

A. SUPPLEMENTARY METHODS.

Rationale for selection an AML genome for sequencing. We selected a specific genome for study based on the following criteria:

1. *De novo* AML in an adult patient less than 60 years old. Patients enrolled in the study were greater than 18 years of age, and with no antecedent history of myelodysplastic syndrome or prior chemotherapy or radiation therapy for cancer, since treatment-related AML may have a different genetic basis than *de novo* AML. Further, several Cooperative Group retrospective analyses have suggested that both the clinical course and the biology of AML in patients over the age of 60 years may be clinically different from those patients less than age 60 with *de novo* AML¹.
2. High sample purity. The morphologic and flow cytometric criteria for M1 AML required high sample purity (e.g. a blast count of >80%) to eliminate the need to further purify malignant cells.
3. A simple, diploid genome. A normal karyotype was required, and we further required the sample to have no somatic copy number alterations based on Affymetrix 6.0 SNP array analysis comparing the tumor and associated normal skin sample.
4. Typical features of M1 AML. We required a gene expression profile typical of M1 AML with normal cytogenetics, to assure that the sample was not an outlier for this subtype.
5. Adequate sample abundance. Adequate amounts of high quality DNA and RNA had to be available both from the bone marrow-derived tumor sample, and from the normal tissue (obtained from a skin biopsy at the time of tumor banking) to assure that all genomic studies could be completed.
6. At least one known AML-specific somatic mutation to serve as a positive control. This assures that the patient has the correct diagnosis, and that DNA samples are suitable for the detection of somatic mutations.
7. IRB approved informed consent for whole genome sequencing.

From the 371 cases of AML banked on our AML protocol as of 12/06, 60 had the FAB M1 subtype. 21 met some of these criteria, and 6 met them all. From these 6, we selected one case for study, UPN 933124, primarily because the patient's initial bone marrow sample contained virtually 100% myeloblasts, eliminating concerns that contaminating normal cells would complicate our analyses. Although this case was carefully selected, the presentation of AML was typical, and the case is representative of most patients with M1 AML, as described below.

Case presentation of UPN 933124. A previously healthy Caucasian female in her late 50s presented with a sudden onset of sore throat and easy bruisability. There was no antecedent history of prior cancer, or administration of cytotoxic drugs and/or radiation therapy. The patient's family history was obtained by a trained genetic counselor and was positive for late onset cancers in her mother and several of her mother's siblings, including a maternal uncle with a reported case of AML that was confirmed by death certificate. The patient's sister has Essential Thrombocythemia (*JAK2 V617F* negative). None of the patient's children had a history of cancer. The patient's pedigree is being withheld from this report to protect the privacy of the family. The patient was found to have a peripheral blood white blood cell count of 105,000 cells per microliter with 85% blasts on presentation. The hemoglobin was 12.1 grams per deciliter, and platelet count was 20,000 per microliter. A bone marrow biopsy and aspirate revealed 100% blasts, and confirmed the diagnosis of AML, FAB M1 subtype (Supplementary Fig. 1). Flow cytometry revealed a single population of myeloblasts with cell surface expression of CD-13, -33 and -117, and absent expression of CD-34 and HLA-DR (Supplementary Fig. 1). Cytogenetic analysis of tumor cells revealed a normal 46 XX karyotype. After informed consent was obtained, she was enrolled in a cooperative group protocol for treatment of AML in adults less than age 60 (Cancer and Leukemia Group B (CALGB) #19808), and a Washington University Human Studies Committee-approved protocol for tissue banking in patients with MDS and AML (HSC #01-1014). Specimens of bone marrow, peripheral blood, and skin were collected and cryopreserved for genomic studies. After emergent leukopheresis, she underwent induction chemotherapy with cytarabine, daunorubicin, and etoposide. A bone marrow biopsy repeated 2 weeks later demonstrated chemoablation. After myeloid recovery, a repeat bone marrow biopsy demonstrated first complete remission. On the basis of her intermediate risk cytogenetics, she was assigned to undergo consolidation chemotherapy with high dose cytarabine and etoposide, followed by filgrastim to facilitate autologous hematopoietic stem cell mobilization and collection. She subsequently underwent dose-intensive busulfan and etoposide conditioning, followed by autologous hematopoietic stem cell rescue. Upon recovery from autologous stem cell transplantation, she was assigned to receive maintenance immunotherapy with IL-2. Despite these multiple therapies, she relapsed 11 months after her initial diagnosis. After failing to respond to a salvage regimen of mitoxantrone, etoposide, and cytarabine, she received fludarabine, high dose Ara-C, idarubicin and gemtuzumab ozogamicin. 10 days later she received high dose cyclophosphamide conditioning (60 mg/kg daily x 2) and then single dose total body irradiation (550 cGy; 30 cGy/min) followed by allogeneic hematopoietic stem cell transplantation from an HLA-matched sibling donor. A bone marrow biopsy performed one month after transplantation confirmed second complete remission. Unfortunately, her disease relapsed four months later. Gemtuzumab ozogamicin was again administered, followed by two donor lymphocyte infusions, without clearance of circulating blasts. She was then successively treated with azacytidine, low dose cytarabine, and decitabine, again without appreciable response, and expired 24 months after her initial diagnosis. Informed consent for whole genome sequencing was subsequently obtained from her next of kin and family members, using a revised Washington University Human Studies Committee-approved protocol that specifically addresses privacy and confidentiality issues associated with whole genome sequencing (HSC #01-1014, Amendment # 17).

Informed consent and tissue banking. Informed consent (obtained in accordance with the Declaration of Helsinki) for tissue banking was obtained from the patient at initial clinical presentation, and she was enrolled on WU HSC protocol #01-1014, “Tissue Acquisition for Analysis of Genetic Progression Factors in Hematologic Diseases,” for which all patients with MDS and AML referred to the Siteman Cancer Center at Washington University since 11/03 have been routinely approached. Eligibility criteria included age greater than 18, more than 30% myeloblasts in the bone marrow, and the absence of antecedent chemotherapy, radiation therapy, or history of myelodysplasia. Samples with 2 or fewer clonal cytogenetic abnormalities were given preference for inclusion in the study. Peripheral blood and bone marrow were obtained for analysis of tumor cells, and a 6-mm punch biopsy of skin was obtained for analysis of unaffected somatic cells. Following collection, de-identified tissue specimens were labeled with a randomly generated Unique Patient Identifier (UPN), and then transported to the Siteman Cancer Center Tissue Procurement Core. Details of clinical presentations and subsequent courses are maintained prospectively in a database by a study-dedicated data manager, to whom knowledge of patient identity is restricted. Subsequent to the death of patient 933124, the WU HSC approved an amendment to the protocol that incorporated language specifically permitting the use of banked DNA for whole genome sequencing, to which her husband consented. The patient’s sister also consented to provide a detailed family history, which was obtained by a trained genetic counselor, as explicitly outlined in the revised protocol.

DNA was prepared from all tissues using standard protocols. One hundred *de novo* AML samples from patients with the same entry criteria were also obtained from Cancer and Leukemia Group B (CALGB)²; similar numbers of patients with the M0/1, M2, M3, and M4 French-American-British (FAB) subtype were selected for analysis. The institutional review boards of all participating institutions approved this study.

Expression Analysis. Bone marrow aspirates were obtained from properly consented AML patients, and RNA was prepared from the unfractionated snap-frozen cell pellets. Total cellular RNA was purified using the Trizol reagent (Invitrogen, Carlsbad, CA), quantified using UV spectroscopy (Nanodrop Technologies), and qualitatively assessed using a BioAnalyzer 2100 and RNA NanoChip assay (Agilent Technologies, Palo Alto, CA). For all samples, 2 ug total cellular RNA was amplified, labeled, and hybridized to Affymetrix Human U133 Plus 2.0 Array GeneChip microarrays (Affymetrix Inc, Santa Clara, CA) using standard protocols from Affymetrix and the Siteman Cancer Center Multiplexed Gene Analysis Core Facility. To perform interarray comparisons, the raw scan data from each microarray were scaled to a target intensity of 1500 using the Affymetrix GCOS 1.2 (MAS 5) statistical algorithm. Scaled data from each array were exported to the Siteman Cancer Center Bioinformatics Server (<http://bioinformatics.wustl.edu>), merged with updated gene annotation data for each probeset on the array, and downloaded for further data visualization and analysis³. The complete dataset has been

analyzed in detail in a separate study (J.E.P., N.R.Grieselhuber, L.W.Chang, M.Murakami, W.Yuan, D.C.L, R.N., M.A.W., T.J.L, manuscript in preparation) and has been publicly deposited (GEO accession number, GSE10358).

Expression data was filtered to eliminate all probes with <10% present calls across all samples and those without significant variation across FAB subclasses. These were next subjected to unsupervised clustering using Ward's method in Spotfire DecisionSite (Spotfire Inc, Somerville, MA)³.

Copy Number and Loss of Heterozygosity Analysis. Peripheral blood and bone marrow were obtained for analysis of tumor cells, and a 6-mm punch biopsy of skin was obtained for analysis of unaffected somatic cells. DNA was prepared from both tissues using standard protocols. DNA was quantified using UV spectroscopy (Nanodrop Technologies), and qualitatively assessed using standard agarose gel electrophoresis. For all samples, 0.5 ug DNA was digested with Nsp and Sty enzymes, amplified, fragmented and labeled, and hybridized to Affymetrix Genome-Wide Human SNP 6.0 Array GeneChip microarrays (Affymetrix Inc, Santa Clara, CA) using standard protocols from Affymetrix and the Siteman Cancer Center Multiplexed Gene Analysis Core Facility. Array image data was analyzed using Affymetrix GCOS 1.4 operating software and Genotyping Console 2.1 to derive .cel data files, which were exported to the Siteman Cancer Center Bioinformatics Server (<http://bioinformatics.wustl.edu>), and downloaded for further data visualization and analysis. The complete dataset has been analyzed in detail in a separate study (M.J.W., J.E.P., R.R., R.N., M.A.W., T.J.L, manuscript in preparation) and will be publicly deposited.

Partek Genomic Suite (Partek Inc, Saint Louis, MO) was used for both copy number and loss of heterozygosity (LOH) analysis. SNP array signal intensity data for both SNP and CN probes (1.8 million) were analyzed to determine copy number using the standard Partek workflow for paired sample copy number analysis, which calculates the ratio of tumor copy number to that of paired normal/skin. The birdseed v2.0 algorithm was used to derive genotypes from SNP probes (900,000). Genotyping results from paired skin and tumor samples were compared to find regions of LOH. The results of both analyses were plotted as a heatmap using Spotfire DecisionSite (Spotfire Inc, Somerville, MA).

Library construction. Illumina genomic DNA libraries were produced according to the manufacturer's supplied protocol, using a 150-200 bp insert size and Illumina single end read adapters. To reduce redundant reads resulting from PCR biasing of the libraries, we adopted a strategy whereby each library was used to produce only 10 flow cells. Following titration, each library was diluted and amplified into clusters according to the Illumina Cluster Station protocols. Each flow cell was analyzed on an Illumina Genome Analyzer, using 32 bp read length for the tumor genome reads and a 35 bp read length for the

normal genome reads. The Illumina processing pipeline, run in a parallelized fashion, performed all post-run quality filtering and base calling. The completed high quality reads were added to our LIMS database.

Algorithms for detecting SNVs. The basic approach for the Decision Tree algorithm for SNV detection was outlined in the Results section. Specifically, SNVs were retained for additional analysis when they were supported by 7 or more unique start-site reads and where the SNV has a Maq quality score of more than 30, or if all of the following were true:

- (1) the SNV had read support with a maximum base quality greater than 26,
- (2) was supported by 3 or more unique start-site reads supporting the SNV, and
- (3) the reads supporting the SNV had an average base quality greater than 16.

Since the skin sample contained contaminating leukemic cells (based on our deep readcount analysis on 454), a requirement of no support for a variant in skin was overly strict and resulted in false negatives. Instead, we used a one-sided binomial test as an unbiased ranking of each potential somatic variant. Due to the small number of reads at each position in skin, we accepted any SNV as a potential tumor variant if it had no supporting reads in skin, or had a p-value greater than 0.65, since the number of skin reads was greater than the coverage-adjusted rate of contamination (about 10% of the skin sample was derived from circulating leukemic cells). This yielded a set of high confidence, potentially somatic variants that were validated by re-sequencing PCR-generated amplicons on the ABI 3730xl platform).

Assigning conservation scores for somatic mutations. The conservation score listed for each mutation in Table 2 is the posterior probability that the element in question is conserved across 28 vertebrate genomes, as calculated by PhastCons * using a phylogenetic hidden Markov model (phylo-HMM)⁴.

Stringent indel detection. We first compared all Solexa tumor and normal reads to the full human genome (hg 36) using ELAND through the Solexa read processing pipeline, removing all reads that did not match to the human genome with 0, 1 or 2 mismatches to a file. These non-matched reads (242,352,962 reads for tumor genome, 105,026,540 for normal) were then compared to a transcriptome database that was derived from the AceView database using cross_match (version 080214.1, Phil Green, personal communication). The following cross_match parameters were used: gap_init: -4, gap_ext: -3, ins_gap_ext: -3, del_gap_ext: -3, -minscore 8 -discrep_lists -minmatch 12 -maxmatch 14 -raw -word_raw -masklevel -1 -tags, with matches assigned a score of +1 and mismatches a score of -2.

We filtered the initial cross_match alignments by first identifying specific indels that were identified either within the normal or tumor read sets as follows:

1. Keep only alignments where ≤ 2 bp are unaligned at their 3' end and no more than 1 bp is unaligned at the 5' end,
2. Keep only indels for which at least two reads identifying the discrepant base have a quality value of $>Q19$ in that position.
3. Remove any putative indel sites with
 - a. more than one indel type at the same position,
 - b. all reads starting or ending at the same site,
 - c. more than one high quality discrepancy ($>Q19$) within the alignment,
 - d. all supporting reads obtained from the same flow cell,
 - e. more than one indel site identified per transcript (frequency greater than 1 per 500 bp).

The clusters of reads identifying putative indels that survived this stringent filtering then were aligned against the human genome (hs36) using cross_match (with the parameters listed above), to ensure best-in-genome position. For the normal set, there were 5,967 putative indels identified by 33,592 reads, and for the tumor set there were 7,781 putative indels identified by 61,350 reads. Of these, 81,843 reads were aligned against build hs36 and 98% (80,265) of them identified the same indel at the same position in the genome as had been identified in the transcriptome alignment. Further, for the indels identified by these 80,264 reads to be considered valid, the best-in-genome alignment for at least one of the reads in an indel-identifying cluster had to have an alignment score of ≥ 3 more than the next best alignment score for that read. We also required that each of the 4 bp on either side of the indel site had to have quality values of $>Q15$. The final list of putative indels required at least one such high quality read to support it.

The positions of indels were correlated with both REFSEQ and dbSNP such that those indel sites occurring within bases annotated as REFSEQ CDS and UTR regions and not already annotated in dbSNP were validated. Some indels were found only in the tumor or normal reads, while others were found in both read sets. In total, we identified 7746 indels within REFSEQ CDSs according to these criteria and filters, 4031 of which were also in dbSNP and 3715 of which are novel. Of these, 183 overlap REFSEQ CDS (91 are in dbSNP and 92 are novel). Of these, 10 were only found in the tumor set. In REFSEQ UTR regions, 1357 indels (871 are in dbSNP; 486 are novel) were identified specifically in the tumor sample. Of those 486 novel UTR indels, 338 were found only in the tumor set of which 114 met the criteria of a

≥ 3 better than the next best alignment score in the genome. Of those, 40 consisted of reads with at least 3 start sites and were validated.

Manual review was finally used to filter lists of potential indels identified by computational methods. The criteria listed below enabled the refinement of the computational methods, as well as minimized the number of validation assays necessary to discover the sequence variants in this project. This procedure enabled manual reviewers to make decisions on the likelihood that computationally indicated variants were based on solid sequence data, or rather due to sequencing artifacts from the sequencing instruments. This type of review has been employed for many years in the “finishing” process of DNA sequencing pipelines, and was helpful in informing this discovery pipeline.

The process involved the creation of consed databases, which are alignment and visual displays of the human reference sequences as well as the aligned sequences to the indicated regions of the genome. The variants (either putative indels or single nucleotide variants) as well as supporting sequences are displayed. Manual reviewers inspect the potential detected variants, and based on specific criteria, determine whether the region is worthy of further sequence validation. The criteria include: 1) at least two unique start site sequences containing the sequence variant and 2) at least one high quality base sequence or one sequence from each strand (double stranded) representing the variant. The only exceptions to those criteria are low quality A's or T's in the sequence, which are discounted regardless of strand because of their prevalence in the sequence. A blat sequence alignment is also performed, and 100% matches elsewhere are discounted.

Estimating the false negative and positive rates for mutation discovery. To estimate the raw false negative rate of SNV discovery in the tumor genome (i.e. the potential number of mutations that were missed by our analysis), we used our reference set of heterozygous SNP calls (concordant between the Affymetrix 6.0 and Illumina genotyping arrays). The same Decision Tree analysis used to detect somatic mutations also detected 40,705 of the 46,494 heterozygous SNPs defined by both the array platforms. The 5,789 bases defined as SNPs by both arrays but not detected in the tumor genome database suggests a raw false negative rate of 12.45% for heterozygous SNPs (which is the most appropriate surrogate for the heterozygous mutations that we are attempting to detect). Since only eight somatic SNVs were verified in coding sequences, this suggests that no more than a few additional non-synonymous mutations are likely to exist in this genome. Assuming that the mutation rate is similar throughout the genome, we infer that the non-coding portion of the tumor genome (98-99% of total genomic content) may contain an additional 500-1000 somatic mutations.

To assess the false positive rate, we determined the number of heterozygous SNPs passing the Decision Tree analysis that intersected genomic positions where both SNP arrays were concordant for the genotype. Of 40,128 predicted heterozygous sites called identically by both SNP arrays (i.e. ‘true’ calls),

only 23 variants from the Decision Tree analysis of tumor DNA did not agree, yielding an estimated false positive rate of 0.06%. The very low estimated false positive rate for true SNPs is very different from the apparent false positive rate of potential somatic variants (152 out of 181). The reason for the apparent contradiction is this: although the false positive rate of the Decision Tree was extremely low, we removed enormous numbers of variants that were almost certainly true positives (i.e. all variants present in the skin, and all variants reported in dbSNP-127 and the Watson and Venter genomes), which greatly enriched our set of potential somatic mutations with false positives. In addition, we purposefully set our filters to allow for some false positives, to minimize the possibility of missing relevant mutations. Based on the large number of false positives detected in the final list of somatic variants, we are confident that very few 'true' non-synonymous mutations were missed by our analysis.

At present, we do not have a 'gold standard' against which we can calibrate the raw false negative rate for indel detection. However, it was reassuring that more than half of the indels detected by our screening method previously had been described in dbSNP, and that most of the putative indels selected for secondary validation (22/24) were found to be correctly called.

Confirmation of somatic mutations by 454 sequencing. Amplicons containing exons with relevant putative mutations were generated from non-amplified genomic DNA from the primary and relapse tumor samples, and the skin biopsy obtained at initial banking (when the patient was overtly leukemic). 454 sequencing of these amplicons was performed as previously described⁵. The sequence bases and quality scores from 454 reads were extracted using the *sffinfo* program. Each FASTA file was aligned to the human reference sequence (hg36) using BLAT⁶. Due to memory constraints, the alignments were performed by chromosome and then compiled into a single output file for each of normal and tumor. FASTA files for each chromosome (1-22, X, Y), with repeats masked to lower case, were obtained from the UCSC Genome Browser Database⁷. BLAT was invoked with the *mask=lower* parameter, which excluded the repetitive sequence during the seeding of alignments. SNP and indel events were identified in the BLAT alignments with the *breakPointRead* program [unpublished]. Variants detected at the somatic sites were combined to tabulate the number of reads supporting reference and variant alleles in each sample.

Detection of copy number variations. To identify potential somatic copy number changes, we searched for regions (≥ 1 kb) having different patterns of coverage in the tumor and skin samples, based on Maq alignments. We found 260 regions in the tumor and skin genomes with no coverage, ranging in size from 1 kb to 20.3 Mb. Not surprisingly, most of them (249 of 260) were found to lie in centromeric, telomeric, and other non-genic regions of the human genome. We identified 24 inherited deletions and two small single copy amplifications in the tumor and skin genomes using sequence read counts and supported these with Affymetrix 6.0 array data analysis (Supplementary Table 1). Examples of inherited heterozygous and homozygous deletions, and an inherited single copy gain (detected on both platforms)

are shown in Supplementary Figure 4. Importantly, no large somatic deletions and/or amplifications were identified with either approach.

B. SUPPLEMENTARY DISCUSSION.

Potential roles of the mutated genes in AML pathogenesis. We detected somatic mutations in eight protein-coding genes. One mutation is in the PTPRT gene, a tyrosine phosphatase/tumor suppressor that is frequently mutated in colon cancer genomes⁸. The mutation in our patient's tumor lies in the phosphatase domain, and is predicted by the SIFT algorithm⁹ to be deleterious to the protein's function. A mutation in CDH24, which encodes cadherin-24 (a member of the type II classical cadherin family) predicts a Y590stop that is likely to cause loss-of-function, since it truncates the protein in extracellular domain 5, proximal to the transmembrane domain. A similar type I cadherin gene CDH1 (E-cadherin) frequently contains a similar germline nonsense mutation in patients with Hereditary Diffuse Gastric Cancer¹⁰. The residual allele of CDH1 can also be silenced by methylation in these tumors¹¹. We also detected a missense mutation in PCLKC (PCDH24), another member of the cadherin superfamily. Expression of PCLKC is reduced in epithelial tumors,¹² possibly by epigenetic silencing, similar to that observed with a related protocadherin (PCDH10) in primary lymphoma samples and cell lines.¹³ Therefore, both mutated cadherin family members may also be behaving as tumor suppressors in this setting. SLC15A1 is a transmembrane peptide transporter that is highly related to SLC11A1 (NRAMP1), SLC19A1, and SLC29A1 (ENT1), all of which are implicated in the transport of small molecules and drugs to cancer cells¹⁴. Biallelic mutations in SCL29A1 that result in a lack of detectable protein recently have been associated with cytarabine resistance in a human leukemia cell line¹⁵, implicating SLC family members in chemotherapy resistance for AML. The SLC15A1 mutation in 933124 is a nonsense mutation that is predicted to cause loss-of-function, perhaps reducing drug transport that contributed to the patient's ultimate resistance to all therapies.

KNDC1 encodes a protein that shares homology with guanine nucleotide exchange factors for small GTPases, such as Ral and Ras. The mutation in this gene lies between the N terminal KIND domain and the GNEF domain, and its effect on protein function is not yet known. GPR123 and EBI2 are members of the G-protein coupled receptor gene family; proteins of this class are known to participate in important cell-cell interactions in a variety of systems¹⁶. The mutations in these genes may somehow alter the ability of the patient's AML cells to interact with bone marrow stromal cells in a way that provides an advantage to cells bearing the mutation, or may alter critical intracellular signaling pathways that affect cell proliferation. Of interest, EBI2 is a constitutively active G-protein linked receptor whose expression is stimulated by Epstein-Barr infection; mutations that alter its activity or trafficking may lead to abnormal growth promotion¹⁷. Although glutamate receptors have not yet been directly implicated in cancer pathogenesis, it is interesting to note that some of these receptors are expressed in embryonic stem cells, where they play a role in self-renewal^{18,19}. Since self-renewal is thought to be an essential

feature of leukemia-initiating cells, it is tempting to speculate that the somatic mutation in GRINL1B may somehow be contributing to this property of the AML clone.

Nature of mutations that may have been missed by the short read method. As discussed in the Results section, our analysis strongly suggests that we may have missed only a very small number of SNVs in coding sequences. We have no 'gold standard' for calibrating our ability to detect indels, but we were reassured that most of the indels selected for validation (22 of 24) were found to be accurate calls. Our studies of copy number variants on high resolution Nimblegen and Affymetrix arrays (and corroborated with Solexa readcount data) makes it unlikely that we have missed somatic deletions or amplifications that are larger than 2-3 Kb in size. Although none of our data have suggested that this genome contains cryptic translocations (e.g. the tumor was cytogenetically normal, and we found no chimeric reads to support translocations, nor evidence of somatic deletions that are often seen at translocation breakpoints), the short read length approach is not optimal for detecting them. Campbell, et al²⁰ have shown that massively-parallel paired end-read sequencing technology can be used to survey the structural variants that may be present in tumor genomes with substantially lower sequence coverage (~12-fold) than was produced in the current study. However, robust read-pair methods were not available when we initiated sequencing of this subject's genome; this method has recently become available, and should provide an optimal interval distance (~2-3 kb) to facilitate the detection of translocations in the future.

Inherited polymorphisms in the 933124 genome. Although the medical history of this patient does not provide definitive evidence for a familial or syndromic origin of her leukemia, the large number of cancer diagnoses in her extended family suggests that there may be an enrichment of susceptibility alleles. Gene association studies in sporadic AML have focused on candidate genes in DNA repair and drug detoxification pathways. In addition, there is emerging evidence that germline polymorphisms in genes that are somatically mutated in sporadic MDS/AML may be low penetrance susceptibility alleles for this disease^{21,22}. We therefore compiled a comprehensive list of genes with known connections to AML pathogenesis from several sources; genes with germline mutations associated with familial or syndromic AML were drawn from the literature²³. Genes mutated in sporadic MDS/AML were selected from a highly curated list of genes with recurrent mutations in human cancer (<http://www.sanger.ac.uk/genetics/CGP/cosmic/>). A total of 119 MDS/AML-associated genes were identified (Supplementary Table 2). Five non-synonymous SNPs in these genes were detected in the tumor and skin genomes and confirmed by capillary re-sequencing (Supplementary Table 3). Two of these sequence variants are in dbSNP and have been implicated in cancer susceptibility. BRCA2 N372H homozygotes are at increased risk for breast cancer²⁴. The p53 R72 variant is a more potent inducer of apoptosis and has higher binding affinity for MDM2 compared to P72²⁵. The data supporting a role for the p53 P72R polymorphism in cancer susceptibility are conflicting, but recent studies have demonstrated an interaction between p53 codon 72 status and the genotype at MDM2 SNP309²⁶, a SNP in the MDM2 promoter that increases mRNA levels, and consequently reduces p53 activity. The

remaining three non-synonymous SNVs are not in dbSNP, although one (FLT3 V194M) was recently detected in AML and shown to lack transforming activity *in vitro*²⁷. 1,834 known and novel synonymous SNPs were detected from -10 to +5 kb from the 5' and 3' annotated boundaries of these genes (Supplementary Table 4). Case: control cohort studies will be required to determine the functional importance and relevance, if any, of these variants for AML pathogenesis.

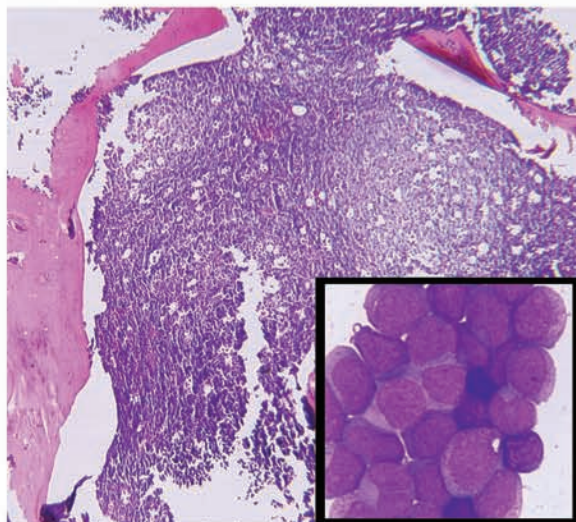
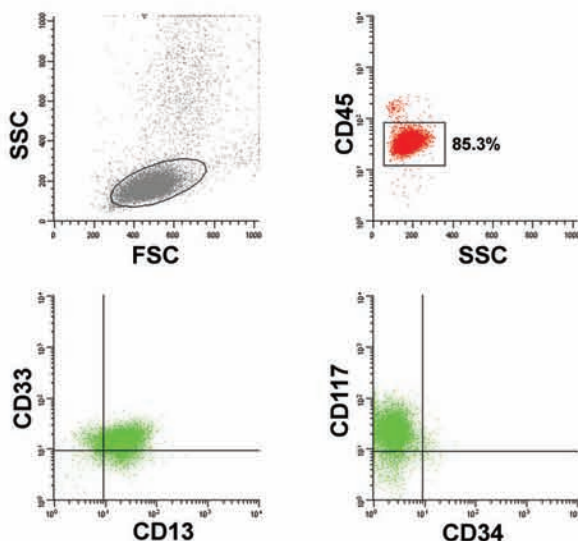
To examine the potential role of germline variants more broadly, we assembled a list of general cancer susceptibility genes from a database of variants associated with human cancer (<http://www.sanger.ac.uk/genetics/CGP/Census/>) and supplemented this with an automated PubMed query using 24 free text terms related to cancer susceptibility. Genes with 10 or more PubMed citations were included. A total of 649 "Cancer Susceptibility" genes were identified (Supplementary Table 5), containing 33 non-synonymous variants detected in the tumor and skin genomes and confirmed by capillary resequencing (Supplementary Table 6). Of the 15 known non-synonymous SNPs in this list, five (including the *TP53* and *BRCA2* alleles) have been associated previously with cancer susceptibility.

Germline copy number variants (CNVs) also may have contributed to AML susceptibility in this patient. Among the genes that lie within intervals constitutionally deleted in this individual are *WWOX*, a tumor suppressor somatically deleted or translocated in a variety of solid tumors^{28,29}. Paralogs of several other genes in the CNVs have been implicated in cancer susceptibility, including *NME1* and *ARL11*, a tumor suppressor mutated in familial CLL, breast, and colorectal cancer (Supplementary Table 1)³⁰⁻³².

C. SUPPLEMENTARY REFERENCES.

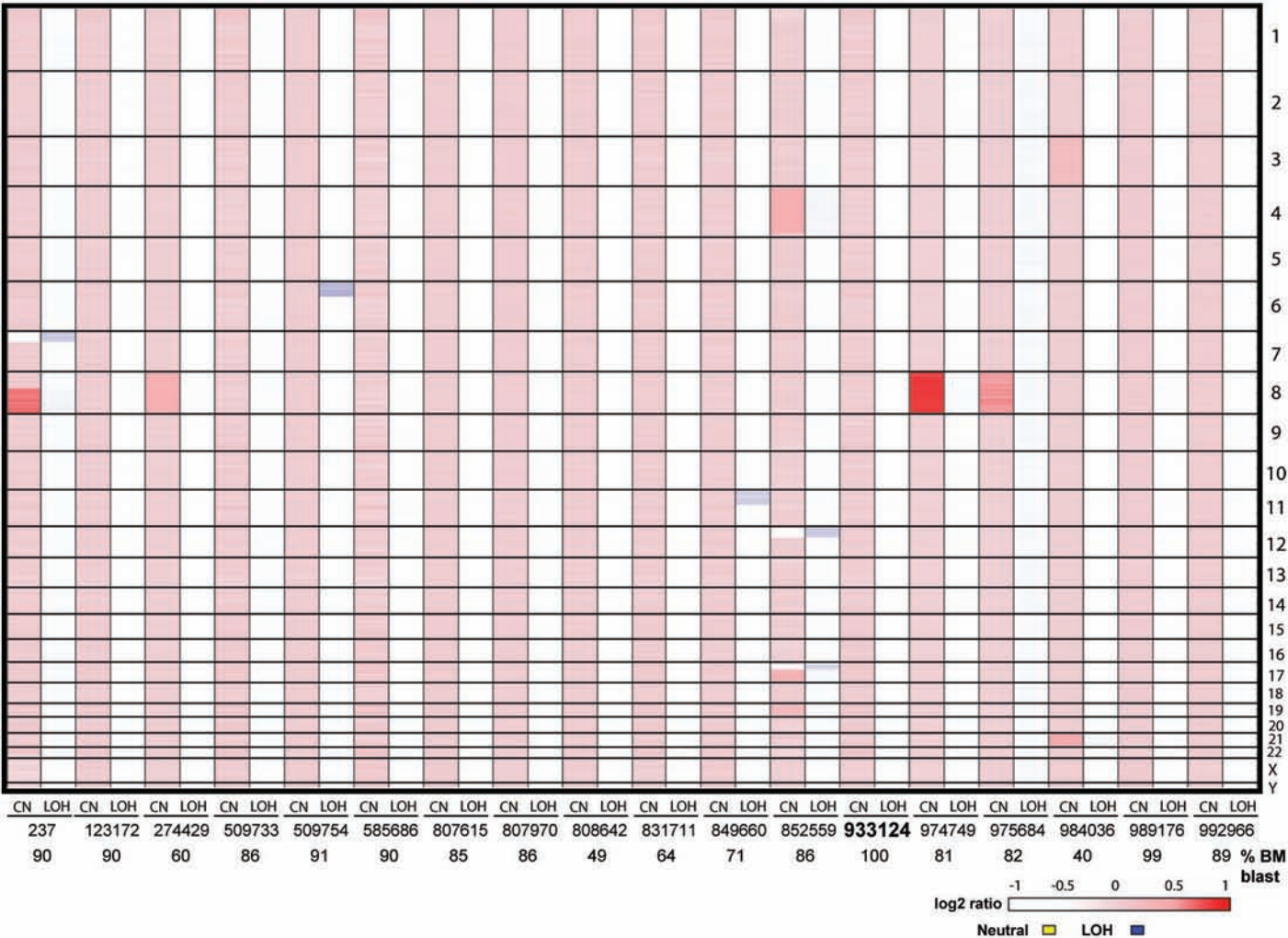
1. Appelbaum, F.R. et al. Age and acute myeloid leukemia. *Blood* **107**, 3481-5 (2006).
2. Link, D.C. et al. Distinct patterns of mutations occurring in de novo AML versus AML arising in the setting of severe congenital neutropenia. *Blood* **110**, 1648-55 (2007).
3. Yuan, W. et al. Commonly dysregulated genes in murine APL cells. *Blood* **109**, 961-70 (2007).
4. Siepel, A. et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* **15**, 1034-50 (2005).
5. Thomas, R.K. et al. High-throughput oncogene mutation profiling in human cancer. *Nat Genet* **39**, 347-351 (2007).
6. Kent, W. BLAT--the BLAST-like alignment tool. *Genome Res* **12**, 656-664 (2002).
7. Karolchik, D. et al. The UCSC Genome Browser Database. *Nucleic Acids Res* **31**, 51-4 (2003).
8. Wang, Z. et al. Mutational analysis of the tyrosine phosphatome in colorectal cancers. *Science* **304**, 1164-6 (2004).
9. Ng, P.C. & Henikoff, S. SIFT: predicting amino acid changes that affect protein function. *Nucl. Acids Res.* **31**, 3812-3814 (2003).
10. Richards, F.M. et al. Germline E-cadherin gene (CDH1) mutations predispose to familial gastric cancer and colorectal cancer. *Hum Mol Genet* **8**, 607-10 (1999).

11. Grady, W.M. et al. Methylation of the CDH1 promoter as the second genetic hit in hereditary diffuse gastric cancer. *Nat Genet* **26**, 16-7 (2000).
12. Okazaki, N., Takahashi, N., Kojima, S., Masuho, Y. & Koga, H. Protocadherin LKC, a new candidate for a tumor suppressor of colon and liver cancers, its association with contact inhibition of cell proliferation. *Carcinogenesis* **23**, 1139-48 (2002).
13. Ying, J. et al. Frequent epigenetic silencing of protocadherin 10 by methylation in multiple haematologic malignancies. *Br J Haematol* **136**, 829-32 (2007).
14. Cheng, Q. et al. Karyotypic abnormalities create discordance of germline genotype and cancer cell phenotypes. *Nat Genet* **37**, 878-82 (2005).
15. Cai, J. et al. Two distinct molecular mechanisms underlying cytarabine resistance in human leukemic cells. *Cancer Res* **68**, 2349-57 (2008).
16. Lagerstrom, M.C. et al. The evolutionary history and tissue mapping of GPR123: specific CNS expression pattern predominantly in thalamic nuclei and regions containing large pyramidal cells. *J Neurochem* **100**, 1129-42 (2007).
17. Sodhi, A., Montaner, S. & Gutkind, J.S. Viral hijacking of G-protein-coupled-receptor signalling networks. *Nat Rev Mol Cell Biol* **5**, 998-1012 (2004).
18. Melchiorri, D. et al. Metabotropic glutamate receptors in stem/progenitor cells. *Neuropharmacology* **53**, 473-80 (2007).
19. Spinsanti, P. et al. Endogenously activated mGlu5 metabotropic glutamate receptors sustain the increase in c-Myc expression induced by leukaemia inhibitory factor in cultured mouse embryonic stem cells. *J Neurochem* **99**, 299-307 (2006).
20. Campbell, P.J. et al. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nat Genet* (2008).
21. Wolfler, A. et al. A functional single-nucleotide polymorphism of the G-CSF receptor gene predisposes individuals to high-risk myelodysplastic syndrome. *Blood* **105**, 3731-6 (2005).
22. Schnittger, S. et al. D324N single-nucleotide polymorphism in the FLT3 gene is associated with higher risk of myeloid leukemias. *Genes Chromosomes Cancer* **45**, 332-7 (2006).
23. Owen, C., Barnett, M. & Fitzgibbon, J. Familial myelodysplasia and acute myeloid leukaemia--a review. *Br J Haematol* **140**, 123-32 (2008).
24. Healey, C.S. et al. A common variant in BRCA2 is associated with both breast cancer risk and prenatal viability. *Nat Genet* **26**, 362-4 (2000).
25. Dumont, P., Leu, J.I., Della Pietra, A.C., 3rd, George, D.L. & Murphy, M. The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. *Nat Genet* **33**, 357-65 (2003).
26. Bougeard, G. et al. Impact of the MDM2 SNP309 and p53 Arg72Pro polymorphism on age of tumour onset in Li-Fraumeni syndrome. *J Med Genet* **43**, 531-3 (2006).
27. Frohling, S. et al. Identification of driver and passenger mutations of FLT3 by high-throughput DNA sequence analysis and functional assessment of candidate alleles. *Cancer Cell* **12**, 501-13 (2007).
28. Aqeilan, R.I. et al. Targeted deletion of Wwox reveals a tumor suppressor function. *Proceedings of the National Academy of Sciences* **104**, 3949-3954 (2007).
29. Paige, A.J.W. et al. WWOX: A candidate tumor suppressor gene involved in multiple tumor types. *Proceedings of the National Academy of Sciences* **98**, 11417-11422 (2001).
30. Chang, C.L. et al. nm23-H1 mutation in neuroblastoma. **370**, 335-336 (1994).
31. Calin, G.A. et al. Familial Cancer Associated with a Polymorphism in ARLTS1. *N Engl J Med* **352**, 1667-1676 (2005).
32. Frank, B. et al. Association of the ARLTS1 Cys148Arg variant with familial breast cancer risk. *Int J Cancer* **118**, 2505-8 (2006).

A.**B.**

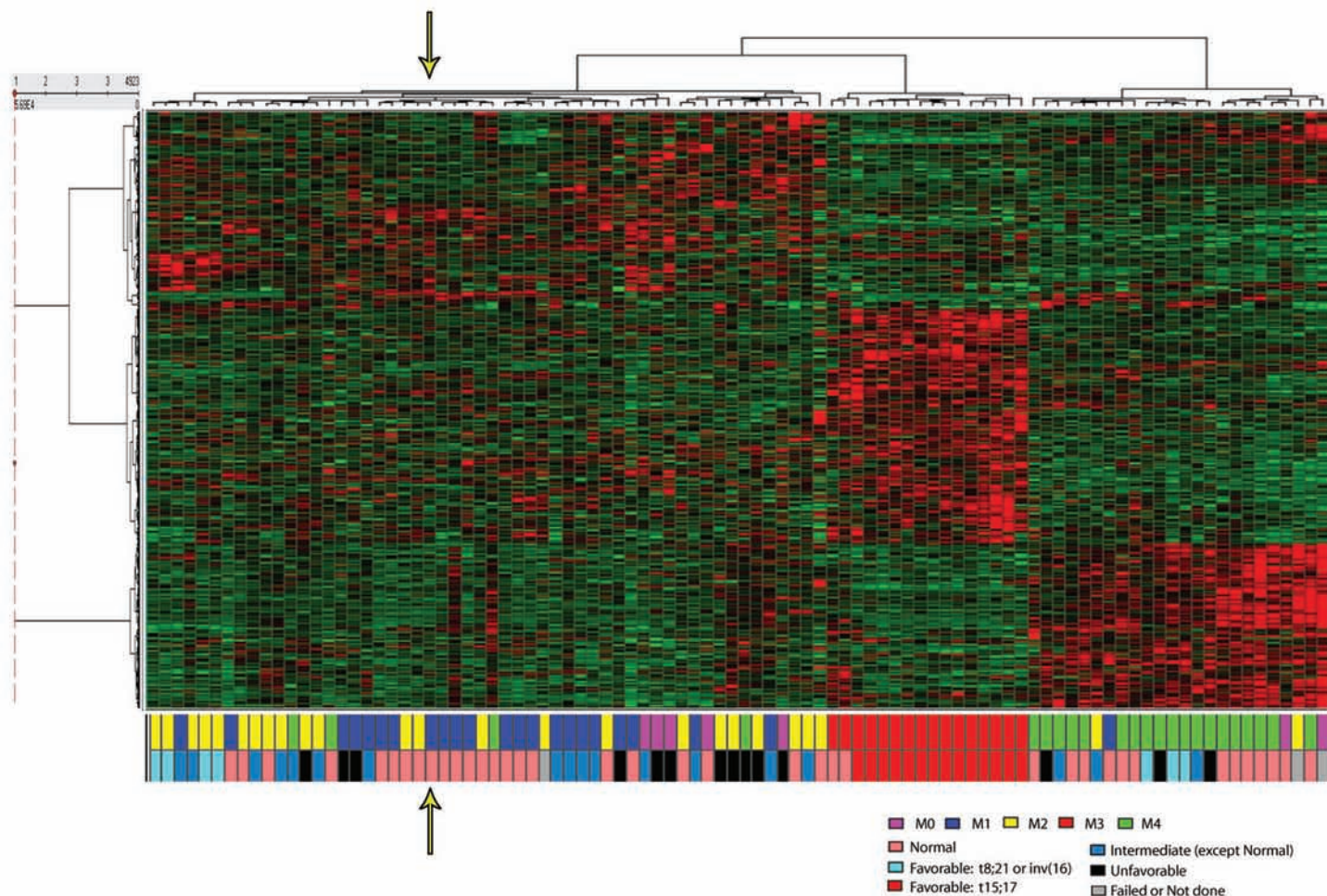
Supplementary Figure 1. Morphology and immunophenotype of AML in patient 933124.

(A) Bone marrow biopsy specimen demonstrates hypercellularity (>90%) with complete replacement (100/100 nucleated cells by manual differential) by a population of immature myeloblasts with folded nuclei and granular cytoplasm (inset). Original magnification, 10X (inset, 60X). (B) Flow cytometric analysis of peripheral blood demonstrates a blast population comprising 85.3% of nucleated cells with a uniform CD13+CD33+, CD177+CD34- immunophenotype. The blasts were variably myeloperoxidase positive and were negative for NBE (not shown).



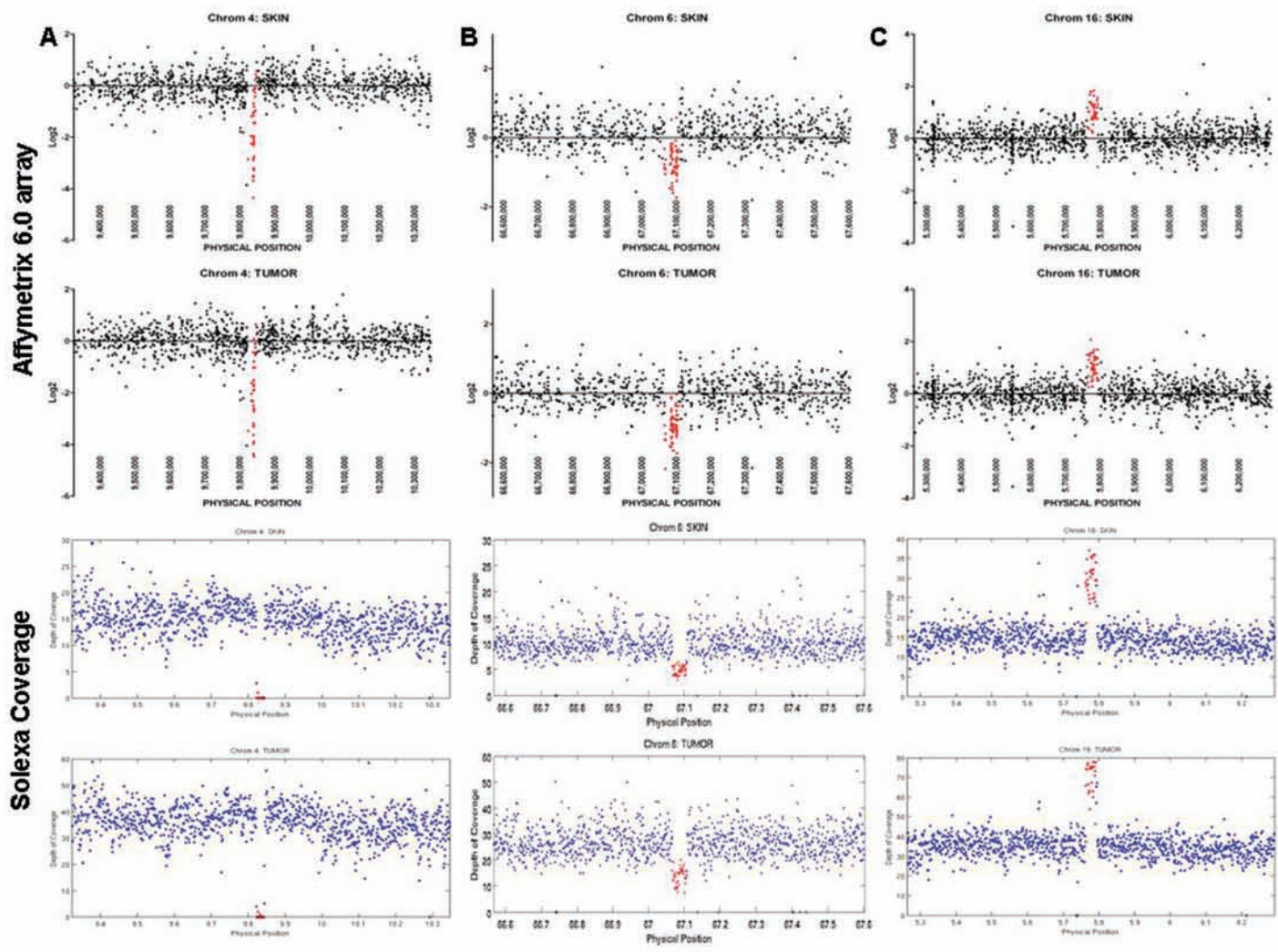
Supplementary Figure 2. Copy number and loss of heterozygosity heatmap for 18 M1 patients.

The results of copy number (CN) and loss of heterozygosity (LOH) analysis of 18 tumor-normal matched sample pairs assayed on Affymetrix SNP 6.0 arrays are shown. The index patient, 933124, is indicated by the bold font. For each sample pair, copy number as log2 ratio is shown on the left, LOH on the right. Copy number is designated by a color range from white to red, with deeper red indicating increasing copy number. LOH is shown in blue, while white indicates a neutral, or non-LOH state.



Supplementary Figure 3. Unsupervised clustering expression heatmap for 111 AML patients.

Expression data from 111 AML patients analyzed using unsupervised clustering (Ward's) method is shown in the heatmap; 92 of the 111 patients are from the Discovery Set of 94 AML cases from Washington University. The other 19 cases were also collected at Washington University using identical criteria for inclusion in the study. The range of Z-score-normalized expression is designated by a green to red color range, with red indicating increasing expression levels. FAB classes (top row under the heatmap) and major cytogenetic categories of the AML samples (bottom row under the heatmap) are shown; the legends for each are shown on the right. The index patient, 933124, is indicated by yellow arrows.



Supplementary Figure 4. Inherited copy number variants detected by Solexa read counts and Affymetrix 6.0 SNP arrays. (A) Homozygous deletion. (B) Heterozygous deletion. (C) Single copy gain.

Supplementary Table 2. MDS/AML candidate genes.

<u>Gene</u>	<u>GeneID</u>	<u>Gene</u>	<u>GeneID</u>	<u>Gene</u>	<u>GeneID</u>
ABL2	27	FLT3	2322	NUP214	8021
AF15Q14	57082	FNBP1	23048	NUP98	4928
ARHGEF12	23365	FUS	2521	PALB2	79728
ARNT	405	GAS7	8522	PDGFRA	5156
BAALC	79870	GATA1	2623	PDGFRB	5159
BLM	641	GFI1	2672	PER1	5187
BRAF	673	GMPS	8833	PICALM	8301
BRCA2	675	GRAF	23092	PML	5371
BRIP1	83990	HAX1	10456	PMX1	5396
CBFA2T1	862	HEAB	10978	PNUTL1	5413
CBFA2T3	863	HLXB9	3110	PRDM16	63976
CBFB	865	HOXA13	3209	PSIP2	11168
CBL	867	HOXA9	3205	PTEN	5728
CDKN2A	1029	HOXC11	3227	PTPN11	5781
CDX2	1045	HOXC13	3229	RARA	5914
CEBPA	1050	HOXD11	3237	RB1	5925
CHIC2	26511	HOXD13	3239	RBM15	64783
CREBBP	1387	JAK2	3717	RPL22	6146
CSF1R	1436	KIT	3815	RPN1	6184
CSF3R	1441	KRAS	3845	RPS19	6223
DDX10	1662	LASP1	3927	RPS24	6229
DEK	7913	LCX	80312	RUNX1	861
DKC1	1736	MDS1	4197	RUNXBP2	799
ELA2	1991	MDS2	259283	SBDS	51119
ELF4	2000	MKL1	57591	SEPT6	23157
EP300	2033	MLF1	4291	SET	6418
ERG	2078	MLL	4297	SMAD4	4089
ETV6	2120	MN1	4330	SOCS1	8651
EVI1	2122	MPL	4352	SPI1	6688
FACL6	23305	MSF	10801	SSH3BP1	10006
FANCA	2175	MYH11	4629	SYK	6850
FANCB	2187	MYST4	23522	TERC	7012
FANCC	2176	NCOA2	10499	TIF1	8805
FANCD2	2177	NF1	4763	TOP1	7150
FANCE	2178	NIN	51199	TP53	7157
FANCF	2188	NOTCH1	4851	TRIP11	9321
FANCG	2189	NPM1	4869	WHSC1L1	54904
FANCI	55215	NRAS	4893	WT1	7490
FANCL	55120	NSD1	64324	ZNF145	7704
FANCM	57697	NUMA1	4926		

Supplementary Table 3. Validated germline non-synonymous variants in MDS/AML genes.

Gene_name	Ensembl_transcript_id	Chromosome	Start_position (B36)	Reference_allele	Variant_allele	in dbSNP?	Variant_type	Amino_acid_change	OMIM
BRCA2	NM_000059	13	31804729	A	C	yes	missense	p.N372H	600185.0013
FLT3	ENST00000380982	13	27524716	C	T	no	missense	p.V194M	
PALB2	NM_024675	16	23545216	G	A	no	missense	p.P864S	
TP53	NM_000546	17	7520197	G	C	yes	missense	p.P72R	*191170.0005
TRIP11	ENST00000267622	14	91510761	G	A	no	missense	p.T1846I	

Supplementary Table 4. Predicted variants from 50k-bp of annotated 5' and 3' boundaries of MDSAM genes.

Reg. Symbol	Transcript Name	Chromosome	2007 Location	Variant Allele	# of tumor reads supporting variant allele	# of skin reads supporting variant allele	Reference Allele	# of tumor reads supporting the reference allele	# of skin reads supporting the reference allele	Variant Type	Amino Acid Change	Location in dMSF-127 (D: No. 1, Yes; Location in Watanabe (D: No. 1, Yes; Location in Venter (D: No. 1, Yes; Location in	Comment		
PRDM16	ENST00000270722	1	2973323 T		22	2	A	23	4	missense	Not Applicable	1	1	0	Unknown
PRDM16	ENST00000270722	1	2980464 C		9	1	T	11	0	missense	Not Applicable	1	1	1	0.001
PRDM16	ENST00000270722	1	2980493 T		18	5	C	10	0	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	2981760 A		14	10	G	28	12	missense	Not Applicable	1	0	0	0.002
PRDM16	ENST00000270722	1	2986175 G		39	4	T	0	0	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	2987335 A		45	11	G	0	0	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3000303 T		9	0	A	0	0	missense	Not Applicable	1	0	0	0.004
PRDM16	ENST00000270722	1	3000713 G		29	6	G	10	11	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3000829 C		25	9	A	22	3	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3001446 G		13	7	C	11	2	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3001489 C		38	20	T	0	0	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3001434 A		13	9	G	21	10	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3001286 C		11	9	T	12	7	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3001317 C		11	5	T	14	5	missense	Not Applicable	1	0	0	0.008
PRDM16	ENST00000270722	1	3001362 C		19	16	T	6	0	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3001463 A		15	10	G	14	1	missense	Not Applicable	1	1	1	0.001
PRDM16	ENST00000270722	1	3000895 A		9	15	G	12	11	missense	Not Applicable	1	1	1	0.002
PRDM16	ENST00000270722	1	3001264 C		26	3	G	10	5	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3000603 C		14	9	A	13	9	missense	Not Applicable	1	1	1	0.003
PRDM16	ENST00000270722	1	3000977 C		25	2	G	12	0	missense	Not Applicable	1	1	1	0.003
PRDM16	ENST00000270722	1	3000339 A		25	11	G	9	0	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3000977 C		25	11	G	9	0	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3000487 A		19	11	G	8	0	missense	Not Applicable	0	0	0	0.003
PRDM16	ENST00000270722	1	3001141 A		26	3	G	10	0	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3011167 T		18	10	C	7	0	missense	Not Applicable	1	0	0	0.001
PRDM16	ENST00000270722	1	3012962 A		14	12	G	14	10	missense	Not Applicable	1	0	0	0.013
PRDM16	ENST00000270722	1	3012863 C		33	2	A	11	0	missense	Not Applicable	1	0	0	0.002
PRDM16	ENST00000270722	1	3014235 C		53	23	A	0	0	missense	Not Applicable	1	0	0	0.01
PRDM16	ENST00000270722	1	3015706 A		18	10	T	4	0	missense	Not Applicable	1	0	0	0.014
PRDM16	ENST00000270722	1	3016120 A		13	12	T	22	7	missense	Not Applicable	1	0	0	0.002
PRDM16	ENST00000270722	1	3016844 C		14	12	G	17	18	missense	Not Applicable	1	0	0	0.002
PRDM16	ENST00000270722	1	3016888 G		12	4	A	24	18	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3016861 C		14	8	T	17	18	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3017082 G		32	11	A	9	0	missense	Not Applicable	1	0	0	0.008
PRDM16	ENST00000270722	1	3017182 A		30	14	G	19	10	missense	Not Applicable	1	0	0	0.001
PRDM16	ENST00000270722	1	3022425 C		18	4	A	16	9	missense	Not Applicable	1	0	0	0.001
PRDM16	ENST00000270722	1	3022494 T		18	7	C	16	10	missense	Not Applicable	1	0	0	0.042
PRDM16	ENST00000270722	1	3024666 G		21	14	G	25	0	missense	Not Applicable	1	0	0	0.008
PRDM16	ENST00000270722	1	3025067 T		13	11	A	0	0	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3026360 C		17	7	T	9	0	missense	Not Applicable	1	0	0	0.004
PRDM16	ENST00000270722	1	3026329 G		18	10	A	20	10	missense	Not Applicable	1	0	0	0.001
PRDM16	ENST00000270722	1	3026475 C		44	26	T	19	17	missense	Not Applicable	1	1	1	Unknown
PRDM16	ENST00000270722	1	3026971 C		19	10	A	23	14	missense	Not Applicable	1	0	0	0.001
PRDM16	ENST00000270722	1	3026323 C		31	6	T	49	6	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3027052 G		17	11	G	2	0	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3027312 G		17	10	A	31	6	missense	Not Applicable	1	1	1	Unknown
PRDM16	ENST00000270722	1	3027977 T		17	11	G	29	3	missense	Not Applicable	1	1	1	Unknown
PRDM16	ENST00000270722	1	3044206 A		33	25	G	16	12	missense	Not Applicable	0	0	0	0.001
PRDM16	ENST00000270722	1	3044969 A		18	4	C	10	0	missense	Not Applicable	0	0	0	0.001
PRDM16	ENST00000270722	1	3047946 C		16	10	T	9	4	missense	Not Applicable	0	0	0	0.007
PRDM16	ENST00000270722	1	3048220 A		18	12	G	6	0	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3054438 T		10	10	G	11	0	missense	Not Applicable	1	0	0	0.002
PRDM16	ENST00000270722	1	3057950 A		16	18	C	9	0	missense	Not Applicable	1	0	0	0.002
PRDM16	ENST00000270722	1	3058077 G		18	10	G	18	0	missense	Not Applicable	1	0	0	0.002
PRDM16	ENST00000270722	1	3058433 C		42	7	T	18	10	missense	Not Applicable	1	1	1	Unknown
PRDM16	ENST00000270722	1	3058968 G		26	8	A	17	13	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3071038 T		25	6	C	6	0	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3071211 C		19	9	A	24	10	missense	Not Applicable	1	0	0	0.005
PRDM16	ENST00000270722	1	3071189 C		15	23	T	9	0	missense	Not Applicable	1	1	1	Unknown
PRDM16	ENST00000270722	1	3071795 T		13	10	C	10	18	missense	Not Applicable	1	0	0	0.002
PRDM16	ENST00000270722	1	3078022 T		46	9	C	0	0	missense	Not Applicable	0	1	1	Unknown
PRDM16	ENST00000270722	1	3078022 C		19	12	C	1	0	missense	Not Applicable	0	1	1	Unknown
PRDM16	ENST00000270722	1	3080750 A		40	10	A	25	0	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3080424 A		18	14	G	9	13	missense	Not Applicable	0	0	0	0.003
PRDM16	ENST00000270722	1	3080838 C		39	7	T	0	0	missense	Not Applicable	0	1	1	0.01
PRDM16	ENST00000270722	1	3080868 T		11	3	C	11	0	missense	Not Applicable	0	1	1	0.002
PRDM16	ENST00000270722	1	3087168 A		14	36	G	308	49	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3087772 A		19	35	G	39	20	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3087833 A		22	9	C	17	0	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3087972 G		18	22	A	20	18	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3088068 A		29	6	G	79	44	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3088068 G		29	25	G	18	10	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3088031 T		38	14	C	56	31	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3088031 T		22	33	G	21	13	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3088489 T		23	16	C	35	27	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3088141 C		46	11	C	0	0	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3088161 C		13	17	C	0	0	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3088495 T		47	8	C	26	5	missense	Not Applicable	0	0	0	0.021
PRDM16	ENST00000270722	1	3089261 A		19	10	A	21	0	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3089434 A		36	9	G	19	11	missense	Not Applicable	1	1	1	Unknown
PRDM16	ENST00000270722	1	3089438 G		23	7	T	17	0	missense	Not Applicable	1	1	1	Unknown
PRDM16	ENST00000270722	1	3089422 T		24	6	G	17	11	missense	Not Applicable	1	1	1	Unknown
PRDM16	ENST00000270722	1	3089465 G		51	21	A	0	0	missense	Not Applicable	1	0	0	0.001
PRDM16	ENST00000270722	1	3089276 G		51	0	T	0	0	missense	Not Applicable	0	1	1	Unknown
PRDM16	ENST00000270722	1	3089277 A		18	16	G	17	0	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3100563 G		51	11	C	0	0	missense	Not Applicable	1	1	1	0.007
PRDM16	ENST00000270722	1	3100543 A		13	12	G	18	0	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3101468 G		19	13	A	21	10	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3101760 T		19	6	C	9	0	missense	Not Applicable	0	0	0	Unknown
PRDM16	ENST00000270722	1	3101133 C		54	27	T	0	0	missense	Not Applicable	1	1	1	0.001
PRDM16	ENST00000270722	1	3101463 G		13	10	G	14	0	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3101962 A		13	16	A	24	3	missense	Not Applicable	1	0	0	0.022
PRDM16	ENST00000270722	1	3101968 G		13	10	G	18	0	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3106776 G		23	8	A	6	0	missense	Not Applicable	1	0	0	Unknown
PRDM16	ENST00000270722	1	3106747 C		19	18	G	18	0	missense	Not Applicable	1	0	0	0.002
PRDM16	ENST00000270722	1	3108752 G												

www.nature.com/nature

ETV6	NA_001887	12	11751437	G	22	6	C	10	4	nonc	Not Applicable	1	0	1	0.222
ETV6	NA_001887	12	11751462	G	26	6	A	17	2	nonc	Not Applicable	1	0	1	Unknown
ETV6	NA_001887	12	11751521	G	27	10	A	18	6	nonc	Not Applicable	1	0	1	0.404
ETV6	NA_001887	12	11751703	T	19	12	G	21	17	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11751840	A	24	5	C	26	7	nonc	Not Applicable	1	0	1	0.001
ETV6	NA_001887	12	11751884	T	11	6	C	29	11	nonc	Not Applicable	1	0	1	Unknown
ETV6	NA_001887	12	11751935	G	11	5	A	31	9	nonc	Not Applicable	1	0	1	Unknown
ETV6	NA_001887	12	11760020	T	47	15	C	0	0	nonc	Not Applicable	1	1	1	0.004
ETV6	NA_001887	12	11760024	G	47	12	A	0	0	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11760460	A	49	17	C	0	0	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11760597	G	49	8	A	0	0	nonc	Not Applicable	1	1	1	0.002
ETV6	NA_001887	12	11760602	T	22	10	C	14	18	nonc	Not Applicable	1	1	0	0.001
ETV6	NA_001887	12	11771240	G	41	6	A	1	6	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11773464	T	13	12	C	32	6	nonc	Not Applicable	1	0	1	0.001
ETV6	NA_001887	12	11775960	G	13	11	A	21	10	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11782094	A	3	6	G	7	6	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11784295	T	20	8	G	11	13	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11787348	T	10	13	A	12	12	nonc	Not Applicable	1	0	1	0.001
ETV6	NA_001887	12	11788071	T	52	13	C	0	0	nonc	Not Applicable	1	1	1	0.126
ETV6	NA_001887	12	11788621	T	11	3	C	21	0	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11794466	G	39	20	A	0	0	nonc	Not Applicable	1	1	1	0.004
ETV6	NA_001887	12	11797071	C	51	5	G	23	2	nonc	Not Applicable	1	1	1	0.027
ETV6	NA_001887	12	11798013	T	18	10	A	22	5	nonc	Not Applicable	1	0	1	Unknown
ETV6	NA_001887	12	11798037	T	20	5	A	20	1	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11798200	G	11	3	A	19	17	nonc	Not Applicable	1	1	1	0.011
ETV6	NA_001887	12	11798686	C	28	6	T	44	1	nonc	Not Applicable	1	1	0	Unknown
ETV6	NA_001887	12	11798807	G	44	12	G	0	0	nonc	Not Applicable	1	1	1	0.001
ETV6	NA_001887	12	11799021	T	23	9	C	15	2	nonc	Not Applicable	1	1	1	0.001
ETV6	NA_001887	12	11799693	A	76	22	G	0	0	nonc	Not Applicable	1	1	1	0.001
ETV6	NA_001887	12	11799796	A	19	6	G	46	1	nonc	Not Applicable	1	1	1	0.001
ETV6	NA_001887	12	11803171	C	17	7	T	14	10	nonc	Not Applicable	1	1	1	0.001
ETV6	NA_001887	12	11804435	T	13	10	C	17	9	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11805106	T	10	1	A	8	4	nonc	Not Applicable	1	1	1	0.017
ETV6	NA_001887	12	11805304	A	19	4	T	31	0	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11818777	G	68	25	A	0	0	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11820034	C	17	1	T	18	4	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11825420	A	13	7	G	21	5	nonc	Not Applicable	1	1	1	0.025
ETV6	NA_001887	12	11827236	T	17	10	A	12	8	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11829295	C	28	21	G	15	4	nonc	Not Applicable	1	1	0	0.217
ETV6	NA_001887	12	11829747	T	42	10	T	0	10	nonc	Not Applicable	1	1	1	0.001
ETV6	NA_001887	12	11830461	T	11	19	A	11	0	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11831792	C	11	24	G	33	10	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11833044	T	65	15	C	0	0	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11833611	G	4	10	A	44	0	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11833979	A	38	9	T	0	0	nonc	Not Applicable	1	1	1	0.028
ETV6	NA_001887	12	11835065	G	12	10	A	1	1	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11835743	C	27	1	G	31	13	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11836117	C	12	11	T	10	7	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11836216	G	41	7	A	15	10	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11836961	C	12	10	A	12	20	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11837286	G	25	9	T	13	12	nonc	Not Applicable	1	1	1	0.007
ETV6	NA_001887	12	11837290	C	12	4	A	14	4	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11838037	C	29	8	T	13	12	nonc	Not Applicable	1	1	0	0.001
ETV6	NA_001887	12	11838210	T	46	2	C	2	3	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11840143	C	46	29	T	0	0	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11847400	G	47	7	A	17	17	nonc	Not Applicable	1	1	1	0.006
ETV6	NA_001887	12	11875654	A	18	0	G	25	13	nonc	Not Applicable	1	0	0	0.002
ETV6	NA_001887	12	11875656	A	18	0	G	25	13	nonc	Not Applicable	1	0	0	0.004
ETV6	NA_001887	12	11875659	A	38	0	G	1	1	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11875774	A	0	2	T	0	2	nonc	Not Applicable	1	1	1	0.001
ETV6	NA_001887	12	11880111	A	35	10	T	14	9	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11880112	A	35	8	A	0	0	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11880796	G	33	10	C	3	3	nonc	Not Applicable	1	1	0	Unknown
ETV6	NA_001887	12	11881323	A	34	10	G	12	4	nonc	Not Applicable	1	1	0	Unknown
ETV6	NA_001887	12	11881511	T	19	8	T	12	0	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11882003	T	19	6	C	12	nonc	Not Applicable	Not Applicable	1	1	0	0.206
ETV6	NA_001887	12	11884829	A	11	9	C	2	6	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11885275	A	45	17	G	1	1	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11885608	C	2	10	A	2	0	nonc	Not Applicable	1	1	1	0.001
ETV6	NA_001887	12	11887232	C	16	1	A	23	9	nonc	Not Applicable	1	1	1	0.004
ETV6	NA_001887	12	11888664	T	13	6	C	7	6	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11892897	C	39	2	G	0	0	nonc	Not Applicable	1	1	1	0.063
ETV6	NA_001887	12	11893211	A	6	17	A	16	17	nonc	Not Applicable	1	1	1	Unknown
ETV6	NA_001887	12	11893726	G	47	23	T	0	0	nonc	Not Applicable	1	1	1	0.014
ETV6	NA_001887	12	11893864	C	12	10	A	2	0	nonc	Not Applicable	1	1	1	0.003
ETV6	NA_001887	12	11893863	G	43	13	A	0	0	3_prime_untranslated_region	Not Applicable	1	0	0	0.002
ETV6	NA_001887	12	11893906	C	42	10	T	0	0	3_prime_untranslated_region	Not Applicable	1	0	0	0.002
KRAS	NA_012000	12	11893946	A	27	4	C	20	1	nonc	Not Applicable	1	1	1	0.001
KRAS	NA_012000	12	26257613	T	13	1	G	12	1	3_prime_untranslated_region	Not Applicable	1	1	1	0.773
KRAS	NA_012000	12	26257644	G	30	0	A	17	4	3_prime_untranslated_region	Not Applicable	1	1	1	Unknown
KRAS	NA_012000	12	26257696	A	30	7	C	1	7	nonc	Not Applicable	1	1	1	Unknown
KRAS	NA_012000	12	26271020	G	20	6	C	6	6	nonc	Not Applicable	1	1	1	Unknown
KRAS	NA_012000	12	26271292	A	14	1	C	1	1	nonc	Not Applicable	1	1	1	0.002
KRAS	NA_012000	12	26281184	C	25	5	G	10	1	nonc	Not Applicable	1	0	0	0.025
KRAS	NA_012000	12	26281995	A	12	4	C	12	4	nonc	Not Applicable	1	1	0	0.007
KRAS	NA_012000	12	26283710	T	12	5	C	26	3	nonc	Not Applicable	1	1	0	Unknown
KRAS	NA_012000	12	26290466	T	26	1	A	1	1	nonc	Not Applicable	1	1	1	0.001
KRAS	NA_012000	12	26297968	C	15	5	T	9	3	nonc	Not Applicable	1	1	0	0.001
KRAS	NA_012000	12	26297969	A	23	4	T	6	6	nonc	Not Applicable	1	1	0	Unknown
HDC13	NA_011410	12	50210308	A	19	10	G	0	0	nonc	Not Applicable	1	1	0	0.235
HDC13	NA_011410	12	50210309	G	38	18	C	1	1	nonc	Not Applicable	1	1	1	0.019
HDC13	NA_011410	12	50212360	G	10	1	A	1	1	nonc	Not Applicable	1	1	1	Unknown
HDC13	NA_011410	12	50212444	A	47	14	C	0	0	3_prime_untranslated_region	Not Applicable	1	1	1	Unknown
PTF1L1	NA_001834	12	11141284	C	3	11	A	38	1	nonc	Not Applicable	1	1	1	Unknown
PTF1L1	NA_001834	12	11141880	T	16	9	A	25	11	nonc	Not Applicable	0	0	0	Unknown
PTF1L1	NA_001834	12	11142064	A	11	10	A	13	10	nonc	Not Applicable	0	0	0	Unknown
PTF1L1	NA_001834	12	11147780	G	14	7	A	2	2	nonc	Not Applicable	1	0	0	0.061
C22C2	ENY70000000000020	12	27442617	T	44	9	G	0	0	p.D205	Not Applicable	1	0	0	0.001
C22C2	ENY70000000000020	12	27442617	T	44	25	C	1	0	nonc	Not Applicable	1	0	0	Unknown
C22C2	ENY70000000000020	12	27442618	G	19	1	C	0	0	nonc	Not Applicable	1	0	0	Unknown
FLT3	ENY70000000000002	13	27446120												

NRV	ENV70000004544	14	3035759 G	57	12 A	0	0	0	Not Applicable	1	0	1	0.001
NRV	ENV70000004544	14	3035824 C	30	9 G	0	5	0	Not Applicable	0	0	0	0.043
NRV	ENV70000004544	14	3035824 T	30	5 C	13	12	0	Not Applicable	1	0	0	0.005
TRIP1	ENV70000007622	14	8533338 G	11	4 T	1	8	0	Not Applicable	1	1	0	0.005
TRIP1	ENV70000007622	14	8533333 T	34	11 A	34	11	0	Not Applicable	1	0	0	0.005
TRIP1	ENV70000007622	14	8533376 G	35	5 T	3	9	0	Not Applicable	0	0	0	0.001
TRIP1	ENV70000007622	14	8533373 C	11	30 C	1	30	0	Not Applicable	0	0	0	0.001
TRIP1	ENV70000007622	14	8533290 C	18	2 T	23	2	0	Not Applicable	1	1	0	0.005
TRIP1	ENV70000007622	14	8533340 A	32	3 C	32	3	0	Not Applicable	1	1	0	0.005
TRIP1	ENV70000007622	14	8533354 G	10	7 C	7	2	0	Not Applicable	1	0	0	0.004
TRIP1	ENV70000007622	14	8533352 C	11	14 A	11	14	0	Not Applicable	1	1	0	0.002
TRIP1	ENV70000007622	14	8533351 C	30	7 T	20	2	0	Not Applicable	1	0	0	0.01
CASC5	NA_175589	15	3895955 A	40	10 C	40	10	0	Not Applicable	1	1	0	0.001
CASC5	NA_175589	15	3895792 T	46	16 C	0	0	0	Not Applicable	1	1	1	0.001
CASC5	NA_175589	15	3895988 T	42	12 C	1	0	0	Not Applicable	1	1	1	0.002
CASC5	NA_175589	15	3895924 G	45	14 A	0	0	0	Not Applicable	1	1	1	0.002
CASC5	NA_175589	15	3895933 C	68	25 T	0	0	0	Not Applicable	1	1	1	0.002
CASC5	NA_175589	15	3895112 T	50	13 G	1	0	0	Not Applicable	1	1	1	0.002
CASC5	NA_175589	15	3895110 C	59	14 T	0	0	0	Not Applicable	1	1	1	0.002
CASC5	NA_175589	15	3895227 T	44	10 C	0	0	0	Not Applicable	1	1	1	0.002
CASC5	NA_175589	15	3895338 C	40	9 T	0	0	0	Not Applicable	1	1	1	0.019
CASC5	NA_175589	15	3895103 T	44	14 G	1	0	0	Not Applicable	1	1	1	0.002
CASC5	NA_175589	15	3895136 C	54	30 T	0	0	0	Not Applicable	1	1	1	0.009
PNL	NA_033328	15	7336480 T	17	7 C	0	0	0	Not Applicable	1	0	1	0.004
PNL	NA_033328	15	7336482 G	46	18 A	0	0	0	Not Applicable	1	0	0	0.002
PNL	NA_033328	15	7336484 A	62	18 T	0	0	0	Not Applicable	1	0	0	0.002
PNL	NA_033328	15	7336482 A	46	16 C	0	1	0	Not Applicable	1	0	1	0.001
PNL	NA_033328	15	7336488 G	49	17 A	0	0	0	Not Applicable	1	0	0	0.001
PNL	NA_033328	15	7336483 A	49	17 A	0	0	0	Not Applicable	1	0	0	0.001
PNL	NA_033328	15	7336483 G	49	17 A	0	0	0	Not Applicable	1	0	0	0.001
PNL	NA_033328	15	7336484 G	47	18 C	0	1	0	Not Applicable	1	0	1	0.001
PNL	NA_033328	15	7336485 G	44	9 C	0	0	0	Not Applicable	1	1	1	0.001
PNL	NA_033328	15	7337352 G	46	36 A	0	0	0	Not Applicable	1	0	1	0.001
PNL	NA_033328	15	7337378 A	38	26 G	0	0	0	Not Applicable	1	0	0	0.001
PNL	NA_033328	15	7337367 A	52	35 T	0	0	0	Not Applicable	1	0	1	0.001
PNL	NA_033328	15	7337374 T	52	5 C	20	9	0	Not Applicable	1	0	0	0.006
PNL	NA_033328	15	7337428 C	13	20 T	16	7	0	Not Applicable	1	0	1	0.008
PNL	NA_033328	15	7337335 C	13	20 T	0	0	0	Not Applicable	1	0	0	0.008
PNL	NA_033328	15	7337244 T	59	16 C	0	0	0	Not Applicable	1	1	1	0.002
PNL	NA_033328	15	7337255 T	46	20 T	0	0	0	Not Applicable	1	1	1	0.002
FANCJ	NA_033328	15	8722222 G	11	10 A	0	0	0	Not Applicable	1	0	0	0.002
FANCJ	NA_033328	15	8722225 T	63	7 C	0	0	0	Not Applicable	1	0	0	0.024
BLM	NA_090007	15	8602781 T	1	5 T	8	0	0	Not Applicable	1	1	0	0.002
BLM	NA_090007	15	8602894 C	17	4 T	24	12	0	Not Applicable	1	0	0	0.018
BLM	NA_090007	15	8602894 T	18	5 C	18	5	0	Not Applicable	1	0	0	0.018
BLM	NA_090007	15	8613328 A	35	4 G	10	0	0	Not Applicable	1	0	0	0.02
BLM	NA_090007	15	8613347 T	33	12 C	13	0	0	Not Applicable	1	0	0	0.027
BLM	NA_090007	15	8613249 G	15	4 A	4	0	0	Not Applicable	1	1	0	0.145
BLM	NA_090007	15	8613278 T	39	5 C	18	0	0	Not Applicable	1	0	0	0.187
BLM	NA_090007	15	8613274 T	32	2 C	27	2	0	Not Applicable	1	1	0	0.005
BLM	NA_090007	15	8613223 T	31	1 C	31	0	0	Not Applicable	1	0	0	0.007
BLM	NA_090007	15	8613301 T	35	3 C	2	0	0	Not Applicable	1	0	0	0.005
BLM	NA_090007	15	8613301 G	32	6 A	21	0	0	Not Applicable	1	0	0	0.005
BLM	NA_090007	15	8613761 C	26	5 G	15	6	0	Not Applicable	1	0	0	0.002
BLM	NA_090007	15	8613763 T	11	9 C	14	1	0	Not Applicable	1	0	0	0.001
BLM	NA_090007	15	8613761 G	17	4 A	11	3	0	Not Applicable	1	1	0	0.005
BLM	NA_090007	15	8613834 C	35	8 C	11	0	0	Not Applicable	1	0	0	0.001
BLM	NA_090007	15	8614015 A	20	5 G	11	7	0	Not Applicable	1	0	0	0.089
BLM	NA_090007	15	8614235 A	24	5 C	27	14	0	Not Applicable	1	0	0	0.004
BLM	NA_090007	15	8614460 T	13	2 T	9	2	0	Not Applicable	1	0	0	0.013
BLM	NA_090007	15	8614575 T	14	3 A	17	5	0	Not Applicable	1	0	0	0.005
BLM	NA_090007	15	8614789 A	14	10 G	10	6	0	Not Applicable	1	0	0	0.005
BLM	NA_090007	15	8615351 A	11	8 C	22	1	0	Not Applicable	1	0	0	0.014
BLM	NA_090007	15	8615417 C	11	7 C	18	13	0	Not Applicable	1	0	0	0.018
BLM	NA_090007	15	8615428 C	15	2 C	7	0	0	Not Applicable	1	0	0	0.003
BLM	NA_090007	15	8615431 C	58	15 T	6	0	0	Not Applicable	1	0	0	0.003
BLM	NA_090007	15	8615433 C	59	17 G	13	0	0	Not Applicable	1	0	0	0.003
BLM	NA_090007	15	8615437 A	17	5 G	12	8	0	Not Applicable	1	0	0	0.005
CHERP	ENV70000003070	16	3178448 C	1	27 C	1	0	0	Not Applicable	1	0	0	0.005
CHERP	ENV70000003070	16	3178485 C	27	10 G	6	0	0	Not Applicable	1	0	0	0.009
CHERP	ENV70000003070	16	3178508 G	18	14 C	12	0	0	Not Applicable	1	0	0	0.005
CHERP	ENV70000003070	16	3893303 T	35	8 C	21	4	0	Not Applicable	1	0	0	0.005
CHERP	ENV70000003070	16	3893329 C	30	7 T	0	0	0	Not Applicable	1	0	0	0.005
CHERP	ENV70000003070	16	3893365 A	41	7 A	0	0	0	Not Applicable	1	1	1	0.005
CHERP	ENV70000003070	16	3893321 G	33	12 A	0	0	0	Not Applicable	1	0	0	0.005
CHERP	ENV70000003070	16	3897064 C	84	24 G	0	0	0	Not Applicable	1	1	1	0.005
MYH1	NA_05146114	16	2573583 T	42	10 C	1	0	0	Not Applicable	1	0	0	0.005
MYH1	NA_05146114	16	2573582 T	42	11 C	1	0	0	Not Applicable	1	0	0	0.005
MYH1	NA_05146114	16	2573124 A	70	29 G	1	0	0	Not Applicable	1	0	1	0.005
MYH1	NA_05146114	16	2573124 T	70	29 G	1	0	0	Not Applicable	1	0	1	0.005
MYH1	NA_05146114	16	2573562 C	68	9 A	0	0	0	Not Applicable	1	0	0	0.017
MYH1	NA_05146114	16	2573571 C	68	9 A	0	0	0	Not Applicable	1	0	0	0.017
MYH1	NA_05146114	16	2573571 G	68	9 A	0	0	0	Not Applicable	1	0	0	0.017
MYH1	NA_05146114	16	2573571 T	68	9 A	0	0	0	Not Applicable	1	0	0	0.017
MYH1	NA_05146114	16	2573571 A	13	9 G	18	13	0	Not Applicable	1	0	0	0.002
MYH1	NA_05146114	16	2573727 G	50	13 A	1	0	0	Not Applicable	1	1	0	0.005
MYH1	NA_05146114	16	2573705 C	46	11 C	0	0	0	Not Applicable	1	0	0	0.005
MYH1	NA_05146114	16	2573787 T	46	12 C	1	0	0	Not Applicable	1	0	0	0.005
MYH1	NA_05146114	16	2573782 C	17	8 C	10	8	0	Not Applicable	1	0	0	0.005
MYH1	NA_05146114	16	2573789 A	12	7 G	10	8	0	Not Applicable	1	0	0	0.005
MYH1	NA_05146114	16	2573788 T	17	8 C	10	8	0	Not Applicable	1	0	0	0.005
MYH1	NA_05146114	16	2573788 A	16	6 G	4	0	0	Not Applicable	1	0	0	0.001
MYH1	NA_05146114	16	2573788 T	14	12 C	12	4	0	Not Applicable	1	0	0	0.001
MYH1	NA_05146114	16	2573788 A	25	6 C	26	2	0	Not Applicable	1	0	0	0.003
MYH1	NA_05146114	16	2573822 G	8	9 C	14	12	0	Not Applicable	1	0	0	0.003
MYH1	NA_05146114	16	2573828 T	30	11 G	17	13	0	Not Applicable	1	0	0	0.005
MYH1	NA_05146114	16	2573828 A	30	11 G	17	13	0	Not Applicable	1	0	0	0.005
MYH1	NA_05146114	16	2573828 T	13	19 G	25	0	0	Not Applicable	1	0	0	0.001
MYH1	NA_05146114	16	2573828 A	13	19 G	25	0	0	Not Applicable	1	0	0	0.001
MYH1	NA_05146114	16	2573828 T	16	10 C	16	7	0	Not Applicable	1	0	0	0.005
MYH1	NA_05146114	16	2573828 A	16	10 C	16	7	0	Not Applicable	1	0	0	0.005
MYH1	NA_05146114	16	2573828 T	50	6 T	0	0	0	Not Applicable	1	1	1	0.001
MYH1	NA_05146114	16	2573828 A	50	6 T	0	0	0	Not Applicable	1	1	1	0.001
MYH1	NA_05146114	16	2573828 T	39	8 C	17	5	0	Not Applicable	1	0	0	0.006
MYH1	NA_05146114	16	2573828 A	39	8 C	17	5	0	Not Applicable	1	0	0	0.006
MYH1	NA_05146114	16	2573828 T	11	12 G	14	9	0	Not Applicable	1	0	0	0.005
MYH1	NA_05146114	16	2573828 A	11	12 G	14	9	0	Not Applicable	1	0	0	0.005
MY													

NP1	ENV70000058273	17	2617028 A	24	6 G	21	11	Unknown	1	1	1	Unknown
NP1	ENV70000058273	17	2617272 A	12	6 G	16	6	Unknown	1	1	0	0.043
NP1	ENV70000058273	17	26173246 G	29	7 G	17	2	Unknown	1	1	0	0.002
NP1	ENV70000058273	17	26177438 C	11	6 T	16	7	Unknown	1	1	1	Unknown
NP1	ENV70000058273	17	26180665 G	15	4 A	15	3	Unknown	1	1	1	Unknown
NP1	ENV70000058273	17	26186453 A	10	1	4	Unknown	1	1	1	1	Unknown
NP1	ENV70000058273	17	26178975 T	9	5 C	9	5	Unknown	1	1	1	Unknown
NP1	ENV70000058273	17	26174506 G	23	3 A	20	6	Unknown	1	1	1	0.1
LASP1	ENV70000058348	17	34203086 T	10	7 C	20	10	Unknown	1	1	1	0.12
LASP1	ENV70000058348	17	34203083 C	67	3 F T	1	0	Unknown	1	1	1	0.001
LASP1	ENV70000058348	17	34203088 G	8	10 A	1	10	Unknown	1	1	1	0.001
LASP1	ENV70000058348	17	34203097 C	10	8 T	21	10	Unknown	1	1	1	0.001
LASP1	ENV70000058348	17	34203094 C	4	5 A	18	14	Unknown	1	1	1	0.001
LASP1	ENV70000058348	17	34203207 C	23	4 G	26	5	Unknown	1	1	0	0.003
LASP1	ENV70000058348	17	34203168 C	12	15 A	15	4	Unknown	1	1	0	0.003
LASP1	ENV70000058348	17	34203222 T	12	5 C	15	8	Unknown	1	1	0	0.001
LASP1	ENV70000058348	17	34203075 A	38	6 G	22	3	Unknown	1	1	0	0.001
LASP1	ENV70000058348	17	34203222 A	12	7 G	9	9	Unknown	1	1	0	0.002
LASP1	ENV70000058348	17	34203207 C	25	13 A	14	14	Unknown	1	1	0	Unknown
LASP1	ENV70000058348	17	34203161 C	14	8 A	14	14	Unknown	1	1	0	Unknown
LASP1	ENV70000058348	17	34203247 G	12	15 A	21	10	Unknown	0	0	0	Unknown
LASP1	ENV70000058348	17	34203051 G	40	18 A	0	5	Unknown	1	1	0	Unknown
LASP1	ENV70000058348	17	34203042 G	12	8 A	10	12	Unknown	1	1	0	0.003
LASP1	ENV70000058348	17	34203425 G	38	4 A	16	10	Unknown	1	1	0	Unknown
LASP1	ENV70000058348	17	34203057 C	46	16 T	0	0	Unknown	1	1	1	Unknown
LASP1	ENV70000058348	17	34207748 A	11	6 G	22	1	Unknown	1	1	0	0.002
LASP1	ENV70000058348	17	34203026 T	12	7 C	18	8	Unknown	1	1	0	0.008
LASP1	ENV70000058348	17	34203425 G	21	9 A	25	2	Unknown	1	1	0	0.008
LASP1	ENV70000058348	17	34203462 T	22	4 A	8	9	Unknown	0	0	0	0.001
LASP1	ENV70000058348	17	34203076 G	11	3 A	1	1	Unknown	1	1	1	0.001
LASP1	ENV70000058348	17	34211863 G	26	5 C	9	7	Unknown	1	1	1	0.414
LASP1	ENV70000058348	17	34210875 A	44	11 G	0	0	Unknown	1	1	1	0.002
LASP1	ENV70000058348	17	34210876 G	25	10 A	54	5	Unknown	1	1	1	0.003
LASP1	ENV70000058348	17	34210866 T	12	5 C	20	2	Unknown	1	1	1	0.008
LASP1	ENV70000058348	17	34213855 A	26	9 G	12	6	Unknown	1	1	1	Unknown
LASP1	ENV70000058348	17	34213856 G	34	17 A	29	4	Unknown	1	1	1	0.006
LASP1	ENV70000058348	17	34213862 T	27	10 G	19	9	Unknown	1	1	1	0.007
LASP1	ENV70000058348	17	34213863 G	22	8 A	12	6	Unknown	1	1	1	0.015
LASP1	ENV70000058348	17	34213864 T	12	4 A	10	10	Unknown	1	1	1	Unknown
LASP1	ENV70000058348	17	34213865 T	22	7 A	32	10	Unknown	1	1	0	Unknown
LASP1	ENV70000058348	17	34213870 G	20	6 G	28	14	Unknown	1	1	1	Unknown
LASP1	ENV70000058348	17	34213871 G	10	15 A	14	12	Unknown	1	1	1	Unknown
LASP1	ENV70000058348	17	34213872 A	11	4 G	11	12	Unknown	1	1	0	0.001
RNA	ENV7000000315169	17	30172345 A	11	4 G	11	12	Unknown	1	1	0	0.001
RNA	ENV7000000315169	17	30172346 A	11	4 G	11	12	Unknown	1	1	0	0.001
BMPI1	ENV7000000315169	17	30172346 T	30	8 A	62	13	Unknown	0	0	0	0.182
BMPI1	NA_013043	17	57123778 G	7	8 A	1	1	Unknown	0	0	0	0.001
BMPI1	NA_013043	17	57123783 G	46	9 G	0	0	Unknown	1	1	1	Unknown
BMPI1	NA_013043	17	57123788 C	17	11 A	17	4	Unknown	1	1	1	0.004
BMPI1	NA_013043	17	57123791 T	10	5 G	23	4	Unknown	1	1	0	0.003
BMPI1	NA_013043	17	57123807 G	19	6 A	19	13	Unknown	1	1	0	0.043
BMPI1	NA_013043	17	57123808 C	11	13 T	13	15	Unknown	1	1	1	0.004
BMPI1	NA_013043	17	57123822 G	8	12 T	17	4	Unknown	1	1	1	Unknown
BMPI1	NA_013043	17	57123823 T	18	4 C	14	8	Unknown	1	1	1	Unknown
BMPI1	NA_013043	17	57123794 A	13	7 C	16	1	Unknown	1	1	0	0.007
BMPI1	NA_013043	17	57123804 T	35	2 A	8	8	Unknown	1	1	1	Unknown
BMPI1	NA_013043	17	57123885 G	14	9 A	14	1	Unknown	1	1	1	Unknown
BMPI1	NA_013043	17	57123856 C	10	8 G	19	11	Unknown	1	1	1	Unknown
BMPI1	NA_013043	17	57123867 A	10	6 G	19	4	Unknown	1	1	2	0.011
BMPI1	NA_013043	17	57123863 G	51	15 A	0	0	Unknown	1	1	0	0.007
BMPI1	NA_013043	17	57123862 A	11	11 A	1	1	Unknown	1	1	1	0.004
BMPI1	NA_013043	17	57123863 A	16	9 C	16	4	Unknown	1	1	1	0.003
BMPI1	NA_013043	17	57123872 G	19	6 A	19	1	Unknown	1	1	1	0.001
BMPI1	NA_013043	17	57123796 T	14	3 G	9	9	Unknown	1	1	1	0.001
BMPI1	NA_013043	17	57123828 G	18	7 T	11	2	Unknown	1	1	0	Unknown
BMPI1	NA_013043	17	57123830 T	11	2 G	27	1	Unknown	1	1	0	0.004
BMPI1	NA_013043	17	57123833 G	13	2 A	5	7	Unknown	0	0	1	0.001
BMPI1	NA_013043	17	57123865 A	1	5 C	24	2	Unknown	1	1	0	0.001
BMPI1	NA_013043	17	57123870 C	24	10 A	38	18	Unknown	1	1	1	1
BMPI1	NA_013043	17	57123862 A	11	5 C	0	0	Unknown	1	1	1	0.001
BMPI1	NA_013043	17	57123833 A	42	5 C	1	0	Unknown	1	1	0	Unknown
BMPI1	NA_013043	17	57123837 A	6	10 A	13	1	Unknown	1	1	0	Unknown
BMPI1	NA_013043	17	57123838 G	53	12 A	0	0	Unknown	1	1	0	Unknown
BMPI1	NA_013043	17	57123843 T	17	18 A	1	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803629 C	43	29 A	0	0	Unknown	1	1	0	0.001
SEPT9	NA_096040	17	72803671 G	22	20 C	0	1	Unknown	1	1	0	0.001
SEPT9	NA_096040	17	72803726 G	39	26 A	0	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803812 C	38	16 A	0	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803906 C	54	8 G	0	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803940 C	12	12 A	0	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803969 G	51	12 A	1	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803934 G	10	9 A	0	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803631 C	50	0	0	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803637 C	10	20 C	0	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803622 C	34	14 A	29	3	Unknown	1	1	0	0.001
SEPT9	NA_096040	17	72803670 C	23	12 C	0	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803670 T	48	5 A	0	0	Unknown	1	1	1	Unknown
SEPT9	NA_096040	17	72803698 A	12	12 C	0	0	Unknown	1	1	1	0.004
SEPT9	NA_096040	17	72803638 T	43	13 C	0	1	Unknown	1	1	1	0.007
SEPT9	NA_096040	17	72804633 T	51	5 C	1	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72804687 G	12	16 A	0	0	Unknown	1	1	1	0.001
SEPT9	NA_096040	17	72804795 G	53	20 A	1	1	Unknown	1	1	1	0.001
SEPT9	NA_096040	17	72804804 G	3	3 C	27	18	Unknown	1	1	1	0.007
SEPT9	NA_096040	17	72803767 C	61	0	0	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803780 T	17	0	0	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803784 C	43	15 G	0	0	Unknown	1	1	1	Unknown
SEPT9	NA_096040	17	72803833 A	7	7 C	19	1	Unknown	1	1	1	Unknown
SEPT9	NA_096040	17	72803823 T	25	15 C	16	9	Unknown	1	1	1	0.002
SEPT9	NA_096040	17	72803856 A	4	4 C	12	12	Unknown	1	1	0	0.007
SEPT9	NA_096040	17	72803823 G	19	1 A	15	0	Unknown	1	1	0	0.047
SEPT9	NA_096040	17	72803822 C	24	4 A	0	0	Unknown	1	1	0	0.001
SEPT9	NA_096040	17	72803823 G	17	5 T	21	2	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803847 C	12	12 C	9	1	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803823 G	13	5 C	17	21	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803824 T	12	5 C	17	21	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803845 A	16	6 G	10	10	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803846 T	12	13 C	13	1	Unknown	1	1	0	0.001
SEPT9	NA_096040	17	72803870 C	10	6 T	11	9	Unknown	1	1	1	Unknown
SEPT9	NA_096040	17	72803784 T	1	1 C	20	3	Unknown	1	1	0	0.004
SEPT9	NA_096040	17	72803845 G	25	9 A	5	5	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803856 A	29	0	26	3	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803874 T	9	9 C	11	21	Unknown	1	1	0	0.007
SEPT9	NA_096040	17	72803864 T	16	17 C	13	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803793 A	14	3 C	13	3	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803867 C	54	8	0	0	Unknown	1	1	0	Unknown
SEPT9	NA_096040	17	72803867 C	22								

Foresight Ex 1042
Foresight v Personalis
IPR2023-00546

MD51	NA_004901	3	17075721	C	26	9	T	16	5	error	Not Applicable	1	0	0	Unknown	
MD51	NA_004901	3	17075887	T	47	11	G	0	0	error	Not Applicable	1	0	0	0.998	
MD51	NA_004901	3	17077027	A	19	8	A	42	2	error	Not Applicable	1	0	0	Unknown	
MD51	NA_004901	3	17078025	T	14	8	C	14	5	error	Not Applicable	1	0	0	Unknown	
MD51	NA_004901	3	17080441	G	10	17	C	10	1	error	Not Applicable	1	0	0	Unknown	
MD51	NA_004901	3	17080513	T	25	3	C	4	error	Not Applicable	1	0	0	0.129		
MD51	NA_004901	3	17080518	G	10	7	C	10	0	error	Not Applicable	1	0	0	0.186	
MD51	NA_004901	3	17080595	C	14	1	T	29	2	error	Not Applicable	1	0	0	1	
MD51	NA_004901	3	17081395	C	10	6	T	13	0	error	Not Applicable	1	0	0	0.789	
MD51	NA_004901	3	17083061	G	24	1	A	40	0	error	Not Applicable	1	0	0	0.001	
MD51	NA_004901	3	17083066	T	13	11	C	13	0	error	Not Applicable	1	0	0	Unknown	
MD51	NA_004901	3	17083413	T	32	1	C	21	4	error	Not Applicable	1	0	0	Unknown	
MD51	NA_004901	3	17083797	D	32	1	C	21	0	error	Not Applicable	1	0	0	Unknown	
MD51	NA_004901	3	17083878	T	37	14	C	6	error	Not Applicable	1	0	0	0.001		
MD51	NA_004901	3	17086022	C	32	3	C	17	12	error	Not Applicable	1	0	0	0.001	
CHC2	NA_004902	4	5403002	T	62	16	G	1	0	error	Not Applicable	0	0	0	0.001	
POGFR	NA_004906	4	54031772	C	62	13	T	0	0	error	Not Applicable	1	0	0	0.004	
POGFR	NA_004906	4	54032622	G	4	9	A	0	0	error	p.P50	1	1	0	0.023	
POGFR	NA_004906	4	54032624	C	52	4	T	0	0	error	Not Applicable	1	1	0	0.024	
POGFR	NA_004906	4	5403424	A	72	7	G	0	0	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_004906	4	54042424	A	20	19	T	0	0	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_004906	4	54044465	A	20	17	C	0	0	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_004906	4	54047822	T	46	11	C	0	0	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_004906	4	54048204	G	46	11	C	0	0	error	Not Applicable	1	1	0	0.009	
POGFR	NA_004906	4	54051237	G	46	11	C	0	0	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_004906	4	54054058	C	40	10	T	0	0	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_004906	4	5405611	T	54	14	C	0	0	error	Not Applicable	1	1	0	0.001	
POGFR	NA_004906	4	54056148	C	50	10	T	0	0	error	Not Applicable	1	1	0	0.079	
POGFR	NA_004906	4	5405707	A	14	11	C	0	0	error	Not Applicable	1	1	0	0.019	
POGFR	NA_004906	4	5405708	T	50	10	T	1	0	0	3_prime_untranslated_region	Not Applicable	1	1	0	0.019
KIT	NA_005022	4	55210260	G	18	1	C	16	9	error	Not Applicable	1	0	0	0.048	
KIT	NA_005022	4	55214515	A	9	3	G	9	0	error	Not Applicable	1	0	0	Unknown	
KIT	NA_005022	4	55221458	G	12	1	A	9	2	error	Not Applicable	1	0	0	Unknown	
KIT	NA_005022	4	55221459	G	26	4	A	19	6	error	Not Applicable	1	0	0	0.001	
KIT	NA_005022	4	55231656	C	14	1	G	28	11	error	Not Applicable	1	0	0	Unknown	
KIT	NA_005022	4	55232712	T	18	4	C	20	10	error	Not Applicable	1	0	0	0.002	
KIT	NA_005022	4	55251402	C	39	9	A	0	0	error	Not Applicable	1	1	0	Unknown	
KIT	NA_005022	4	55252087	T	13	12	C	0	0	error	Not Applicable	1	1	0	Unknown	
KIT	NA_005022	4	55252191	G	50	7	A	0	0	error	Not Applicable	1	1	0	0.046	
KIT	NA_005022	4	55252795	A	60	9	C	3	0	error	Not Applicable	1	1	0	Unknown	
KIT	NA_005022	4	55270625	C	4	4	T	17	4	error	Not Applicable	1	0	0	0.002	
CFSTR	NA_005011	5	14643730	G	21	5	G	18	7	2_prime_untranslated_region	Not Applicable	1	0	0	Unknown	
CFSTR	NA_005011	5	14643736	A	19	5	G	17	6	3_prime_untranslated_region	Not Applicable	1	0	0	Unknown	
CFSTR	NA_005011	5	14643803	C	19	6	G	18	7	alt	Not Applicable	0	0	0	1	
CFSTR	NA_005011	5	14643804	A	22	4	C	20	10	error	p.L85	1	0	0	0.002	
CFSTR	NA_005011	5	14643878	A	23	9	G	4	error	Not Applicable	1	0	0	Unknown		
CFSTR	NA_005011	5	14643879	G	28	10	T	28	10	error	Not Applicable	1	1	0	Unknown	
CFSTR	NA_005011	5	14643783	G	24	8	C	12	error	Not Applicable	1	0	0	Unknown		
CFSTR	NA_005011	5	14643827	T	18	17	C	8	17	error	Not Applicable	1	0	0	Unknown	
CFSTR	NA_005011	5	14643913	C	27	3	T	18	10	error	Not Applicable	1	0	0	0.073	
CFSTR	NA_005011	5	14643954	A	11	9	G	10	10	error	Not Applicable	1	0	0	0.003	
CFSTR	NA_005011	5	14643430	G	9	9	A	9	9	error	Not Applicable	1	0	0	Unknown	
CFSTR	NA_005011	5	14643709	A	12	7	C	2	0	error	Not Applicable	1	0	0	Unknown	
CFSTR	NA_005011	5	14647821	G	20	7	A	25	12	error	Not Applicable	1	0	0	0.001	
CFSTR	NA_005011	5	14645064	A	20	12	C	20	12	error	Not Applicable	1	0	0	0.001	
CFSTR	NA_005011	5	14645065	A	20	12	T	12	error	Not Applicable	1	0	0	0.001		
CFSTR	NA_005011	5	14645703	A	19	8	G	14	6	error	p.T54	1	0	0	0.002	
CFSTR	NA_005011	5	14648423	C	14	4	G	11	13	error	Not Applicable	1	0	0	0.003	
CFSTR	NA_005011	5	14648424	C	12	7	C	12	12	error	Not Applicable	1	0	0	Unknown	
CFSTR	NA_005011	5	14648622	A	12	9	C	9	9	error	Not Applicable	1	0	0	0.005	
CFSTR	NA_005011	5	14648674	C	20	7	C	20	10	error	Not Applicable	1	0	0	0.043	
CFSTR	NA_005011	5	14648720	G	20	2	A	24	16	alt	p.P20	1	0	0	Unknown	
CFSTR	NA_005011	5	14648725	A	20	7	C	18	11	error	Not Applicable	1	0	0	Unknown	
CFSTR	NA_005011	5	14648726	G	26	8	A	14	error	Not Applicable	1	1	0	Unknown		
CFSTR	NA_005011	5	14648789	T	13	2	C	4	error	Not Applicable	1	0	0	Unknown		
CFSTR	NA_005011	5	14648787	G	13	2	C	3	error	Not Applicable	1	1	0	0.001		
CFSTR	NA_005011	5	14648788	A	11	4	C	1	error	Not Applicable	1	1	0	0.001		
CFSTR	NA_005011	5	14648791	A	38	15	G	17	4	error	Not Applicable	1	0	0	Unknown	
CFSTR	NA_005011	5	14648801	C	27	8	C	2	error	Not Applicable	1	0	0	0.002		
CFSTR	NA_005011	5	14648808	T	29	21	T	7	error	Not Applicable	1	0	0	Unknown		
CFSTR	NA_005011	5	14648803	T	19	9	T	19	10	error	Not Applicable	1	0	0	Unknown	
CFSTR	NA_005011	5	14648895	A	18	10	T	12	error	Not Applicable	1	0	0	Unknown		
CFSTR	NA_005011	5	14647248	T	13	14	C	17	11	error	Not Applicable	1	0	0	Unknown	
CFSTR	NA_005011	5	14647249	C	74	33	G	0	0	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_005009	5	14647546	C	19	12	T	22	10	3_prime_untranslated_region	Not Applicable	1	0	0	0.071	
POGFR	NA_005009	5	14647554	G	21	4	A	22	18	error	Not Applicable	1	0	0	Unknown	
POGFR	NA_005009	5	14647723	C	19	21	T	12	11	error	Not Applicable	1	0	0	0.033	
POGFR	NA_005009	5	14647785	T	19	11	C	12	error	Not Applicable	1	0	0	0.002		
POGFR	NA_005009	5	14647809	C	12	1	C	6	1.0E-11	error	Not Applicable	1	0	0	0.002	
POGFR	NA_005009	5	14648423	T	17	9	C	9	9	error	Not Applicable	1	0	0	Unknown	
POGFR	NA_005009	5	14648424	C	17	6	A	29	6	error	Not Applicable	1	0	0	0.001	
POGFR	NA_005009	5	14648623	A	29	10	G	12	error	Not Applicable	1	0	0	0.004		
POGFR	NA_005009	5	14648624	T	29	9	T	12	error	Not Applicable	1	0	0	Unknown		
POGFR	NA_005009	5	14648625	A	30	10	A	29	7	error	Not Applicable	1	1	0	0.001	
POGFR	NA_005009	5	14648626	T	16	10	A	19	6	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_005009	5	14648627	A	16	9	G	29	14	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_005009	5	14648628	C	30	10	T	29	8	error	Not Applicable	1	1	0	0.001	
POGFR	NA_005009	5	14648629	A	30	10	A	29	8	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_005009	5	14648630	T	30	10	T	29	8	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_005009	5	14648631	C	30	10	C	29	8	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_005009	5	14648632	A	30	10	A	29	8	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_005009	5	14648633	T	30	10	T	29	8	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_005009	5	14648634	C	30	10	C	29	8	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_005009	5	14648635	A	30	10	A	29	8	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_005009	5	14648636	T	30	10	T	29	8	error	Not Applicable	1	1	0	Unknown	
POGFR	NA_005009	5	14648637	A	30	10	A	29	8	error	Not Applicable	1	1	0	Unknown	
NPFL	NA_005020	5	17075721	A	51	20	C	2	1	error	Not Applicable	1	1	0	0.004	
NPFL	NA_005020	5	17075723	G	54	22	A	0	0	error	Not Applicable	1	1	0	0.013	
NPFL	NA_005020	5	17075457	T	50	0	C	0	0	error	Not Applicable	1	1	0	0.001	
NPFL	NA_005020	5	17075462	T	51	0	C	0	0	error	Not Applicable	1	1	0	0.001	
NPFL	NA_005020	5	17075653	G	47	6	A	0	0	error	Not Applicable	1	0	0	Unknown	
NPFL	NA_005020	5	17075678	A	9	10	G	0	0	error	Not Applicable	1	0	0	Unknown	
ND5L	NA_024005	5	17053008	A	60	19	C	0	0	error	Not Applicable	1	0	0	Unknown	
ND5L	NA_024005	5	17053108	C	14</											

[illegible]

Supplementary Table 5. Cancer susceptibility genes.

ABCB1	5243	BARD1	580	CDC2	983	CYR61	3491	FANCA	2175
ABCB5	340273	BAX	581	CDC73	79577	DAPK1	1612	FANCB	2187
ABCC1	4363	BCAR1	9564	CDH1	999	DAXX	1616	FANCC	2176
ABCC2	1244	BCAR3	8412	CDK2	1017	DCK	1633	FANCD2	2177
ABCC3	8714	BCL2	596	CDK4	1019	DCLRE1A	9937	FANCE	2178
ABCC4	10257	BCL2A1	597	CDK6	1021	DCLRE1B	64858	FANCF	2188
ABCC5	10057	BCL2L1	598	CDK7	1022	DCLRE1C	64421	FANCG	2189
ABCC6	368	BCL2L11	10018	CDKN1A	1026	DDB1	1642	FANCI	55215
ABCG2	9429	BCR	613	CDKN1B	1027	DDB2	1643	FANCL	55120
ABL1	25	BHD	201163	CDKN2A	1029	DHCR24	1718	FANCM	57697
ACACA	31	BID	637	CDKN2B	1030	DIABLO	56616	FAS	355
ACE	1636	BIRC2	329	CDKN2D	1032	DLG5	9231	FASLG	356
ACP5	54	BIRC3	330	CETN2	1069	DMC1	11144	FCGR2A	2212
ADH1B	125	BIRC4	331	CFLAR	8837	DNAJC15	29103	FCGR3A	2214
ADH1C	126	BIRC5	332	CFTR	1080	DNMT1	1786	FCGR3B	2215
ADIPOQ	9370	BIRC7	79444	CHAF1A	10036	DNMT3B	1789	FEN1	2237
ADRB2	154	BLM	641	CHEK1	1111	DPYD	1806	FGF2	2247
ADRB3	155	BMP2	650	CHEK2	11200	DUT	1854	FGFR2	2263
AGT	183	BMPR1A	657	CHUK	1147	E2F1	1869	FGFR3	2261
AHR	196	BNIP3	664	CLDN4	1364	ECGF1	1890	FH	2271
AIP	9049	BRAF	673	CLK2	1196	EGF	1950	FHIT	2272
AKAP13	11214	BRCA1	672	CLU	1191	EGFR	1956	FLT3	2322
AKR1C1	1645	BRCA2	675	COL18A1	80781	EIF4E	1977	FN1	2335
AKR1C3	8644	BRIP1	83990	COMT	1312	EIF4EBP1	1978	FOLH1	2346
AKT1	207	BSG	682	CRABP1	1381	ELA2	1991	FOLR1	2348
AKT2	208	BUB1B	701	CRAC1	338377	ELAC2	60528	FOXO1A	2308
ALDH2	217	C19orf40	91442	CREB1	1385	EME1	146956	FOXO3A	2309
ALKBH2	121642	CA9	768	CREBBP	1387	EME2	197342	FPGS	2356
ALKBH3	221120	CAPN10	11132	CSAG2	728461	ENPP1	5167	FRAP1	2475
ALOX12	239	CARD15	64127	CSE1L	1434	EP300	2033	GCLC	2729
ALOX5	240	CASP1	834	CSF1R	1436	EPHB2	2048	GCNT1	2650
AMACR	23600	CASP10	843	CSNK2A1	1457	EPHX1	2052	GDF15	9518
APAF1	317	CASP3	836	CTLA4	1493	EPO	2056	GH1	2688
APC	324	CASP7	840	CTNNB1	1499	ERBB2	2064	GHRL	51738
APEX1	328	CASP8	841	CXCL12	6387	ERBB3	2065	GNB3	2784
APEX2	27301	CASP9	842	CXCR4	7852	ERBB4	2066	GPC3	2719
APOB	338	CAT	847	CYBA	1535	ERCC1	2067	GPI	2821
APOE	348	CAV1	857	CYLD	1540	ERCC2	2068	GPX1	2876
APP	351	CBS	875	CYP11A1	1583	ERCC3	2071	GSK3B	2932
APT-X	54840	CCL2	6347	CYP17A1	1586	ERCC4	2072	GSTA1	2938
AR	367	CCL5	6352	CYP19A1	1588	ERCC5	2073	GSTM1	2944
ARL11	115761	CCND1	595	CYP1A1	1543	ERCC6	2074	GSTM3	2947
ASNS	440	CCND3	896	CYP1A2	1544	ERCC8	1161	GSTO1	9446
ATM	472	CCNE1	898	CYP1B1	1545	ESR1	2099	GSTO2	119391
ATP7B	540	CCNH	902	CYP2A13	1553	ESR2	2100	GSTP1	2950
ATR	545	CCR2	1231	CYP2A6	1548	EXO1	9156	GSTT1	2952
ATRIIP	84126	CCR5	1234	CYP2C19	1557	EXT1	2131	GTF2H1	2965
AXIN2	8313	CD14	929	CYP2C9	1559	EXT2	2132	GTF2H2	2966
BACH1	571	CD34	947	CYP2D6	1565	F2	2147	GTF2H3	2967
BAD	572	CD36	948	CYP2E1	1571	F5	2153	GTF2H4	2968
BAG1	573	CD40	958	CYP3A4	1576	FABP2	2169	GTF2H5	404672
BAK1	578	CD44	960	CYP3A5	1577	FADD	8772	H2AFX	3014

Supplementary Table 5. Cancer susceptibility genes.

HDAC1	3065	INS	3630	MKI67	4288	NOTCH1	4851	PRKCB1	5579
HFE	3077	INSR	3643	MLH1	4292	NPM1	4869	PRKCD	5580
HGF	3082	IRF1	3659	MLH3	27030	NQO1	1728	PRKCE	5581
HGS	9146	IRS1	3667	MLL	4297	NR1I2	8856	PRKCI	5584
HIF1A	3091	IRS2	8660	MMP1	4312	NR3C1	2908	PRKDC	5591
HLA-A	3105	ITGA2	3673	MMP13	4322	NR4A1	3164	PRNP	5621
HLA-B	3106	ITGB1	3688	MMP14	4323	NTHL1	4913	PTCH	5727
HLA-C	3107	ITGB3	3690	MMP2	4313	NUDT1	4521	PTEN	5728
HLA-DPB1	3115	JAK2	3717	MMP3	4314	OGG1	4968	PTGS1	5742
HLA-DQA1	3117	JUN	3725	MMP7	4316	P2RX7	5027	PTGS2	5743
HLA-DQB1	3119	KCNQ1OT1	10984	MMP9	4318	PAK1	5058	PTK2	5747
HLA-DRB1	3123	KDR	3791	MMS19	64210	PALB2	79728	PTPN1	5770
HMOX1	3162	KIT	3815	MNAT1	4331	PALLD	23022	PTPN13	5783
HRAS	3265	KLF6	1316	MPG	4350	PARP1	142	PTPRJ	5795
HRPT2	3279	KLK3	354	MPL	4352	PARP2	10038	RAB6C	84084
HSD17B1	3292	KLRK1	22914	MPO	4353	PCNA	5111	RAD1	5810
HSD3B2	3284	KRAS	3845	MRE11A	4361	PDCD8	9131	RAD17	5884
HSPA1B	3304	LAPTM4B	55353	MS	4397	PDGFRA	5156	RAD18	56852
HSPA5	3309	LEP	3952	MSH2	4436	PEBP1	5037	RAD23A	5886
HSPB1	3315	LEPR	3953	MSH3	4437	PER1	5187	RAD23B	5887
HSPD1	3329	LGALS3	3958	MSH4	4438	PGR	5241	RAD50	10111
HTATIP2	10553	LIG1	3978	MSH5	4439	PHB	5245	RAD51	5888
HUS1	3364	LIG3	3980	MSH6	2956	PHOX2B	8929	RAD51C	5889
ICAM1	3383	LIG4	3981	MSR1	4481	PIK3CA	5290	RAD51L1	5890
ID1	3397	LRP1	4035	MTHFR	4524	PIK3CG	5294	RAD51L3	5892
IFNG	3458	LSP1	4046	MTR	4548	PLAU	5328	RAD52	5893
IFNGR1	3459	LTA	4049	MTRR	4552	PML	5371	RAD54B	25788
IGF1	3479	LYN	4067	MUS81	80198	PMS1	5378	RAD54L	8438
IGF1R	3480	MAD2L1	4085	MUTYH	4595	PMS2	5395	RAD9A	5883
IGF2	3481	MAD2L2	10459	MVP	9961	PMS2L3	5387	RAF1	5894
IGF2R	3482	MADD	8567	MX1	4599	PNKP	11284	RALBP1	10928
IGFBP1	3484	MADH4	4089	MYC	4609	POLB	5423	RARA	5914
IGFBP3	3486	MAP2K1	5604	MYCL1	4610	POLD1	5424	RARB	5915
IKBKB	3551	MAP3K1	4214	MYCN	4613	POLE	5426	RASSF1	11186
IL10	3586	MAP3K5	4217	MYH1	4619	POLG	5428	RB1	5925
IL12A	3592	MAPK1	5594	NAT1	9	POLH	5429	RBBP8	5932
IL12B	3593	MAPK14	1432	NAT2	10	POLI	11201	RBL2	5934
IL13	3596	MAPK3	5595	NBS1	4683	POLK	51426	RDM1	201299
IL18	3606	MAPK8	5599	NCOA3	8202	POLL	27343	RECQL	5965
IL1A	3552	MBD4	8930	NCOR2	9612	POLM	27434	RECQL4	9401
IL1B	3553	MBL2	4153	NEIL1	79661	POLN	353497	RECQL5	9400
IL1R1	3554	MC1R	4157	NEIL2	252969	PON1	5444	RELA	5970
IL1RN	3557	MCL1	4170	NEIL3	55247	PPARA	5465	RET	5979
IL2	3558	MDC1	9656	NF1	4763	PPARG	5468	RETN	56729
IL24	11009	MDM2	4193	NF2	4771	PPARGC1A	10891	REV1L	51455
IL4	3565	MEN1	4221	NFE2L2	4780	PPP1R13L	10848	RFC1	5981
IL4R	3566	MET	4233	NFKB1	4790	PPP1R1B	84152	RNASEL	6041
IL6	3569	MGMT	4255	NFKBIA	4792	PRDM2	7799	RPA1	6117
IL6R	3570	MICA	4276	NFKBIL1	4795	PRKAB1	5564	RPA2	6118
IL8	3576	MICB	4277	NME1	4830	PRKACA	5566	RPA3	6119
IL8RA	3577	MIF	4282	NOS2A	4843	PRKAR1A	5573	RPA4	29935
IL8RB	3579	MITF	4286	NOS3	4846	PRKCA	5578	RPS6KA1	6195

Supplementary Table 5. Cancer susceptibility genes.

RPS6KB1	6198	THBS1	7057	UGT1A7	54577
RRM1	6240	TIMP1	7076	VDR	7421
RRM2B	50484	TIMP2	7077	VEGF	7422
RUNX1	861	TLR1	7096	VHL	7428
RXRA	6256	TLR10	81793	WRN	7486
SBDS	51119	TLR2	7097	WT1	7490
SDHB	6390	TLR4	7099	WWOX	51741
SDHC	6391	TLR6	10333	XAB2	56949
SDHD	6392	TLR9	54106	XAF1	54739
SERPINE1	5054	TMC6	11322	XBP1	7494
SFN	2810	TNF	7124	XPA	7507
SFTP	6439	TNFAIP1	7126	XPC	7508
SH2D1A	4068	TNFAIP3	7128	XRCC1	7515
SHBG	6462	TNFRSF10A	8797	XRCC2	7516
SHC1	6464	TNFRSF10B	8795	XRCC3	7517
SHFM1	7979	TNFRSF10C	8794	XRCC4	7518
SHH	6469	TNFRSF10D	8793	XRCC5	7520
SHMT1	6470	TNFRSF11B	4982	XRCC6	2547
SKP2	6502	TNFRSF1A	7132	YBX1	4904
SLC11A1	6556	TNFRSF1B	7133	ZAP70	7535
SLC19A1	6573	TNFSF10	8743	ZBTB16	7704
SLC29A1	2030	TNFSF11	8600	ZNF350	59348
SMAD2	4087	TNFSF13B	10673	ZNRD1	30834
SMAD3	4088	TOP1	7150		
SMARCB1	6598	TOP2A	7153		
SMUG1	23583	TOP2B	7155		
SNAI2	6591	TOP3A	7156		
SOCS1	8651	TOPBP1	11073		
SOCS3	9021	TP53	7157		
SOD1	6647	TP53BP1	7158		
SOD2	6648	TP73	7161		
SP1	6667	TP73L	8626		
SPO11	23626	TRAF1	7185		
SPP1	6696	TREX1	11277		
SRC	6714	TSC1	7248		
SRD5A2	6716	TSC2	7249		
SRI	6717	TSG101	7251		
STAT1	6772	TSHR	7253		
STAT3	6774	TSPY1	7258		
STK11	6794	TTK	7272		
STK6	6790	TUBB	203068		
SUFU	51684	TUBB3	10381		
SULT1A1	6817	TWIST1	7291		
TCEAL1	9338	TXN	7295		
TCF1	6927	TYMS	7298		
TDG	6996	UBE2A	7319		
TDP1	55775	UBE2B	7320		
TERT	7015	UBE2N	7334		
TGFB1	7040	UBE2V2	7336		
TGFBR1	7046	UGCG	7357		
TGFBR2	7048	UGT1A1	54658		
TGM2	7052	UGT1A6	54578		

Supplementary Table 6. Validated germline non-synonymous variants in cancer susceptibility genes.

Gene_name	Chromosome	dbSNP(0:no; 1:yes)	Start_position (B36)	Variant_allele	Reference_allele	Ensembl_transcript_id	Variant_type	Amino_acid_change	Polyphen_prediction	Watson or Venter?	Conservation	OMIM
ABCB5	7	0	20734602	T	A	ENST00000382055	missense	p.I530L	NULL		0	0.804
ABCC1	16	0	16049580	T	G	NM_004996	missense	p.R433S	NULL		0	0.972
ABCC6	16	1	16159100	T	C	NM_001171	missense	p.R1268Q	NULL	VENTER	0.992	*603234.0003 and *603234.0011
BRCA2	13	1	31804729	C	A	NM_000059	missense	p.N372H	NULL		0	0.004 600185.0013
CFTR	7	1	116986769	A	G	NM_000492	missense	p.V470M	NULL		0	1 *602421.0023
COL18A1	21	0	45735616	G	C	ENST00000359759	missense	p.P1122R	NULL		0	0.718
CYP19A1	15	0	49291948	A	C	NM_000103	missense	p.R375L	NULL		0	1
DLG5	10	1	79286611	C	T	ENST00000372392	missense	p.Q30R	NULL	VENTER/WATSON	0.268	*604090.0001
DPYD	1	1	98121473	A	G	ENST00000370192	missense	p.R29C	NULL	VENTER/WATSON	0.887	274270.0004
ELAC2	17	1	12855734	A	G	NM_018127	missense	p.S217L	NULL		0	0.004 *605367.0001
ERCC6	10	0	50410617	A	G	NM_000124	missense	p.R134V	NULL		0	0.873
FLT3	13	0	27524716	T	C	ENST00000380982	missense	p.V194M	NULL		0	0.486
HGS	17	0	77261521	G	C	ENST00000329138	missense	p.T7S	NULL		0	1
HLA-C	6	0	31345103	C	T	ENST00000376235	missense	p.T363A	NULL	VENTER	NULL	
HLA-C	6	0	31346238	T	G	ENST00000376235	missense	p.P208H	NULL	VENTER		0.25
IGFBP3	7	0	45920990	A	G	NM_001013398	missense	p.T283I	NULL		0	0.157
IL1RN	2	0	113606755	A	G	NM_173841	missense	p.A127T	NULL		0	0.894
IL24	1	0	205142944	G	T	ENST00000367095	missense	p.L181V	NULL		0	0.638
LEPR	1	1	65809029	G	A	NM_002303	missense	p.K109R	NULL	VENTER	0.989	*601007.0004
LEPR	1	1	65831101	G	A	NM_002303	missense	p.Q223R	NULL	VENTER/WATSON	0.998	*601007.0003
MC1R	16	1	88513345	T	G	ENST00000304984	missense	p.V60L	NULL	WATSON	1	*155555.0006
MC1R	16	1	88513441	A	G	ENST00000304984	missense	p.V92M	NULL		0	0.997 *155555.0002
MCL1	1	0	148817589	G	T	NM_021960	missense	p.M231L	NULL		0	1
MTHFR	1	1	11777063	G	T	ENST00000376583	missense	p.E470A	NULL	WATSON	0.993	607093.0004
MYH1	17	0	10351978	T	C	NM_005963	missense	p.G640S	NULL		0	0.988
PALB2	16	0	23545216	A	G	NM_024675	missense	p.P864S	NULL		0	0.998
PLAU	10	1	75343107	C	T	NM_002658	missense	p.L141P	NULL	VENTER/WATSON	0.017	*191840.0001
PON1	7	1	94775382	C	T	NM_000446	missense	p.Q192R	NULL	VENTER/WATSON	NULL	168820.0001
SDHD	11	1	111462875	A	G	NM_003002	missense	p.G12S	NULL		0	0.239 *602690.0011
TGFBR1	9	0	100934683	G	A	NM_004612	missense	p.I139V	NULL		0	1
TP53	17	1	7520197	C	G	NM_000546	missense	p.P72R	NULL	VENTER/WATSON	0.752	*191170.0005
TRAF1	9	0	122715716	G	A	NM_005658	missense	p.M139T	NULL		0	0.034
WRN	8	0	31124149	T	C	NM_000553	missense	p.S1141L	NULL		0	0.001