Whole genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution and clonal tides

Short title: Sequencing Reveals Natural History of Myeloma

Jan B. Egan¹, Chang-Xin Shi¹, Waibhav Tembe², Alexis Christoforides², Ahmet Kurdoglu², Shripad Sinari², Sumit Middha³, Yan Asmann³, Jessica Schmidt¹, Esteban Braggio¹, Jonathan J. Keats², Rafael Fonseca¹, P. Leif Bergsagel¹, David W. Craig², John D. Carpten², A. Keith Stewart¹.

Affiliations:
1 – Mayo Clinic Arizona, Division of Hematology-Oncology, 13400 E. Shea Blvd Scottsdale, Arizona 85259
2 – Translational Genomics Research Institute, 445 N. Fifth Street, Phoenix, Arizona 85004
3 – Mayo Clinic, Department of Health Sciences Research, 200 First Street SW, Rochester, Minnesota 55905

Corresponding Author:
A. Keith Stewart
Mayo Clinic Collaborative Research Building
13400 E Shea Blvd
Scottsdale, AZ 85259-5494
stewart.keith@mayo.edu
Phone: 480-301-4411
Fax: 480-301-8387

Scientific Category: Lymphoid Neoplasia
ABSTRACT

The longitudinal evolution of a myeloma genome from diagnosis to plasma cell leukemia has not previously been reported. We used whole genome sequencing (WGS) on four purified tumor samples and patient germline DNA drawn over a five year period in a t(4;14) multiple myeloma patient. Tumor samples were acquired at diagnosis, first relapse, second relapse and end-stage secondary plasma cell leukemia (sPCL). In addition to the t(4;14), all tumor time points also shared 10 common single nucleotide variants (SNV) on WGS comprising shared initiating events. Interestingly, we observed genomic sequence variants that waxed and waned with time in progressive tumors suggesting the presence of multiple independent, yet related clones at diagnosis that rose and fell in dominance. Five newly acquired SNV, including truncating mutations of \textit{RB1} and \textit{ZKSCAN3}, were observed only in the final sPCL sample suggesting leukemic transformation events. This longitudinal WGS characterization of the natural history of a high-risk myeloma patient demonstrated tumor heterogeneity at diagnosis with shifting dominance of tumor clones over time and has also identified potential mutations contributing to myelomagenesis as well as transformation from myeloma to overt extramedullary disease such as sPCL.
INTRODUCTION

Multiple myeloma is a treatable mature B-cell malignancy and progress in improving survival has been rapid in recent years. Despite this, therapy is usually not curative and further insights into the genomic basis for disease and its progression are critical. An early event in myelomagenesis is the dysregulation of Cyclin D1. This dysregulation results from primary Immunoglobulin Heavy Chain (IgH) translocations with partners MMSET/FGFR3, CCND1, CCND3, cMAF and MAFB or an alternative pathway of chromosomal hyperdiploidy. Beyond these initiating events common mutations that accompany progression include: multiple chromosomal losses including deletion of RB1 and TP53, amplification of chromosome 1q, mutations in the NF-κB signaling pathway or in K- and N-RAS. PTEN deletions have been reported in 20-33% of sPCL cases, which is significantly higher than the incidence in myeloma (6%) suggesting PTEN deletion may contribute to leukemic transformation. The presence of t(4;14), t(14;16), chromosome 13 deletion, deletion of 17p, amplification of chromosome 1 and hypodiploidy are associated with high risk/poor prognosis.

The Multiple Myeloma Research Consortium (MMRC) recently reported the whole genome and the exome sequences of 38 myeloma patients. Along with common mutations in known genes such as NRAS, KRAS and TP53, novel mutations were detected relatively frequently in several genes including e.g. FAM46C and DIS3. In addition, a significant clustering of mutations was observed in histone modifying enzymes. Of clinical translational relevance for that study was the first report of BRAF kinase domain mutations in myeloma, with obvious implications for the treatment of disease.
While these recent findings from the MMRC\textsuperscript{19} have provided a detailed characterization of the myeloma genome at a single time point, the genomic basis of the progression of myeloma over time from diagnosis to end stage disease\textsuperscript{20,21} remains unclear. Therefore to further understand the natural history of disease genome progression from diagnosis of myeloma to terminal secondary plasma cell leukemia (sPCL), we conducted a longitudinal study of a single myeloma patient for whom four tumor samples were obtained throughout the disease course. Whole genome sequencing (WGS) conducted on these samples has identified 36 validated novel single nucleotide variants (SNV), 27 of which are located in genes that are mutated in the MMRC cohort\textsuperscript{19} and/or the COSMIC database\textsuperscript{22} but do not share the same point mutations. PolyPhen-2\textsuperscript{23}/SIFT\textsuperscript{24} predict 18 of these 27 SNV are damaging. In addition, 10 of the 27 SNV are shared at all tumor time points while five SNV are unique to the sPCL sample and may be associated with leukemic transformation. Furthermore, mate-pair sequencing (MPS) of the sPCL has identified 79 structural variants.

**MATERIAL AND METHODS**

Detailed methods are available in the supplemental material file. Methods in brief are presented here.

**Samples**

Patient samples were acquired with patient consent in accordance with the Declaration of Helsinki with approval from the Mayo Clinic Institutional Review Board. Iliac crest bone marrow samples were acquired at diagnosis, first and second relapse. Peripheral blood was
obtained at the time of second relapse for use as germline control and again at the time of sPCL diagnosis for circulating tumor cells. Bone marrow and peripheral blood samples were treated with ACK lysis buffer to remove red cells and CD138+ cell populations were isolated using anti-CD138 antibodies on a Robocept (StemCell Technologies). Tumor cell purity was estimated using a slide based kappa/lambda assay. Purified tumor cells were preserved either in RNALater (Qiagen) or as dry pellets at -80°C. DNA was isolated from preserved samples with the Puregene kit (Qiagen) following manufacturer’s protocol.

**Whole Genome Sequencing**

Single end WGS was completed for the constitutional, diagnostic, first relapse and second relapse samples on the Life Technologies SOLiD™ platform. Paired end WGS and mate pair sequencing (MPS) was conducted for the sPCL sample on the Illumina® HiSeq 2000 platform. All sequencing reads were aligned to NCBI36/hg18, SNV were identified and pair wise comparison made between the tumor samples and the germline sample for the identification of somatic variants. After visual confirmation in the Integrative Genomics Viewer25, the SNV were validated by capillary sequencing in all samples. The MMRC cohort19 of 38 myeloma genomes and the COSMIC database22 of multiple tumor types were then queried to determine, within larger tumor populations, the frequency of all somatic mutations in the genes containing validated somatic SNV in this case study. Comparative pathway analysis was conducted to identify potentially shared pathways among the SNV unique to specific time points. Mate pair analysis was conducted for the identification of structural rearrangements.
Microarrays and Fluorescent *in-situ* Hybridization

DNA samples from all tumor time points were fragmented, labeled and hybridized to the 244A Human Genome CGH Microarray (Agilent) according to the manufacturer’s recommendation. Aberration calling was performed in Genomic WorkBench v6.5 (Agilent Technologies) using the ADM-2 algorithm with a 5.5 threshold and 0.2 log2 and 3 probe filters.

RNA samples from diagnosis and sPCL were submitted to the Mayo Clinic genomics core for gene expression studies. Samples were labeled and hybridized to the HG-U133Plus 2.0 genechip (Affymetrix). Expression estimates were extracted using the MAS5.0 algorithm implemented in Expression Console using default parameters.

The expression level of identified candidate genes was analyzed across public databases including: the Multiple Myeloma Research Consortium database (http://www.broadinstitute.org/mmmgp/home), GlaxoSmithKline Cancer Cell Line Genomic Profiling Dataset (https://array.nci.nih.gov/caarray/project/woost-00041) and our Mayo clinic myeloma research laboratory database (unpublished data). Fluorescent *in-situ* Hybridization (FISH) assays were conducted at the time of diagnosis in a clinical laboratory in compliance with federal regulatory standards. The microarray datasets associated with this manuscript have been deposited in GEO as a super series under accession GSE36825 the individual platforms are available individually under GSE36822, GSE36823, and GSE36824

RESULTS

Case history
A 67 year-old woman was diagnosed with multiple myeloma and presented with bone marrow containing 25% plasma cells harboring a t(4;14)(p16;q32) and chromosome 13 deletion detected by FISH. The patient was initially treated with lenalidomide and low dose dexamethasone with very good response but progression of disease first observed at 22 months. Bone marrow was collected for the first relapse sample at 23 months. The patient then received treatment with carfilzomib on a clinical trial for the next four months with good response. Treatment was then discontinued to pursue stem cell mobilization that was ultimately unsuccessful. At 33 months, disease progression was again observed and the patient was treated with bortezomib in combination with an experimental agent SGN-40 for two months but therapy was discontinued due to toxicity. Collection of the bone marrow for the second relapse sample as well as peripheral blood for constitutional genome analysis was conducted at 39 months. Treatment with melphalan, prednisone and bortezomib was then begun and continued until month 43 and was again discontinued due to toxicity. For the next four months, the patient received various combination chemotherapies with limited success. At month 53 the myeloma progressed to sPCL and treatment was discontinued and patient ultimately referred to hospice. Peripheral blood was collected for the sPCL sample at month 54.

Array comparative genomic hybridization (aCGH) of the diagnostic and the first relapse samples in this patient contained numerous copy number abnormalities which in total encompassed 3968 genes of which 1235 were expressed at diagnosis (Keats, et al., companion manuscript this issue). We wished to interrogate the specific genes that may be involved in myelomagenesis, disease progression and drug sensitivity or resistance. We therefore decided to conduct whole
genome sequencing (WGS) to further investigate the individual genetic events associated with the natural history of progressive myeloma.

**Whole Genome Sequencing**

WGS was conducted for germline DNA and at four time points during the tumor evolution including: diagnosis, first relapse (23 months), second relapse (39 months) and sPCL (54 months). For the first three tumor time points and germline samples, an average of $2.45 \times 10^9$ single end reads/sample were sequenced with an average of $1.76 \times 10^9$ reads/sample (71.8%) aligning to the hg18 reference (Table 1). The sPCL sample had $1.04 \times 10^9$ paired end reads/sample with $0.97 \times 10^9$ reads/sample (93.1%) aligning to the hg18 reference. Average coverage of aligned sequence was 30X prior to duplicate removal and 20X after removal of duplicates. Over two million constitutional SNV were identified in the germline sample and tumor samples had over 14,000 additional somatic SNV with the sPCL sample containing a much larger number of variants than the other samples. The majority of the SNV were located within introns and intergenic space. From the somatic SNV located in exons, 124 candidate, nonsynonymous SNV were identified that were tumor specific, 36 of which validated with capillary sequencing (Table 2, Suppl Table 1). Of the 36 SNV that validated, 27 were located in genes containing somatic mutations in the MMRC 38 myeloma genomes\textsuperscript{19} and/or the COSMIC database\textsuperscript{22} of multiple tumor types. While the genes containing the 27 SNV were also somatically mutated in these other tumor populations, no somatic SNV were shared between these populations and this work. Eighteen (68%) of these 27 SNV were predicted as potentially damaging and three (11%) were predicted as truncating amino acid changes. Of those 27 genes, nine (33%) were located in genes identified as mutated in both the MMRC cohort and COSMIC
Identification of SNV driving myelomagenesis

SNV shared at all tumor time points are likely to be tumor initiating events. Fifteen of the validated SNV were shared at all tumor time points, with 10 of these 15 presenting in somatically mutated genes in the MMRC cohort and/or the COSMIC database that include: AFF1, ATXN1, CNGA3, COL2A1, CSMD3, KRT9, LRRC4C, MAGI1, MYPN, and RNF145 (Table 2). Of particular interest is AFF1, which is widely expressed in myeloma, and was also identified as mutated in the MMRC cohort. AFF1 is a fusion partner with MLL in both pediatric and adult acute leukemias that has been implicated in the alteration of histone methylation signatures in mice. In this patient the AFF1 gene contains a missense mutation (R163H) that Mutation Taster predicts may result in the loss of phosphoserine sites downstream of a putative altered splice site.

Identification of SNV driving leukemic transformation

SNV unique to the sPCL sample may contribute to leukemic transformation from myeloma to sPCL. Seven SNV were unique to sPCL, of which five were present in genes somatically mutated in the MMRC cohort and/or the COSMIC database and include: RB1, TNN, TUBB8, ZKSCAN3 and ZNF521 (Table 2). Two genes of particular interest are RB1 and ZKSCAN3,
which were somatically mutated in both the MMRC cohort\textsuperscript{19} and the COSMIC database\textsuperscript{22}. A deletion of \textit{RB1} was detected in this patient at diagnosis, thus the SNV in the sPCL sample is hemizygous as it occurred in the only remaining copy of \textit{RB1}. The mutations in \textit{RB1} (E287*) and \textit{ZKSCAN3} (E28*) are predicted to result in truncated proteins.\textsuperscript{23,28} Although \textit{RB1} is expressed at both diagnosis and sPCL, the \textit{RB1} truncation at the sPCL time point is predicted by Mutation Taster\textsuperscript{28} to result in a protein lacking critical binding pockets and interaction domains. The well-conserved KRAB and SCAN domains within \textit{ZKSCAN3} are predicted by Mutation Taster to be eliminated with this truncation. Furthermore, no expression of \textit{ZKSCAN3} was detected in this patient.

**Clonal evolution over time**

While 10 common SNV were shared at all tumor time points, we also observed variants which were only detectable at alternating time points i.e. diagnosis and second relapse, or first relapse and sPCL suggesting the waxing and waning of different clones with time and treatment (Figure 2, Suppl Table 1). At diagnosis, one SNV, \textit{PDE4DIP}, was unique, not found at later time points. Biologically this is difficult to explain and may be artifact although the mutation was identified both in libraries prepared from genomic DNA and from capillary sequencing which utilized whole genome amplified DNA. This SNV is different from the somatic mutations present in other tumor types within the COSMIC database\textsuperscript{22} including a single chronic lymphocytic leukemia case. Six SNV were shared by only the diagnostic and the second relapse samples of which five SNV were located in genes with somatic mutations in the MMRC cohort\textsuperscript{19} and/or the COSMIC database\textsuperscript{22} including: \textit{C12orf42}, \textit{DOK5}, \textit{PARD3B}, \textit{PPFIBP1} and \textit{ZNF557}. The first relapse and sPCL samples did not share any unique SNV beyond the 10 observed at all tumor
time points. Array CGH and Fluorescent \textit{in-situ} Hybridization data described in detail in a companion manuscript in this issue (Keats et al.) indicate however that these time points share a common progenitor that is not detectable with these validated, coding SNV. The first relapse sample had seven unique SNV not found at any other time point of which six were present in genes somatically mutated in the MMRC cohort and/or the COSMIC database that include: \textit{ATXN1, CACNA1S, DSC1, PCDH7, PTPRD} and \textit{TLR9} while the sPCL had seven SNV unique only to the sPCL sample of which five were present in genes containing somatic mutations in the MMRC cohort and/or the COSMIC database that include: \textit{RB1, TNN, TUBB8, ZKSCAN3}, and \textit{ZNF521}. Comparative pathway analysis revealed no shared pathways between the genes containing SNV unique to specific time points. While the evolutionary divergence of the diagnostic and second relapse samples is relatively small with only one unique SNV, the evolutionary divergence of the first relapse and sPCL samples is much larger with the samples acquiring six and five mutations respectively.

\textbf{Identification of structural variants}

The sPCL mate-pair analysis revealed the presence of 79 structural variants including the diagnostic t(4;14), three rearrangements involving chromosomes 10 and 12, as well as complex structural rearrangement in regions on chromosomes 11 and 12 (Figure 3, Suppl Table 2). All of the translocation breakpoints in chromosome 12 occur in regions with multiple breakpoints indicating complex structural rearrangement. This suggests intra-chromosomal rearrangement likely occurred within chromosome 12 prior to a translocation event with chromosome 10. Interpretation of these complex events is beyond the scope of this study and we suggest not pertinent to the main message. Thirteen inversions were detected, ten of which occurred in gene
rich chromosomal regions including **MLLT11** and **NOTCH2NL** which are expressed in myeloma. **MLLT11** is a translocation partner in acute leukemia\(^{29}\) and over-expression has been associated with poor outcome in pediatric acute myeloid leukemia and adult myelodysplastic syndrome.\(^{30}\) A translocation was identified between **NOTCH2NL** and **LCLAT1** in the MMRC cohort\(^{19}\) suggesting this NOTCH signaling pathway protein may be prone to structural rearrangement. Half of the ten biallelic deletions observed occurred in non-coding regions while the remaining deletions included deletion of one or more genes including: **BIRC2**, **BIRC3** and **TP53**.

**DISCUSSION**

The comprehensive study of myeloma genomes in the MMRC cohort\(^{19}\) has confirmed the heterogeneity of myeloma tumors and identified potential pathways and mutations of interest for further study, but was unable to address genomic changes over time and specifically changes that contribute to the leukemic transformation leading to sPCL. This work reports several potential drivers of myeloma as well as initiators of leukemic transformation from myeloma to sPCL.

Potential initiating mutations would be expected to present at all time points. One of these genes, **AFF1**, expressed in myeloma and also mutated in the MMRC cohort, was identified in our study to share the same SNV at all tumor time points. AFF1 (also referred to as AF4) is involved in RNA polymerase II mediated transcription. Specifically, it associates with positive transcription elongation factor b (P-TEFb), and ENL as part of the AEP complex (AF4/ENL/P-TEFb)\(^{31}\) that regulates transcription.\(^{32,33}\) The AEP complex co-localizes with MLL at the HOXA9 promoter.\(^{31}\) In myeloma cell lines, **HOXA9** expression was inversely related to H3K24me3 levels and cells lines depleted of **HOXA9** demonstrated a growth disadvantage.\(^{19}\) Therefore, a damaging
mutation in AFF1, which is predicted to disrupt phosphoserine sites downstream of the mutation, could result in dysregulation of HOXA9 leading to dysregulated cell growth and contributing to myelomagenesis.

Variants that do not present until the final sPCL time point are potential candidates for leukemic transformation. One of these, ZKSCAN3, has been demonstrated to modulate CCND2 in myeloma cells with increased levels of CCND2 observed with ZKSCAN3 overexpression in U266 myeloma cells.\textsuperscript{34} In contrast, the opposite is expected in the sPCL as the truncating SNV in ZKSCAN3 observed in the sPCL is predicted to result in a loss of function protein. Thus, in the absence of functional ZKSCAN3 protein, we would expect decreased levels of CCND2 with subsequent increased levels of other cyclin D proteins potentially leading to dysregulation of the G1/S checkpoint.

Another gene with a truncating variant that does not present until sPCL is \textit{RB1}, a critical regulator of the cell cycle. Monoallelic deletion of \textit{RB1} has been implicated in myelomagenesis\textsuperscript{11-13} and was present in the chromosome 13 deletion detected at diagnosis in this patient. The SNV in the remaining \textit{RB1} allele of this patient is predicted to result in a loss of function protein, as all critical binding and interaction domains are lost. Consequently, the functional loss of RB1 in the remaining allele is significant as the resulting protein is expected to be unable to regulate the critical G1/S cell cycle checkpoint. Furthermore, a biallelic deletion in \textit{TP53}, a key tumor suppressor that regulates cell cycle arrest, was also identified only in the sPCL sample resulting in the absence of two key regulators of the cell cycle, RB1 and TP53. Taken together, truncating mutations in both ZKSCAN3 and \textit{RB1} together with deletion of \textit{TP53}
could result in catastrophic dysregulation of cell cycle checkpoints. This combined with a mutation in \textit{AFF1} could confer the growth advantage necessary for leukemic transformation from myeloma to sPCL.

The presence of variants detectable only at alternating time points suggests that all clonal progenitors were present at diagnosis, but selection pressures from treatment and clonal evolution caused dominance of these clones to rise and fall over time. Our work suggests the diagnostic and the second relapse clones demonstrate little divergence over time as only one SNV differentiates the two clones. In contrast, the first relapse and the sPCL clones have six and five unique SNV respectively, a total of 11 SNV distinguishing the two clones, suggesting greater evolutionary divergence with time and aggressiveness of disease. Taken together, this suggests common initiating events represented by the 10 shared SNV, but divergence in clonal evolution with at least two dominant clones at diagnosis that then evolve in clonal tides detected only at alternating time points. This suggests a selective pressure of therapeutic drugs. Of interest, correlation of genomic change with therapy suggests that the dominant diagnostic clone was sensitive to lenalidomide but as that clone declined, a new resistant clone emerged resulting in the first relapse. Furthermore, the first relapse clone was responsive to carfilzomib but not the original diagnostic clone that re-emerged. In addition, the presence of unique SNV at the final sPCL time point that were undetectable at earlier time points suggests tumor evolution that involved the acquisition of novel mutations that enabled leukemic transformation to occur. In addition, myeloma can present as multiple, discrete foci that are not adequately represented in serial iliac crest bone marrow samplings; thus, the clonal heterogeneity of myeloma may in fact be even greater than that reported here.
While the findings reported here are limited to a single case study, they are the first to study genomic evolution of a myeloma over time and demonstrate tumor heterogeneity and clonal dynamics of myeloma. Furthermore, these findings suggest possible driver mutations that may contribute to myelomagenesis and to leukemic transformation. In order to further understand the natural history and progression of myeloma to sPCL, additional studies in larger cohorts of serially collected patient samples are warranted.
ACKNOWLEDGEMENTS

We would like to thank Greg Ahmann for his assistance with processing samples for this project. We would like to thank GlaxoSmithKline for the use of the Cancer Cell Line Genomic Profiling Dataset. We are indebted to the Translational Genomics Research Institute High-performance Bio-computing division for making available the supercomputing resources funded by NIH grants 1S10RR25056-01 and 1S10RR023390-01. This work was supported by NIH grant CA133115-01 and Mayo Clinic Cancer Center.

AUTHORSHIP CONTRIBUTIONS:

JBE conducted sequencing, analysis and manuscript preparation. WT, AC, AK, SS, SM, YA conducted sequencing alignment, and post-alignment analysis. CXS, JS, EB, JJK conducted analysis and validation. DWC participated in analysis and manuscript preparation. RF and PLB participated in manuscript preparation. JDC and AKS designed the study, participated in analysis and manuscript preparation.

DISCLOSURE OF CONFLICTS OF INTEREST

RF has received a patent for the prognostication of MM based on genetic categorization of the disease. He has received consulting fees from Medtronic, Otsuka, Celgene, Genzyme, BMS, Lilly, Millenium and AMGEN. He also has sponsored research from Cylene and Onyx.

The other authors do not have conflicts of interest to disclose.
REFERENCES


Table 1. Summary of whole genome sequencing metrics

<table>
<thead>
<tr>
<th>Metric</th>
<th>Normal</th>
<th>Diagnostic</th>
<th>Relapse 1</th>
<th>Relapse 2</th>
<th>PCL†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total reads</td>
<td>2742938922</td>
<td>2171920328</td>
<td>2238325993</td>
<td>2662347429</td>
<td>1039706340</td>
</tr>
<tr>
<td>Aligned reads</td>
<td>2035868820</td>
<td>1551256330</td>
<td>1577742203</td>
<td>1878505981</td>
<td>968454017</td>
</tr>
<tr>
<td>Percent aligned</td>
<td>74.2%</td>
<td>71.4%</td>
<td>70.5%</td>
<td>70.6%</td>
<td>93.1%</td>
</tr>
<tr>
<td>Coverage depth (raw)</td>
<td>34</td>
<td>26</td>
<td>27</td>
<td>32</td>
<td>33</td>
</tr>
<tr>
<td>Coverage depth (duplicates removed)</td>
<td>24</td>
<td>17</td>
<td>18</td>
<td>23</td>
<td>--</td>
</tr>
<tr>
<td>Total SNV</td>
<td>2646506</td>
<td>2294806</td>
<td>2313110</td>
<td>2523352</td>
<td>3540671</td>
</tr>
<tr>
<td>Somatic SNV</td>
<td>---</td>
<td>14550</td>
<td>14226</td>
<td>16090</td>
<td>991172</td>
</tr>
<tr>
<td>Somatic SNV unique to sample</td>
<td>---</td>
<td>9426</td>
<td>9089</td>
<td>10088</td>
<td>983125</td>
</tr>
<tr>
<td>Somatic SNV/MB</td>
<td>---</td>
<td>4.72</td>
<td>4.62</td>
<td>5.22</td>
<td>321.60</td>
</tr>
<tr>
<td>Small somatic indels</td>
<td>---</td>
<td>2828</td>
<td>2592</td>
<td>2230</td>
<td>548404</td>
</tr>
</tbody>
</table>

* SOLiD 50 bp single end reads
† Illumina 101 bp paired end reads
Table 2. Summary of somatic SNV validated by capillary sequencing

<table>
<thead>
<tr>
<th>Samples mutation in</th>
<th>Chromosome Position (NCBI36)</th>
<th>Allele change</th>
<th>Amino acid change</th>
<th>Gene</th>
<th>Functional prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>98363078 G &gt; T</td>
<td>R75L</td>
<td>CNGA3</td>
<td></td>
<td>Benign</td>
</tr>
<tr>
<td>3</td>
<td>65431115 G &gt; A</td>
<td>T281M</td>
<td>MAGI1</td>
<td></td>
<td>Damaging</td>
</tr>
<tr>
<td>4</td>
<td>20461383 C &gt; T</td>
<td>V32M isoform 5</td>
<td>KCNIP4</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>4</td>
<td>88187220 G &gt; A</td>
<td>R163H</td>
<td>AFF1</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>5</td>
<td>158567354 A &gt; G</td>
<td>L11S</td>
<td>RNF145</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>6</td>
<td>16435083 G &gt; T</td>
<td>Q480K</td>
<td>ATXN1</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>6</td>
<td>31658110 C &gt; G</td>
<td>A22P</td>
<td>LTBP</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>8</td>
<td>11343928 C &gt; T</td>
<td>D2050N isoform 1</td>
<td>CSMD3</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>9</td>
<td>12688569 T &gt; G</td>
<td>L276R</td>
<td>TYRP1</td>
<td></td>
<td>Benign</td>
</tr>
<tr>
<td>10</td>
<td>69552003 C &gt; T</td>
<td>P268S</td>
<td>MYPN</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>11</td>
<td>40094082 T &gt; C</td>
<td>R113G</td>
<td>LRRRC4</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>12</td>
<td>46654167 G &gt; A</td>
<td>T1361M isoform 2</td>
<td>COL2A1</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>16</td>
<td>30106313 T &gt; A</td>
<td>L249Q</td>
<td>CORO1A</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>17</td>
<td>35541834 C &gt; T</td>
<td>S482L</td>
<td>MSL1</td>
<td></td>
<td>Benign</td>
</tr>
<tr>
<td>17</td>
<td>36978007 T &gt; C</td>
<td>I443V</td>
<td>KRT9</td>
<td></td>
<td>Benign</td>
</tr>
<tr>
<td>18</td>
<td>12300979 A &gt; T</td>
<td>L68F</td>
<td>TUBB6</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>18</td>
<td>26977706 C &gt; T</td>
<td>G329D</td>
<td>DSC</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>19</td>
<td>6257886 G &gt; A</td>
<td>R26Q</td>
<td>DOK</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>19</td>
<td>7034018 T &gt; C</td>
<td>G179W isoform b</td>
<td>G186W isoform a</td>
<td></td>
<td>Unknown</td>
</tr>
<tr>
<td>19</td>
<td>32634825 G &gt; A</td>
<td>E28*</td>
<td>ZKSCAN3</td>
<td></td>
<td>Truncating</td>
</tr>
<tr>
<td>10</td>
<td>84006 C &gt; A</td>
<td>G109V</td>
<td>TUBB8</td>
<td></td>
<td>Damage</td>
</tr>
<tr>
<td>13</td>
<td>47835092 G &gt; T</td>
<td>E28*</td>
<td>RB1</td>
<td></td>
<td>Truncating</td>
</tr>
<tr>
<td>17</td>
<td>73731163 A &gt; T</td>
<td>R82*</td>
<td>BIRC5</td>
<td></td>
<td>Truncating</td>
</tr>
<tr>
<td>18</td>
<td>21059590 C &gt; T</td>
<td>E764K</td>
<td>ZNF521</td>
<td></td>
<td>Damage</td>
</tr>
</tbody>
</table>

* indicates presence of a STOP codon
† indicates genes identified as mutated in the MMRC cohort
‡ indicates genes identified as having confirmed somatic mutations in the COSMIC database

For personal use only. From www.bloodjournal.org by guest on January 3, 2018. For personal use only.
FIGURE LEGENDS

Figure 1. Frequency of somatic mutations in genes of interest within other tumor cohorts
Genes containing SNV identified in this case study patient were compared against the MMRC cohort of 38 myeloma genomes and the COSMIC database of other tumor types to determine the frequency of all somatic mutations in these genes within larger populations. This graph illustrates the frequency (%) of all somatic mutations in the MMRC cohort and in other tumors present in the COSMIC database. Frequencies with less than 0.5% were counted as absent and were not included in the graph.

Figure 2. Clonal divergence of validated variant alleles
Graph illustrating the presence of shared and unique SNV at each tumor time point. There are 15 variants common to all time points and shared by a common ancestor. Six variants are common to only the diagnostic and second relapse while no variants are common to the first relapse and sPCL. The greatest divergence is observed between the first relapse and sPCL, with seven unique variants detected in each sample.

Figure 3. Circos plot depicting the summary of structural variation in the sPCL
The center, line plots indicate the presence of discordant read pairs. The middle ring contains the array CGH plot and the outermost ring indicates mutated or deleted genes.
Figure 1

![Graph showing frequency of MMRC and COSMIC confirmed somatic mutations across various genes.]

Figure 2

![Diagram illustrating the genetic changes from normal cell to first, second, and third relapse.]

- **15 shared SNV (all tumors):** AFF1, ADK1, COL2A1, CORD1A, CNGA3, CSMD3, DFB, MAG2, MEL1, NCMP, PKR1, LRRC4C, MTP1P, BNF145, FYP1
- **6 shared SNV (Diagnostic & second relapse):** ACER1, C10orf42, DOK5, PVRG38, PFFBP1, ZNF557
- **1 unique SNV (Diagnostic only):** PDE4DIP
- **7 unique SNV (First relapse):** ADK1, CACNA1S, DSC1, KDM5P, PVRG2, TUB, TUBB6
- **7 unique SNV (Third relapse):** BRICS, BRK, SUB1, TNP, TUBB5, ZNF543, ZNF527
Figure 3
Whole genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution and clonal tides

Jan B. Egan, Chang-Xin Shi, Waibhav Tembe, Alexis Christoforides, Ahmet Kurdoglu, Shripad Sinari, Sumit Middha, Yan Asmann, Jessica Schmidt, Esteban Braggio, Jonathan J. Keats, Rafael Fonseca, P. Leif Bergsagel, David W. Craig, John D. Carpten and A. Keith Stewart