similarly, Dr. Shendure states that "Schmitt is merely mentioning in some introductory remarks that different techniques employing DNA sequencing have been used by others to detect fetal aneuploidy." Ex. 2023 ¶ 20. These statements fail to read the entire background section of Schmitt as a whole, and thus lack specificity and credibility. *See Rohm & Haas Co. v. Brotech Corp.*, 127 F.3d 1089,

TEMPUS EX1002.262

Relying on Dr. Quackenbush's testimony, Patent Owner also argues that "there would be no good reason to apply Schmitt's DCS method to detecting fetal aneuploidy," because fetal aneuploidy is not a "rare mutation" and using DCS to detect mutations that are not rare "increase[s] the expense and complexity of the method." PO Resp. 35–37 (citing Ex. 2025 ¶¶ 106–107); PO Sur-Reply 10–11. As an initial matter, we note that, as of the earliest priority date of the '822 patent, ordinarily skilled artisans knew that barcode-based deep-sequencing techniques had been used to screen cfDNA for genetic defects such as aneuploidy successfully. Ex. 1002 ¶ 113; Ex. 1058, 72–74 (describing the use of unique molecular identifiers to detect trisomy 21); Ex. 1059, 1 (describing the use of barcode-based massively parallel sequencing for the detection of aneuploidy). But even if Patent Owner is correct that fetal aneuploidy does not qualify as a "rare mutation," the lack of economic feasibility due to increased complexity is not a teaching away from using Schmitt's DCS method for that purpose. *See In re Farrenkopf*, 713 F.2d 714, 718 (Fed. Cir. 1983) ("That a given combination would not be made by business[wo]men for economic reasons does not mean that persons skilled in the art would not make the combination because of some technological incompatibility. Only the latter fact would be relevant.").

Finally, Patent Owner argues that Schmitt, Fan, and Forshew are, at most, analogous art and that their use of next-generation sequencing "is not evidence that [an ordinarily skilled artisan] would have wanted to—or could have—combined these methods." PO Resp. 37–38 (citing Ex. 2025 ¶¶ 108–

1092 (Fed. Cir. 1997) ("Nothing in the rules or in our jurisprudence requires the fact finder to credit the unsupported assertions of an expert witness.").

TEMPUS EX1002.263

110). We note that analogous art can, in some instances, demonstrate a reason to combine references, such as where "[o]ne skilled in the art would naturally look to prior art addressing the same problem as the invention at hand, and . . . would find an appropriate solution." *In re ICON Health & Fitness, Inc.*, 496 F.3d 1374, 1380 (Fed. Cir. 2007). Here, Schmitt teaches that deep-sequencing techniques have been used "in a variety of fields," including in "prenatal screening for fetal aneuploidy" and in "early detection of cancer." Ex. 1011, 1:38–45. Schmitt teaches, however, that deep sequencing "has limitations" in that it produces artifactual mutations resulting from DNA damage, PCR errors, and sequencing errors. *Id.* at 1:56–2:62. Schmitt teaches that the disclosed DCS method (which includes the "hybrid" embodiment discussed above) "reduces or eliminates artifactual mutations" associated with the prior art. *Id.* An ordinarily skilled artisan looking to screen cfDNA for prenatal or cancerous mutations with high accuracy and sensitivity would have naturally looked to improvements in the field of next-generation sequencing, such as Schmitt's DCS method. Ex. 1001 ¶¶ 99–111; Ex. 1104 ¶¶ 32–39. Thus, although the fact that the prior-art references are analogous is not dispositive on the issue of motivation to combine, we maintain that an ordinarily skilled artisan reading Schmitt would have understood that Schmitt's DCS method was a reasonable substitute for the deep-sequencing and error-correction techniques used in the prior art—such as the next-generation sequencing techniques Fan and Forshew used for prenatal cfDNA screening and for tumor cfDNA screening, respectively. *Id.*

In sum, we find that the facts here constitute a case of "the simple substitution of one known element for another or the mere application of a known technique to a piece of prior art ready for the improvement." *KSR*,

TEMPUS EX1002.264

550 U.S. at 417. Thus, for all the above reasons, we find that a preponderance of the evidence supports Petitioner's contention that an ordinary artisan would have been motivated to combine the teachings of the Schmitt references with those of Fan or Forshew.

b. *Reasonable Expectation of Success*

We next consider whether Petitioner has shown by a preponderance of the evidence that an ordinarily skilled artisan would have had a reasonable expectation of success. "The reasonable expectation of success requirement refers to the likelihood of success in combining references to meet the limitations of the claimed invention." *Intelligent Bio-Sys., Inc. v. Illumina Cambridge Ltd.*, 821 F.3d 1359, 1367 (Fed. Cir. 2016).

In this case, the question before us is whether the ordinarily skilled artisan would have had a reasonable expectation that Schmitt's DCS method could be used successfully with a population of cfDNA molecules. Ex. 1001, 62:18–48 (claim 1). In making our findings as to "reasonable expectation of success," we keep in mind that we cannot demand absolute certainty. *See Medichem, S.A. v. Rolabo, S.L.*, 437 F.3d 1157, 1165 (Fed. Cir. 2006) ("While the definition of 'reasonable expectation' is somewhat vague, our case law makes clear that it does not require a certainty of success."); *see also Pfizer*, 480 F.3d at 1364 ("[C]ase law is clear that obviousness cannot be avoided simply by a showing of some degree of unpredictability in the art so long as there was a reasonable probability of success.").

Upon consideration of the entire record, we are persuaded by and adopt Petitioner's contentions (and Dr. Gabriel's testimony) that an ordinarily skilled artisan would have had a reasonable expectation of success in using Schmitt's DCS method to detect genetic mutations in cfDNA. *See*

TEMPUS EX1002.265

Pet. 37–41; Ex. 1002 ¶¶ 112–120. In particular, we agree with Dr. Gabriel's testimony that all the necessary techniques for carrying out the claimed method steps of the '822 patent were known and required only ordinary skill to perform. *See* Ex. 1002 ¶¶ 116–117. In this regard, the '822 patent itself teaches that the techniques needed to implement the claimed method were well known and routine in the art. *See* Ex. 1001, 36:1–6 (stating that "cell free polynucleotides may be isolated and extracted using a variety of techniques known in the art" including by "commercially available kits"); *id.* at 38:20–25 (stating that "assignment of unique or non-unique identifiers, or molecular barcodes in reactions of this disclosure may follow methods and systems described by, for example," prior-art patents and patent applications); *id.* at 38:41–48 (stating that "PCR for sequencing may be performed using any means, including but not limited to use of commercial kits provided by Nugen (WGA kit), Life Technologies, Affymetrix, Promega, Qiagen and the like"); *id.* at 30:8–12 (stating that "subsequent sequencing of cell free polynucleotides" may be done "by techniques known in the art"); *id.* at 35:27–31 (as to mapping and aligning sequences, stating that "sequences can be interrogated using the genome brow[s]er available" online); *id.* at 45:64–46:5 (stating that "[c]onsensus sequences can be generated from families of sequence reads by any method known in the art").

We also agree with Dr. Gabriel's testimony that Schmitt itself reasonably suggests that the DCS method can be used successfully for several different types of applications by

> explain[ing] that the DCS method does not require substantive modification of existing sequencing workflows and can serve as a general technique for next generation sequencing:

TEMPUS EX1002.266

> [T]he DCS approach can be generalized to nearly any sequencing platform because a double-stranded SMI tag can be incorporated into other existing adaptors, or for sequencing approaches that do not require adaptors, a double-stranded SMI tag can be ligated onto duplex DNA sample prior to sequencing [… and therefore the] compatibility of DCS with existing sequencing workflows, the potential for greatly reducing the error rate of DNA sequencing, and the multitude of applications for the double-stranded SMI sequences validate DCS as a technique that may play a general role in next generation DNA sequencing.

Ex. 1002 ¶ 114 (quoting Ex. 1011, 18:44–61) (alteration in original).

In sum, "[t]his is not a situation where the prior art gave no direction on how to reach a successful result." *In Re: Copaxone Consol. Cases*, 906 F.3d 1013, 1026 (Fed. Cir. 2018). Instead, Schmitt suggests to the skilled artisan to use the DCS method as a general technique for next generation sequencing and error correction, and the '822 patent evinces that the technical details for isolating, barcoding, amplifying, sequencing, and analyzing cfDNA molecules were known to be routine in the art. These factors persuade us that an ordinarily skilled artisan would have had a reasonable expectation of success.

We are not persuaded by Patent Owner's arguments and evidence to the contrary. *See* PO Resp. 38–53. We observe that many of Patent Owner's arguments are, in essence, that the prior art fails to provide proof that using the DCS method to analyze cfDNA would have been successful. *See, e.g.*, PO Resp. 38–41 (characterizing Petitioner's evidence of the successful prior-art use of barcode-based sequencing and error correction for cfDNA irrelevant because none "use[d] the Schmitt method," and arguing that "[e]ven if other methods using [next-generation sequencing] had been

applied to cfDNA, it does not establish that the different Schmitt method is applicable to cfDNA"); *id.* at 41–42 (arguing that "Schmitt does not tag, much less analyze fragments, similar to the naturally fragmented size of cfDNA"); *id.* at 45 (arguing that "Schmitt applies the DCS method exclusively to samples of cellular DNA" and "does not even mention applying DCS to cfDNA, let alone provide any corresponding analysis").

But the test for obviousness does not require "actual success" in the prior art. *See Endo Pharm. Inc. v. Actavis LLC,* 922 F.3d 1365, 1379 (Fed. Cir. 2019) ("As we have repeatedly held, obviousness requires only a reasonable expectation of success, not proof of actual success."). "[O]nly a reasonable expectation of success, not a guarantee, is needed." *Pfizer*, 480 F.3d 1364; *see also Par Pharm., Inc. v. TWi Pharms., Inc.*, 773 F.3d 1186, 1198 (Fed. Cir. 2014) ("The reasonable expectation of success requirement for obviousness does not necessitate an absolute certainty for success."). And here, we find that a preponderance of the evidence supports Petitioner's contention that an ordinarily skilled artisan would have had a *reasonable*, even if not absolute, expectation that the DCS method could have been used to analyze cfDNA successfully for the reasons discussed above.

Patent Owner also argues an ordinarily skilled artisan would not have had a reasonable expectation of success in substituting genomic DNA with cfDNA because the skilled artisan would have considered Schmitt's DCS method to be incompatible with cfDNA. *See* PO Resp. 41–53. In this regard, Patent Owner argues that "there are several critical aspects of the Schmitt method (e.g., random shearing, optimal fragment size, randomly generated tags, numerous processing steps) that would be recognized as being incompatible with cfDNA sample analysis." *Id.* at 46 (citing Ex. 2023 ¶¶ 21–22). We address these alleged incompatibilities individually below.

First, Patent Owner argues that DCS is incompatible with cfDNA because "Schmitt describes the optimal range of fragment sizes for DCS is 200–500 bp"—a size that is greater than that of naturally occurring cfDNA (i.e., about 140–170 bp). *Id.* at 41–42, 46–47. This argument is not persuasive because nothing in Schmitt limits the applicability of DCS to DNA fragments of approximately 200–500 bp in size. Instead, Schmitt clearly teaches as part of the DCS method "size-selecting for *appropriate length* fragments." Ex. 1011, 3:53–56 (emphasis added). That the appropriate length for fragments analyzed in Example 1 of Schmitt was 200–500 bp does not teach an ordinarily skilled artisan away from the use of other, differently sized, fragments for different needs. *See* Ex. 1011, 19:57–20:15 (describing sequencing library preparation for Example 1); *see also* Ex. 1104 ¶¶ 48–50.

Second, Patent Owner argues that Schmitt's hybrid DCS method is incompatible with cfDNA because that method relies on a random shearing step to label the DNA. PO Resp. 42 (citing Ex. 1011, 9:9–13, 25:64–65, 17:41–45). Patent Owner also argues that Dr. Gabriel conceded that shearing is incompatible with cfDNA. *Id.* at 50 (citing Ex. 2021, 158:4–12; Ex. 2022, 105:11–20; Ex. 2023 ¶¶ 33–34). As discussed above, however, Patent Owner's arguments about shearing are not persuasive. *Supra* § IV.E.1.b. The record evidence amply supports Petitioner's contention that an ordinarily skilled artisan would have understood that cfDNA is already fragmented by nature and, thus, would have simply omitted the shearing step when tagging cfDNA with Schmitt's short 4-mer tags. Ex 1002 ¶¶ 118, 143; Ex. 1052, 94 (describing plasma DNA library preparation and stating that "plasma DNA molecules [are] already fragmented in nature, so no additional fragmentation step [is] required"); Ex. 1104 ¶ 46. Again, "[a] person of

ordinary skill is also a person of ordinary creativity, not an automaton." *KSR*, 550 U.S. at 421.[16]

Third, Patent Owner argues that Schmitt's DCS method "requires large quantities of input DNA that are simply not available when working with cfDNA due to minute quantities that are reasonably obtained from patients." PO Resp. 46, 49 (citing Ex. 2023 ¶ 32). These large amounts of DNA are necessary, Patent Owner argues, "[i]n view of inefficiencies in the method and steps requiring discarding significant amounts of sample and data." *Id.* at 49. In particular, Patent Owner argues that "Schmitt instructs additional steps that discard sample and data during processing and analysis." PO Resp. 47 (citing Ex. 1011, 20:42–43, 20:47–50, 20:53–55, 20:65, 21:4–18); *see also id.* at 47–49 (discussing SMI tags as "an additional source of data loss). Patent Owner argues that "[t]he difference in the amount of DNA between what is used in the Schmitt method and what is practicable to obtain from a patient is not trivial" and "would require fundamental modification and redesign of the DCS method." *Id.* at 49–50.

Although we have carefully considered Patent Owner's arguments and evidence on this topic, we remain persuaded that an ordinarily skilled artisan

---

[16] Moreover, although a subtle point, we disagree with Patent Owner that an ordinarily skilled artisan would have understood *the process of shearing* to be a critical aspect of the hybrid method. It is true that the hybrid method uses a combination of sheared ends and a shorter n-mer tag to create a "unique" molecular identifier. Ex. 1011, 9:9–13. But the lesson the skilled artisan would have drawn from Schmitt is that the *sequences* at the ends of the DNA molecule plus the non-unique *sequences* of the 4-mer create a unique identifier. *Id.*; *see also* Ex. 1104 ¶ 47. Again, ligation of adaptors and tags to DNA sequences (whether the DNA is blunt-ended or contains overhanging molecules) was a routine, conventional activity in the art as of the earliest priority date of the '822 patent. *See, e.g.*, Ex. 1011, 19:5–20:15; Ex. 1004, 16–34; Ex. 1048, 16270–271.

TEMPUS EX1002.270

would have had a reasonable expectation of success. In this regard, the "sample" or "data" to which Patent Owner refers relates to digital information, not physical molecules of DNA. Ex. 1002 ¶¶ 174–176, 191–192 (describing the process of filtering out sequence reads that do not meet certain quality thresholds). Dr. Shendure conceded during cross-examination that "the actual manipulations are related to the removal of data rather than molecules." Ex. 1102, 75:22–79:17. In addition, as discussed above, Schmitt teaches that "the DCS approach can be generalized to nearly any sequencing platform" and emphasizes the "compatibility of DCS with existing sequencing workflows." Ex. 1011, 18:44–61. Schmitt states that the compatible sequencing platforms are, for example, "the Illumina sequencing platform, ABI SOLiD sequencing platform, Pacific Biosciences sequencing platform, 454 Life Sciences sequencing platform, Ion Torrent sequencing platform, Helicos sequencing platform, and nanopore sequencing technology." *Id.* at 15:32–38. Petitioner points out that Fan used approximately 1–8 ng of cfDNA in the Solexa/Illumina sequencing platform with success, Pet. Reply 24 (citing Ex. 1048, 16270–71), thus suggesting that any necessary modifications of these well-known sequencing platforms would have been a routine activity for an ordinarily skilled artisan.

Finally, Patent Owner argues that several articles published after Schmitt, including one written by the DCS inventors, show that the DCS method "is inefficient and poorly suited for cfDNA samples." PO Resp. 43–45 (citing Ex. 2002; Ex. 2039; Ex. 2040; Ex. 2041; Ex. 1057). But as we pointed out in the Institution Decision, Inst. Dec. 26, a certain tagging efficacy is not recited in claim 1 of the '822 patent, *see* Ex. 1001, 62:17–48

TEMPUS EX1002.271

(claim 1).[17] Patent Owner's argument about efficacy, therefore, is not persuasive. *See Sound View Innovations, LLC v. Hulu, LLC*, 2020 WL 3583556, at *3 (Fed. Cir. July 2, 2020) (holding that "[s]ince the claims here did not require a certain level of practicality, the Board did not err in finding a reasonable expectation of success in meeting the limitations of the claimed invention" (quotation marks and alternations omitted)).

3. Objective Indicia of Non-Obviousness

We must consider any evidence of objective indicia of non-obviousness before reaching our conclusion on obviousness *vel non*. *WBIP, LLC v. Kohler Co.*, 829 F.3d 1317, 1328 (Fed. Cir. 2016). Notwithstanding what the teachings of the prior art would have suggested to one of ordinary skill in the art at the time of the invention, the totality of the evidence submitted, including objective evidence of non-obviousness, may lead to a conclusion that the challenged claims would not have been obvious to one of ordinary skill. *In re Piasecki*, 745 F.2d 1468, 1471–72 (Fed. Cir. 1984). Patent Owner presents evidence of two of these considerations: (1) long-felt need and (2) commercial success. PO Resp. 57–62; PO Sur-Reply 20–23.

a. *Guardant360*

Patent Owner offers the Guardant360 product as evidence for long-felt need and commercial success of the claimed invention. PO Resp. 57. According to Patent Owner, "Guardant360 is a 'liquid biopsy' assay that provides clinically actionable sequence information from cfDNA molecules

[17] In contrast, claim 1 of U.S. Patent 9,902,992 B2, which was at issue in related proceedings IPR2019-00636 and IPR2019-00637, expressly recites "tag[ging] at least 20% of the cfDNA molecules." Institution in those proceedings was denied. *See* IPR2019-00636, Paper 10; IPR2019-00637, Paper 10.

obtained from a routine blood draw from cancer patients." *Id.* at 58 (citing Ex. 2034, 4). Patent Owner argues that Guardant360 embodies "at least claim 1 of the '822 patent," including "convert[ing] cfDNA obtained from a patient into non-uniquely tagged parent polynucleotides which are then amplified and sequenced," *id.* (citing Ex. 2029, 5, 17–18; Ex. 2030, 3539, 3540, 3542, 3544); mapping the resulting reads to a reference genome and grouping into families "based on start and stop position and identifier sequence," *id.* (citing Ex. 2029, 19; Ex. 2030, 3540–3541); and collapsing sequence reads at each genetic locus "to generate a base call" and determining "the frequency of one or more bases at each locus," *id.* (citing Ex. 2029, 19; Ex. 2030, 3544).

b. *Nexus*

At the outset, we determine that we can give Patent Owner's arguments about long-felt need and commercial success no weight in our obviousness analysis. "For objective evidence of secondary considerations to be accorded substantial weight, its proponent must establish a nexus between the evidence and the merits of the claimed invention." *In re Huai-Hung Kao*, 639 F.3d 1057, 1068 (Fed. Cir. 2011) (quotation marks and emphasis omitted); *see also Fox Factory, Inc. v. SRAM, LLC*, 944 F.3d 1366, 1373 (Fed. Cir. 2019) ("In order to accord substantial weight to secondary considerations in an obviousness analysis, the evidence of secondary considerations must have a nexus to the claims, i.e., there must be a legally and factually sufficient connection between the evidence and the patented invention." (quotation omitted)).

Patent Owner argues that we must presume a nexus. PO Resp. 57–58. But a nexus is presumed only when "the patentee shows that the asserted objective evidence is tied to a specific product and that product 'embodies

the claimed features, and is coextensive with them.'" *Fox Factory*, 994 F.3d at 1373 (quoting *Polaris Indus., Inc. v. Arctic Cat, Inc.*, 882 F.3d 1056, 1072 (Fed. Cir. 2018)); *see also WBIP*, 829 F.3d at 1329 (setting forth circumstances in which the presumption of nexus applies). And here, Patent Owner has failed to show with specific and credible evidence that claim 1 and Guardant360 are "coextensive."

For example, Patent Owner states that "Dr. Quackenbush explains [that] the Guardant360 test embodies at least claim 1 of the '822 patent," but provides no actual citations to, or discussion of, Dr. Quackenbush's testimony. PO Resp. 58. We decline to search through Dr. Quackenbush's Declarations for this evidence. *See DeSilva v. DiLeonardi*, 181 F.3d 865, 866–67 (7th Cir. 1999) ("A brief must make all arguments accessible to the judges, rather than ask them to play archeologist with the record."). Patent Owner provides citations to Exhibit 2029[18] and Exhibit 2030[19], but these documents appear to be research articles and do not provide any legal or factual analysis of claim 1 in relationship to Guardant360. Thus, we determine that Patent Owner's arguments fall well short of a persuasive showing of nexus.

---

[18] Richard B. Lanman et al., *Analytical and Clinical Validation of a Digital Sequencing Panel for Quantitative, Highly Accurate Evaluation of Cell-Free Circulating Tumor DNA*, PLOS ONE. 1–27 (Oct. 16, 2015) (Ex. 2029).

[19] Justin I. Odegaard et al., *Validation of a Plasma-Based Comprehensive Cancer Genotyping Assay Utilizing Orthogonal Tissue- and Plasma-Based Methodologies*, 24(15) CLIN. CAN. RES. 1–27 (Aug. 1, 2018) (Ex. 2030).

TEMPUS EX1002.274

4. Conclusion as to Obviousness

In sum, we find that the combination of Schmitt, Schmitt 2012, and Fan or Forshew teaches or suggests each and every element of claims 1–11, 13, and 17–20. We find that an ordinarily skilled artisan would have been motivated to combine Schmitt, Schmitt 2012, and Fan or Forshew and would have had a reasonable expectation of success in achieving the claimed invention. We also find that Patent Owner has failed to persuasively show secondary considerations of non-obviousness for failure to show persuasively that the Guardant360 product is coextensive with claim 1. We are therefore unable to accord Petitioner's evidence any weight. *Fox Factory*, 944 F.3d at 1373.

Thus, after carefully considering the arguments and evidence, we determine that the record as a whole weighs in favor of a conclusion of obviousness of claims 1–11, 13, and 17–20, especially given the disclosures of the art of record in this case and strength of the obviousness case based on the first three *Graham* factors. We find that Petitioner fails to prove by a preponderance of the evidence that claim 12 would have been obvious.

TEMPUS EX1002.275

V. CONCLUSION[20]

Petitioner establishes by a preponderance of the evidence that claims 1–11, 13, and 17–20 of the '822 patent are unpatentable. Petitioner fails to establish by a preponderance of the evidence that claim 12 of the '822 patent is unpatentable.

Claims	35 U.S.C. §	Reference(s)/Basis	Claims Shown Unpatentable	Claims Not Shown Unpatentable
1–13, 17–20	103(a)	Schmitt, Schmitt 2012, and Fan or Forshew	1–11, 13, 17–20	12
Overall Outcome			1–11, 13, 17–20	12

VI. ORDER

In consideration of the foregoing, it is hereby:

ORDERED that claims 1–11, 13, and 17–20 of the '822 patent have been proven to be unpatentable;

ORDERED that claim 12 of the '822 patent has not been proven to be unpatentable; and

FURTHER ORDERED that because this is a Final Written Decision, parties to the proceeding seeking judicial review of the Decision must comply with the notice and service requirements of 37 C.F.R. § 90.2.

---

[20] Should Patent Owner wish to pursue amendment of the challenged claims in a reissue or reexamination proceeding subsequent to the issuance of this decision, we draw Patent Owner's attention to the April 2019 *Notice Regarding Options for Amendments by Patent Owner Through Reissue or Reexamination During a Pending AIA Trial Proceeding*, 84 Fed. Reg. 16,654 (Apr. 22, 2019). If Patent Owner chooses to file a reissue application or a request for reexamination of the challenged patent, we remind Patent Owner of its continuing obligation to notify the Board of any such related matters in updated mandatory notices. *See* 37 C.F.R. § 42.8(a)(3), (b)(2).

For PETITIONER:

Rolando Medina
Eric Marandett
Stephanie Schoenwald
Sophie Wang
CHOATE, HALL & STEWART LLP
Rmedina@choate.com
Emarandett@choate.com
Sschonewald@choate.com
Swang@choate.com

For PATENT OWNER:

Michael Rosato
Steven Parmelee
Sonja Gerrard
WILSON, SONSINI, GOORDICH & ROSATI
Mrosato@wsgr.com
Sparmelee@wsgr.com
Sgerrard@wsgr.com

Electronic Patent Application Fee Transmittal

Application Number:	16575128
Filing Date:	18-Sep-2019
Title of Invention:	SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION
First Named Inventor/Applicant Name:	AmirAli TALASAZ
Filer:	Timothy A Hott/Michelle Chan
Attorney Docket Number:	42534-704.313

Filed as Large Entity

Filing Fees for Utility under 35 USC 111(a)

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Basic Filing:				
PETITION FEE- 37 CFR 1.17(H) (GROUP III)	1464	1	140	140
RCE- 1ST REQUEST	1801	1	1300	1300
Pages:				
Claims:				
Miscellaneous-Filing:				
Petition:				
Patent-Appeals-and-Interference:				

Description	Fee Code	Quantity	Amount	Sub-Total in USD($)
Post-Allowance-and-Post-Issuance:				
Extension-of-Time:				
Miscellaneous:				
		Total in USD ($)		**1440**

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

Decision Date : August 28, 2020

In re Application of :

AmirAli TALASAZ

DECISION ON PETITION

UNDER CFR 1.313(c)(2)

Application No : 16575128

Filed : 18-Sep-2019

Attorney Docket No : 42534-704.313

This is an electronic decision on the petition under 37 CFR 1.313(c)(2), filed August 28, 2020 , to withdraw the above-identified application from issue after payment of the issue fee.

The petition is **GRANTED.**

The above-identified application is withdrawn from issue for consideration of a submission under 37 CFR 1.114 (request for continued examination). See 37 CFR 1.313(c)(2).

Petitioner is advised that the issue fee paid in this application cannot be refunded. If, however, this application is again allowed, petitioner may request that it be applied towards the issue fee required by the new Notice of Allowance.

Telephone inquiries concerning this decision should be directed to the Patent Electronic Business Center (EBC) at 866-217-9197.

This application file is being referred to Technology Center AU 1637 for processing of the request for continuing examination under 37 CFR 1.114 .

Office of Petitions

Electronic Acknowledgement Receipt

EFS ID:	40417455
Application Number:	16575128
International Application Number:	
Confirmation Number:	9440
Title of Invention:	SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION
First Named Inventor/Applicant Name:	AmirAli TALASAZ
Customer Number:	115823
Filer:	Timothy A Hott/Michelle Chan
Filer Authorized By:	Timothy A Hott
Attorney Docket Number:	42534-704.313
Receipt Date:	28-AUG-2020
Filing Date:	18-SEP-2019
Time Stamp:	17:22:37
Application Type:	Utility under 35 USC 111(a)

Payment information:

Submitted with Payment	yes
Payment Type	DA
Payment was successfully received in RAM	$1440
RAM confirmation Number	E20208RH22339204
Deposit Account	
Authorized User	
The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:	

File Listing:

Document Number	Document Description	File Name	File Size(Bytes)/ Message Digest	Multi Part /.zip	Pages (if appl.)
1	Petition automatically granted by EFS	petition-request.pdf	31627 59a48f863741b94537229a492da1e90a365789d1	no	2
Warnings:					
Information:					
2	Request for Continued Examination (RCE)	2020-08-28_GH0002US-CON14_RCE_Conditional.pdf	1364277 74e54024182d0301dcf1eddf276b1651a03ab6ec	no	3
Warnings:					
Information:					
3	Quick Path Information Disclosure Statement	2020-08-28_GH0002US-CON14_QPIDS.pdf	195741 a6453601e1fda99a5bcdd7994d159227ef9d0e9a	no	2
Warnings:					
Information:					
4	Information Disclosure Statement (IDS) Form (SB08)	2020-08-28_GH0002US-CON14_SB08.pdf	1053190 fb59f796f8d94187176bceb91c27e8779d839977	no	4
Warnings:					
Information:					
A U.S. Patent Number Citation or a U.S. Publication Number Citation is required in the Information Disclosure Statement (IDS) form for autoloading of data into USPTO systems. You may remove the form to add the required data in order to correct the Informational Message if you are citing U.S. References. If you chose not to include U.S. References, the image of the form will be processed and be made available within the Image File Wrapper (IFW) system. However, no data will be extracted from this form. Any additional data such as Foreign Patent Documents or Non Patent Literature will be manually reviewed and keyed into USPTO systems.					
5	Other Reference-Patent/App/Search documents	IPR2019-00634-FinalWrittenDecision_2020-08-18.pdf	731090 1bc4df5152cada6534ee7ef7772b4fd478febea8	no	96
Warnings:					
Information:					

6	Other Reference-Patent/App/Search documents	IPR2019-00652-FinalWrittenDecision_2020-08-18.pdf	527384 aea9a8cbac118da8e8684665de91a125db412ffa	no	68
Warnings:					
Information:					
7	Fee Worksheet (SB06)	fee-info.pdf	32289 afa9c87a9900444407b63e7a30f7fc32274089e44	no	2
Warnings:					
Information:					
		Total Files Size (in bytes):	3935598		

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111

If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

Doc Code: PET.AUTO Document Description: Petition automatically granted by EFS-Web	PTO/SB/140 U.S. Patent and Trademark Office Department of Commerce

Electronic Petition Request	**PETITION TO WITHDRAW AN APPLICATION FROM ISSUE AFTER PAYMENT OF THE ISSUE FEE UNDER 37 CFR 1.313(c)**
Application Number	16575128
Filing Date	18-Sep-2019
First Named Inventor	AmirAli TALASAZ
Art Unit	1637
Examiner Name	KENNETH HORLICK
Attorney Docket Number	42534-704.313
Title	SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION

An application may be withdrawn from issue for further action upon petition by the applicant. To request that the Office withdraw an application from issue, applicant must file a petition under this section including the fee set forth in § 1.17(h) and a showing of good and sufficient reasons why withdrawal of the application from issue is necessary.

APPLICANT HEREBY PETITIONS TO WITHDRAW THIS APPLICATION FROM ISSUE UNDER 37 CFR 1.313(c).

A grantable petition requires the following items:
(1) Petition fee; and
(2) One of the following reasons:
(a) Unpatentability of one or more claims, which must be accompanied by an unequivocal statement that one or more claims are unpatentable, an amendment to such claim or claims, and an explanation as to how the amendment causes such claim or claims to be patentable;
(b) Consideration of a request for continued examination in compliance with § 1.114 (for a utility or plant application only); or
(c) Express abandonment of the application. Such express abandonment may be in favor of a continuing application, but not a CPA under 37 CFR 1.53(d).

Petition Fee

- ◯ Small Entity
- ◯ Micro Entity
- ◉ Regular Undiscounted

Reason for withdrawal from issue

◯ One or more claims are unpatentable ◉ Consideration of a request for continued examination (RCE) (List of Required Documents and Fees) ◯ Applicant hereby expressly abandons the instant application (any attorney/agent signing for this reason must have power of attorney pursuant to 37 CFR 1.32(b)).
RCE request,submission, and fee. ☐ I certify, in accordance with 37 CFR 1.4(d)(4) that : The RCE request ,submission, and fee have already been filed in the above-identified application on ☒ Are attached.
THIS PORTION MUST BE COMPLETED BY THE SIGNATORY OR SIGNATORIES I certify, in accordance with 37 CFR 1.4(d)(4) that I am: ◉ An attorney or agent registered to practice before the Patent and Trademark Office who has been given power of attorney in this application. ◯ An attorney or agent registered to practice before the Patent and Trademark Office, acting in a representative capacity. ◯ A sole inventor ◯ A joint inventor; I certify that I am authorized to sign this submission on behalf of all of the inventors as evidenced by the power of attorney in the application ◯ A joint inventor; all of whom are signing this e-petition

Signature	/Timothy A. Hott/
Name	Timothy A. Hott
Registration Number	67740

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/575,128	09/18/2019	AmirAli TALASAZ	42534-704.313	9440

115823 7590 09/09/2020

Wilson Sonsini Goodrich & Rosati / Guardant Health
650 Page Mill Road
Palo Alto, CA 94304

EXAMINER	
HORLICK, KENNETH R	

ART UNIT	PAPER NUMBER
1637	

NOTIFICATION DATE	DELIVERY MODE
09/09/2020	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

Patents@guardanthealth.com
patentdocket@wsgr.com

CORRECTED ***Notice of Allowability***	**Application No.** 16/575,128	**Applicant(s)** TALASAZ, AmirAli	
	Examiner KENNETH R HORLICK	**Art Unit** 1637	**AIA (FITF) Status** Yes

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☑ This communication is responsive to the submission of 08/28/20.
   ☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2. ☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

3. ☑ The allowed claim(s) is/are 31-56 and 58-61 . As a result of the allowed claim(s), you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information , please see **http://www.uspto.gov/patents/init_events/pph/index.jsp** or send an inquiry to **PPHfeedback@uspto.gov.**

4. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

   Certified copies:

   a) ☐ All b) ☐ Some *c) ☐ None of the:

   1. ☐ Certified copies of the priority documents have been received.
   2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
   3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____ .

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____ .

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).

6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☑ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date 8/28/20.
3. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material _____.
4. ☐ Interview Summary (PTO-413), Paper No./Mail Date. _____.
5. ☐ Examiner's Amendment/Comment
6. ☐ Examiner's Statement of Reasons for Allowance
7. ☐ Other _____.

/KENNETH R HORLICK/
Primary Examiner, Art Unit 1637

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed

PTO/SB/08a (02-18)
Approved for use through 11/30/2020. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it contains a valid OMB control number.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT

(Not for submission under 37 CFR 1.99)

Application Number	16575128
Filing Date	2019-09-18
First Named Inventor	AmirAli TALASAZ
Art Unit	1637
Examiner Name	Kenneth R. HORLICK
Attorney Docket Number	42534-704.313

U.S.PATENTS Remove

Examiner Initial*	Cite No	Patent Number	Kind Code[1]	Issue Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					

If you wish to add additional U.S. Patent citation information please click the Add button. Add

U.S.PATENT APPLICATION PUBLICATIONS Remove

Examiner Initial*	Cite No	Publication Number	Kind Code[1]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear
	1					

If you wish to add additional U.S. Published Application citation information please click the Add button. Add

FOREIGN PATENT DOCUMENTS Remove

Examiner Initial*	Cite No	Foreign Document Number[3]	Country Code[2] i	Kind Code[4]	Publication Date	Name of Patentee or Applicant of cited Document	Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear	T[5]
	1							

If you wish to add additional Foreign Patent Document citation information please click the Add button Add

NON-PATENT LITERATURE DOCUMENTS Remove

Examiner Initials*	Cite No	Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published.	T[5]

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	16575128
	Filing Date	2019-09-18
	First Named Inventor	AmirAli TALASAZ
	Art Unit	1637
	Examiner Name	Kenneth R. HORLICK
	Attorney Docket Number	42534-704.313

/K.R.H/	1	IPR2019-00634, Final Written Decision of U.S. Patent 9,840,743, dated August 18, 2020	
/K.R.H/	2	IPR2019-00652, Final Written Decision of U.S. Patent 9,834,822, dated August 18, 2020	

If you wish to add additional non-patent literature document citation information please click the Add button Add

EXAMINER SIGNATURE

Examiner Signature	/KENNETH R HORLICK/	Date Considered	09/03/2020

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

INFORMATION DISCLOSURE STATEMENT BY APPLICANT (Not for submission under 37 CFR 1.99)		
	Application Number	16575128
	Filing Date	2019-09-18
	First Named Inventor	AmirAli TALASAZ
	Art Unit	1637
	Examiner Name	Kenneth R. HORLICK
	Attorney Docket Number	42534-704.313

CERTIFICATION STATEMENT

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

OR

☒ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

See attached certification statement.

☒ The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

A certification statement is not submitted herewith.

SIGNATURE

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

Signature	/Timothy A. Hott/	Date (YYYY-MM-DD)	2020-08-28
Name/Print	Timothy A. Hott	Registration Number	67740

This collection of information is required by 37 CFR 1.97 and 1.98. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 1 hour to complete, including gathering, preparing and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

To: Patents@guardanthealth.com,patentdocket@wsgr.com,
From: PAIR_eOfficeAction@uspto.gov
Cc: PAIR_eOfficeAction@uspto.gov
Subject: Private PAIR Correspondence Notification for Customer Number 115823

Sep 09, 2020 03:37:59 AM

Dear PAIR Customer:

Wilson Sonsini Goodrich & Rosati / Guardant Health
650 Page Mill Road
Palo Alto, CA 94304
UNITED STATES

The following USPTO patent application(s) associated with your Customer Number, 115823 , have new outgoing correspondence. This correspondence is now available for viewing in Private PAIR.

The official date of notification of the outgoing correspondence will be indicated on the form PTOL-90 accompanying the correspondence.

Disclaimer:
The list of documents shown below is provided as a courtesy and is not part of the official file wrapper. The content of the images shown in PAIR is the official record.

Application	Document	Mailroom Date	Attorney Docket No.
16575128	NOA	09/09/2020	42534-704.313
	1449	09/09/2020	42534-704.313

To view your correspondence online or update your email addresses, please visit us anytime at https://sportal.uspto.gov/secure/myportal/privatepair.

If you have any questions, please email the Electronic Business Center (EBC) at EBC@uspto.gov with 'e-Office Action' on the subject line or call 1-866-217-9197 during the following hours:

Monday - Friday 6:00 a.m. to 12:00 a.m.

Thank you for prompt attention to this notice,

UNITED STATES PATENT AND TRADEMARK OFFICE

PATENT APPLICATION INFORMATION RETRIEVAL SYSTEM

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	ISSUE DATE	PATENT NO.	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/575,128	10/06/2020	10793916	42534-704.313	9440

115823 7590 09/16/2020

Wilson Sonsini Goodrich & Rosati / Guardant Health
650 Page Mill Road
Palo Alto, CA 94304

ISSUE NOTIFICATION

The projected patent number and issue date are specified above.

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(application filed on or after May 29, 2000)

The Patent Term Adjustment is 0 day(s). Any patent to issue from the above-identified application will include an indication of the adjustment on the front page.

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Application Assistance Unit (AAU) of the Office of Data Management (ODM) at (571)-272-4200.

APPLICANT(s) (Please see PAIR WEB site http://pair.uspto.gov for additional applicants):

AmirAli TALASAZ, Atherton, CA;
GUARDANT HEALTH, INC., Redwood City, CA;

The United States represents the largest, most dynamic marketplace in the world and is an unparalleled location for business investment, innovation, and commercialization of new technologies. The USA offers tremendous resources and advantages for those who invest and manufacture goods here. Through SelectUSA, our nation works to encourage and facilitate business investment. To learn more about why the USA is the best country in the world to develop technology, manufacture products, and grow your business, visit SelectUSA.gov.

To: Patents@guardanthealth.com,patentdocket@wsgr.com,
From: PAIR_eOfficeAction@uspto.gov
Cc: PAIR_eOfficeAction@uspto.gov
Subject: Private PAIR Correspondence Notification for Customer Number 115823

Sep 17, 2020 04:22:45 AM

Dear PAIR Customer:

Wilson Sonsini Goodrich & Rosati / Guardant Health
650 Page Mill Road
Palo Alto, CA 94304
UNITED STATES

The following USPTO patent application(s) associated with your Customer Number, 115823 , have new outgoing correspondence. This correspondence is now available for viewing in Private PAIR.

The official date of notification of the outgoing correspondence will be indicated on the form PTOL-90 accompanying the correspondence.

Disclaimer:
The list of documents shown below is provided as a courtesy and is not part of the official file wrapper. The content of the images shown in PAIR is the official record.

Application	Document	Mailroom Date	Attorney Docket No.
16575128	ISSUE.NTF	09/16/2020	42534-704.313

To view your correspondence online or update your email addresses, please visit us anytime at https://sportal.uspto.gov/secure/myportal/privatepair.

If you have any questions, please email the Electronic Business Center (EBC) at EBC@uspto.gov with 'e-Office Action' on the subject line or call 1-866-217-9197 during the following hours:

Monday - Friday 6:00 a.m. to 12:00 a.m.

Thank you for prompt attention to this notice,

UNITED STATES PATENT AND TRADEMARK OFFICE

PATENT APPLICATION INFORMATION RETRIEVAL SYSTEM

AO 120 (Rev. 08/10)

TO:	REPORT ON THE FILING OR DETERMINATION OF AN ACTION REGARDING A PATENT OR TRADEMARK
Mail Stop 8 Director of the U.S. Patent and Trademark Office P.O. Box 1450 Alexandria, VA 22313-1450	

In Compliance with 35 U.S.C. § 290 and/or 15 U.S.C. § 1116 you are hereby advised that a court action has been filed in the U.S. District Court District of Delaware on the following

☐ Trademarks or ☑ Patents. (☐ the patent action involves 35 U.S.C. § 292.):

DOCKET NO.	DATE FILED	U.S. DISTRICT COURT
	11/23/2020	District of Delaware

PLAINTIFF	DEFENDANT
Guardant Health, Inc.	Foundation Medicine, Inc.

	PATENT OR TRADEMARK NO.	DATE OF PATENT OR TRADEMARK	HOLDER OF PATENT OR TRADEMARK
1	See Attached		
2			
3			
4			
5			

In the above—entitled case, the following patent(s)/ trademark(s) have been included:

DATE INCLUDED	INCLUDED BY ☐ Amendment ☐ Answer ☐ Cross Bill ☐ Other Pleading	
PATENT OR TRADEMARK NO.	DATE OF PATENT OR TRADEMARK	HOLDER OF PATENT OR TRADEMARK
1		
2		
3		
4		
5		

In the above—entitled case, the following decision has been rendered or judgement issued:

DECISION/JUDGEMENT

CLERK	(BY) DEPUTY CLERK	DATE

Copy 1—Upon initiation of action, mail this copy to Director Copy 3—Upon termination of action, mail this copy to Director
Copy 2—Upon filing document adding patent(s), mail this copy to Director Copy 4—Case file copy

	PATENT OR TRADEMARK NO.	DATE OF PATENT OR TRADEMARK	HOLDER OF PATENT OR TRADEMARK
1	US 10,501,810 B2	12/10/2019	Guardant Health, Inc.
2	US 10,704,085 B2	7/7/2020	Guardant Health, Inc.
3	US 10,704,086 B2	7/7/2020	Guardant Health, Inc.
4	US 10,793,916 B2	10/6/2020	Guardant Health, Inc.
5	US 10,801,063 B2	10/13/2020	Guardant Health, Inc.
6	US 9,840,743 B2	12/12/2017	Guardant Health, Inc.
7	US 9,834,822 B2	12/5/2017	Guardant Health, Inc.

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE

PATENT WITHDRAWAL NOTICE

DATE WITHDRAWN	WITHDRAWAL NUMBER
03/31/2021	59743

The following application has been **WITHDRAWN** from the 09/08/2020 issue.

SERIAL NO.	PATENT NUMBER
16/575,128	10767233

TITLE

SYSTEMS AND METHODS TO DETECT RARE MUTATIONS AND COPY NUMBER VARIATION

NAME AND ADDRESS

AmirAli TALASAZ
Redwood City, CA

REASON FOR WITHDRAWAL

Auto-petition to withdraw - Granted

APPROVED

/Kimberly House/ , Manager
Patent Publication Branch
Office of Data Management

FORM PTO-302 -- (REV. 05-2009)

AO 120 (Rev. 08/10)

TO:	REPORT ON THE FILING OR DETERMINATION OF AN ACTION REGARDING A PATENT OR TRADEMARK
Mail Stop 8 Director of the U.S. Patent and Trademark Office P.O. Box 1450 Alexandria, VA 22313-1450	

In Compliance with 35 U.S.C. § 290 and/or 15 U.S.C. § 1116 you are hereby advised that a court action has been filed in the U.S. District Court District of Delaware on the following

☐ Trademarks or ☑ Patents. (☐ the patent action involves 35 U.S.C. § 292.):

DOCKET NO.	DATE FILED	U.S. DISTRICT COURT
20-1580-LPS	11/23/2020	District of Delaware

PLAINTIFF	DEFENDANT
Guardant Health, Inc.	Foundation Medicine, Inc.

PATENT OR TRADEMARK NO.	DATE OF PATENT OR TRADEMARK	HOLDER OF PATENT OR TRADEMARK
1 See Attached		
2		
3		
4		
5		

In the above—entitled case, the following patent(s)/ trademark(s) have been included:

DATE INCLUDED	INCLUDED BY ☐ Amendment ☐ Answer ☐ Cross Bill ☐ Other Pleading	
PATENT OR TRADEMARK NO.	DATE OF PATENT OR TRADEMARK	HOLDER OF PATENT OR TRADEMARK
1		
2		
3		
4		
5		

In the above—entitled case, the following decision has been rendered or judgement issued:

DECISION/JUDGEMENT
Stip. of Dismissal

CLERK	(BY) DEPUTY CLERK	DATE
John A. Cerino		1-6-2022

Copy 1—Upon initiation of action, mail this copy to Director Copy 3—Upon termination of action, mail this copy to Director
Copy 2—Upon filing document adding patent(s), mail this copy to Director Copy 4—Case file copy

	PATENT OR TRADEMARK NO.	DATE OF PATENT OR TRADEMARK	HOLDER OF PATENT OR TRADEMARK
1	US 10,501,810 B2	12/10/2019	Guardant Health, Inc.
2	US 10,704,085 B2	7/7/2020	Guardant Health, Inc.
3	US 10,704,086 B2	7/7/2020	Guardant Health, Inc.
4	US 10,793,916 B2	10/6/2020	Guardant Health, Inc.
5	US 10,801,063 B2	10/13/2020	Guardant Health, Inc.
6	US 9,840,743 B2	12/12/2017	Guardant Health, Inc.
7	US 9,834,822 B2	12/5/2017	Guardant Health, Inc.

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

GUARDANT HEALTH, INC., Plaintiff, v. FOUNDATION MEDICINE, INC., Defendant.	C. A. No. 20-cv-1580-LPS

JOINT STIPULATION AND PROPOSED ORDER OF DISMISSAL

Pursuant to Federal Rule of Civil Procedure 41(a)(1)(A)(ii), IT IS HEREBY STIPULATED by Plaintiff Guardant Health, Inc. and Defendant Foundation Medicine, Inc. that:

1) All claims, defenses, and counterclaims in these actions are hereby dismissed with prejudice;
2) No party admits liability;
3) Each party shall bear its own costs, attorneys' fees, and expenses incurred in connection with these actions; and
4) The U.S. District Court for the District of Delaware retains jurisdiction to enforce and resolve any disputes arising under the settlement agreement between the parties.

Dated: January 5, 2022

FARNAN LLP

/s/ Brian E. Farnan
Brian E. Farnan (#4089)
Michael J. Farnan (#5165)
919 North Market Street, 12th Floor
Wilmington, DE 19801
(302) 777-0300
bfarnan@farnanlaw.com
mfarnan@farnanlaw.com

Attorneys for Guardant Health, Inc

Respectfully submitted,

MORRIS, NICHOLS, ARSHT & TUNNELL LLP

/s/ Jeremy A. Tigan
Karen Jacobs (#2881)
Jeremy A. Tigan (#5239)
1201 North Market Street
P.O. Box 1347
Wilmington, DE 19899
(302) 658-9200
kjacobs@morrisnichols.com
jtigan@morrisnichols.com

Attorneys for Foundation Medicine, Inc.

IT IS SO ORDERED this 6th day of January, 2022.

The Honorable Leonard P. Stark

TEMPUS EX1002.301