Candidate susceptibility variants in angioimmunoblastic T-cell lymphoma

Donner, Iikki

2019-01

Donner, I , Katainen , R , Kaasinen , E , Aavikko , M , Sipilä , L J , Pukkala , E & Aaltonen ,
L A 2019 , ' Candidate susceptibility variants in angioimmunoblastic T-cell lymphoma ',

http://hdl.handle.net/10138/298283
https://doi.org/10.1007/s10689-018-0099-x

Downloaded from Helda, University of Helsinki institutional repository.
This is an electronic reprint of the original article.
This reprint may differ from the original in pagination and typographic detail.
Please cite the original version.
Candidate susceptibility variants in angioimmunoblastic T-cell lymphoma

Iikki Donner1,2 · Riku Katainen1,2 · Eevi Kaasinen1,2,3 · Mervi Aavikko1,2 · Lauri J. Sipilä1,2 · Eero Pukkala4,5 · Lauri A. Aaltonen1,2

Published online: 10 August 2018 © Springer Nature B.V. 2018

Abstract

Angioimmunoblastic T-cell lymphoma (AITL) is a subtype of peripheral T-cell lymphoma with a poor prognosis: the 5-year survival rate is approximately 30%. Somatic driver mutations have been found in TET2, IDH2, DNMT3A, RHOA, FYN, PLCG1, and CD28, whereas germline susceptibility to AITL has to our knowledge not been studied. The homogenous Finnish population is well suited for studies on genetic predisposition. Here, we performed an exome-wide rare variant analysis in 23 AITL patients. No germline mutations were found in the driver genes, implying that they are not frequently involved in genetic AITL predisposition. Potentially pathogenic variants present in at least two patients and showing significant (p < 0.01) enrichment in our sample set were found in ten genes: POLK, PRKCB, ZNF676, PRRC2B, PCDHGB6, GNL3L, TTC36, OTOG, OSGEPL1, and RASSF9. The most significantly enriched variants, causing p.Lys469Ter in a splice variant of POLK and p.Pro588His in PRKCB, are intriguing candidates as Polk deficient mice display a spontaneous mutator phenotype, whereas PRKCB was recently shown to be somatically mutated in 33% of another peripheral T-cell lymphoma, adult T-cell lymphoma. If validated, our findings would provide new insight into the pathogenesis of AITL, as well as tools for early detection in susceptible individuals.

Introduction

Angioimmunoblastic T-cell lymphoma (AITL) is a subtype of peripheral T-cell lymphoma (PTCL). It accounts for approximately 1–2% of malignant lymphomas and nearly 20% of PTCL [1]. AITL typically afflicts individuals of advanced age (median age at diagnosis 65) and has a poor prognosis with a 5-year survival rate of approximately 30%. The malignant cells originate from T follicular helper (TFH) cells and preserve the major functions of this cell type. In recent years, a number of nodal PTCLs previously classified as PTCL, not otherwise specified (NOS), have been found to have TFH-cell features. Follicular T-cell lymphoma, for example, shows genetic overlap with AITL, which has led to these two being unified under the heading “AITL and other nodal lymphomas of TFH cell origin”, in the recently revised WHO Classification of Tumours of the Haematopoietic and Lymphoid Tissues [2].

The majority of the tumor cells in AITL can be found in lymph nodes, liver, spleen, skin, and bone marrow. The malignancy is characterized by an extensive surrounding tumor microenvironment populated by various reactive cell types such as T- and B-cells, B immunoblasts, plasma cells, eosinophils and macrophages. Tumor cells are often in the minority, comprising as little as 10% of an AITL lesion [3]. Epstein-Barr virus (EBV) is a potentially oncogenic virus associated with both lymphoproliferative disorders and solid tumors. It has been linked in particular to endemic Burkitt’s lymphoma, a pediatric cancer in which EBV coinfection with malaria in early life is a major contributing factor [4].
The majority of AITL cases are also positive for EBV; however, it is rarely the malignant T cells that are infected, but instead the reactive B cells [5, 6].

Recurrent somatic mutations in the epigenetic regulators $TET2$ (47–76% of cases), $IDH2$ (20–45% of cases) and $DNMT3A$ (33% of cases) have been detected in AITL, indicating that epigenetic aberrations are a major driver of AITL lymphomagenesis [7–10]. Mutations in $IDH2$ are restricted to Arg172, and this mutation occurs almost exclusively in AITL. Mutations in $TET2$, $IDH2$ and $DNMT3A$ have been shown to co-occur in AITL, whereas they are mutually exclusive in acute myeloid leukemia. It is noteworthy that the $DNMT3A$ and $TET2$ mutations found in AITL and other hematopoietic malignancies are commonly detected in a clonal subpopulation of blood cells derived from a single progenitor with a growth advantage - a phenomenon termed clonal hematopoiesis - in elderly individuals. Due to a high prevalence of clonal hematopoiesis in healthy elderly individuals, additional genetic or epigenetic changes are likely required to drive lymphomagenesis [11].

Nearly 70% of AITL cases have the specific p.Gly17Val mutation in the small GTPase $RHOA$ [12, 13]. Alterations in diverse T-cell receptor signaling genes ($PLCG1$, $CD28$, $PIK3$ elements, $GTF2I$, $CTNNB1$) were found in approximately half the AITL cases in a study by Vallois et al. 2016. The same study confirmed a high (60%) prevalence of p.Gly17Val substitutions in $RHOA$ in AITL, and also reported a novel substitution p.Lys18Asn in three patients. Recurring mutations in $CD28$ in AITL have been described earlier by Rohr et al. 2016 [15].

To our knowledge, genetic predisposition to AITL has never been studied. In this study we sequenced the germline exomes of 23 AITL patients. We looked for shared rare variants significantly enriched in our sample set when compared to population specific controls, and for germline mutations in genes known to be somatically mutated in AITL.

Materials and methods

Ethics statement

This study was approved by the National Supervisory Authority for Welfare and Health (Valvira; 1423/06.01.03.01/2012), National Institute for Health and Welfare (THL; 151/5.05.00/2017), and the ethics committee of the Hospital District of Helsinki and Uusimaa (HUS; 408/13/03/03/09).

Patients and samples

We utilized the Finnish Cancer Registry (FCR), which was established in 1952 and harbors data on all cancer cases diagnosed in Finland since 1953 [16], to find patients with AITL. Altogether 136 patients had been diagnosed with this malignancy between 1953 and 2015. Non-malignant formalin-fixed paraffin-embedded (FFPE) tissue material was collected from 43 patients. In many cases tissue material was too scarce to obtain adequate amounts of DNA for successful sample preparation. In the end, the exomes of 23 patients (15 male, 8 female) were successfully sequenced. Median age at diagnosis of these patients was 57 years (range 30–84). Sex and age at diagnosis of the patients are listed in Table 1.

During the revision process of this manuscript, we found that one of the patients, AITL42, had mistakenly been listed as having an AITL diagnosis in the FCR. The correct diagnosis was PTCL-NOS, which at the time included also TFH phenotype malignancies. Based on patient records the exact phenotype of the patient’s malignancy was never determined, nor was AITL excluded. Whether the malignant cells had TFH features had not been determined. Because of genetic overlap between these two lymphomas, we decided to retain her data in the study.

Exome sequencing

Genomic DNA was extracted from archival FFPE tissue samples with the phenol–chloroform method and prepared for exome sequencing with the KAPA Hyper Prep Kit (Kapa Biosystems Inc., Wilmington, MA) and SeqCap EZ System (Roche Nimblegen Inc., Madison, WI). We used 300–1000 ng of DNA sheared to an average size of 200–300 bp with the Covaris S220 Focused-ultrasonicator (Covaris Inc., Woburn, MA). For target enrichment we used the 96 Mb SeqCap EZ Exome + UTR (Roche Nimblegen Inc.) kit. Paired-end sequencing was performed with Illumina HiSeq 4000 (Illumina, San Diego, CA) at Karolinska High Throughput Center. Average coverage across all samples was 38 reads, and 87% of targeted areas were covered by ten or more reads.

Data analysis

For a brief description of the pipeline used on the raw sequencing data, please refer to Donner et al. 2017 [17]. Annotation, comparative variant analysis, controlling, and variant visualization were performed with BasePlayer (https://doi.org/10.1101/126482). Gene annotation was based on Ensembl GRCh37 release 84. Minimum coverage was set at ten reads and minimum allelic fraction at 30%. We used two different variant control data sets for MAF filtering: gnomAD (full set) [18] and the gnomAD Finnish subset (n = 12,897). We looked for rare, likely pathogenic variants in genes known to be recurrently somatically mutated in AITL: $TET2$, $DNMT3A$, $IDH2$, $CD28$, $FYN$, $PLCG1$, and $RHOA$. The variants were required to have a MAF < 0.01.
in the two gnomAD sets used. Variants in other genes were required to be present in at least two patients, have a MAF of < 0.01 in the two control sets, and to be significantly (p < 0.01) enriched in our sample set as determined by Fisher’s exact test. Odds ratios and confidence intervals were calculated using conditional Maximum Likelihood Estimate. Effects of missense variants as well as in-frame deletions and insertions were predicted with Provean, SIFT, and PolyPhen-2 [19–21]. Variants were required to be predicted deleterious using the available algorithms’ default threshold values in order to pass filtering.

**Validation by Sanger sequencing**

Candidate variants were validated by Sanger sequencing of DNA extracted from FFPE blocks with the GeneRead DNA FFPE Kit (Qiagen). Primers were designed with Primer3web version 4.0.0 (http://primer3.ut.ee/) and capillary sequencing was performed at the Institute for Molecular Medicine Finland (FIMM) using the BigDye v.3.1 sequencing reaction and ABI3730xl DNA Analyzer electrophoresis (Applied Biosystems, Foster City, CA). Each reaction was performed in three replicates in order to exclude monallelic amplification.

**Results**

We were not able to detect any rare, likely pathogenic variants in genes frequently somatically mutated in AITL: TET2, DNMT3A, IDH2, CD28, PLCG1, FYN, and RHOA. Potentially pathogenic variants present in at least two patients and showing significant enrichment in our sample set were found in ten genes: POLK, PRKCB, ZNF676, PRRC2B, PCDHGB6, GNL3L, TTC36, OTOG, OSGEPL1, and RASSF9. The variants are listed in Table 2 in descending order of their odds ratio. Variants found in PRKCB and OSGEPL1 were both shared by three patients, and variants in the other eight genes were shared by two.

**Discussion**

The Finnish population has reduced genetic heterogeneity and publicly available population control genotypes are plentiful, which makes this population well suited for studies on genetic predisposition. Here, we set out to study genetic susceptibility to AITL.

Genes known to be somatically mutated in AITLs, listed in Swerdlow et al. [2], did not harbor any rare potentially pathogenic germline point mutations or INDELs,
Table 2 Rare (MAF < 0.01) variants significantly enriched among the patients included in this study

<table>
<thead>
<tr>
<th>Position</th>
<th>Base change</th>
<th>rs-code</th>
<th>Gene</th>
<th>Transcriptb</th>
<th>Effect</th>
<th>Gnomad MAF, Finns</th>
<th>OR (95% CI low–high)</th>
<th>p value</th>
<th>Gnomad MAF, all</th>
<th>OR (95% CI low–high)</th>
<th>p value</th>
<th>Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>5:74886314</td>
<td>A/T</td>
<td>rs371305423</td>
<td>POLK</td>
<td>ENST00000515295*</td>
<td>p.Lys469Ter</td>
<td>1.2E–4</td>
<td>368 (30–3683)</td>
<td>3.4E–5</td>
<td>1.0E–4</td>
<td>446 (49–1836)</td>
<td>1.2E–5</td>
<td>AITL7, AITL30</td>
</tr>
<tr>
<td>16:24202451</td>
<td>C/A</td>
<td>rs35631544</td>
<td>PRKCB</td>
<td>ENST00000303531</td>
<td>p.Pro588His</td>
<td>1.3E–3</td>
<td>54 (10–184)</td>
<td>3.6E–5</td>
<td>8.3E–4</td>
<td>81 (16–254)</td>
<td>9.7E–6</td>
<td>AITL24</td>
</tr>
<tr>
<td>19:22363095</td>
<td>T/C</td>
<td>rs201369122</td>
<td>ZNF676</td>
<td>ENST00000397121</td>
<td>p.Lys475Arg</td>
<td>8.5E–4</td>
<td>53 (6–226)</td>
<td>8.4E–4</td>
<td>1.9E–3</td>
<td>23 (3–90)</td>
<td>3.7E–3</td>
<td>AITL14, AITL36</td>
</tr>
<tr>
<td>5:140789505</td>
<td>C/T</td>
<td>rs754052398</td>
<td>PCDHGB6</td>
<td>ENST00000520790</td>
<td>p.Pro579Leu</td>
<td>1.4E–3</td>
<td>31 (3–126)</td>
<td>2.3E–3</td>
<td>2.2E–4</td>
<td>200 (23–778)</td>
<td>5.6E–5</td>
<td>AITL15, AITL17</td>
</tr>
<tr>
<td>X:54566662</td>
<td>TAAA/T</td>
<td>rs748797074</td>
<td>GNL3L</td>
<td>ENST00000336470</td>
<td>p.Lys62del</td>
<td>2.6E–3</td>
<td>29 (5–95)</td>
<td>2.3E–4</td>
<td>5.9E–3</td>
<td>12 (2–37)</td>
<td>2.7E–3</td>
<td>AITL24</td>
</tr>
<tr>
<td>11:118399392</td>
<td>G/A</td>
<td>rs118041771</td>
<td>TTC36</td>
<td>ENST00000302783</td>
<td>p.Gly65Arg</td>
<td>1.9E–3</td>
<td>24 (3–98)</td>
<td>3.6E–3</td>
<td>2.2E–3</td>
<td>20 (2–78)</td>
<td>4.9E–3</td>
<td>AITL18, AITL22</td>
</tr>
<tr>
<td>11:17621218</td>
<td>C/T</td>
<td>rs117005078</td>
<td>OTOG</td>
<td>ENST00000399391</td>
<td>p.Pro1240Leu</td>
<td>1.9E–3</td>
<td>24 (2–98)</td>
<td>3.7E–3</td>
<td>2.9E–3</td>
<td>16 (2–61)</td>
<td>8.0E–3</td>
<td>AITL30, AITL35</td>
</tr>
<tr>
<td>2:190617411</td>
<td>T/TA</td>
<td>rs568165204</td>
<td>OSGEPL1</td>
<td>ENST00000261451</td>
<td>p.Leu378Phes-Ter3</td>
<td>3.1E–3</td>
<td>22 (4–72)</td>
<td>4.4E–4</td>
<td>2.9E–3</td>
<td>24 (5–75)</td>
<td>3.4E–4</td>
<td>AITL31, AITL39, AITL45</td>
</tr>
<tr>
<td>12:86199409</td>
<td>G/A</td>
<td>rs191094027</td>
<td>RASSF9</td>
<td>ENST00000361228</td>
<td>p.Arg127Trp</td>
<td>2.5E–3</td>
<td>18 (2–72)</td>
<td>6.4E–3</td>
<td>4.4E–4</td>
<td>102 (12–402)</td>
<td>2.1E–4</td>
<td>AITL8, AITL24</td>
</tr>
</tbody>
</table>

MAF minor allele frequency, OR odds ratio, CI confidence interval, hom homozygous

aGenome assembly GRCh37
bCanonical transcript, except for *
implying that they are infrequently involved in susceptibility to the disease. However, larger structural variation could not be looked into due to the limitations of exome data. We also looked for enrichment of rare potentially pathogenic variants throughout the exome. The most significantly enriched variant in our sample set was a nonsense change p.Lys469Ter in a non-canonical splice variant (ENST00000515295) of POLK, shared by two patients. The isoform differs from the canonical sequence at amino acids 453–472, and lacks the canonical amino acids 473–870. Although the structure and function of the main isoform of POLK are known, the role of isoform 2, which is affected in our patients, is unknown. POLK is one of four Y-family bypass polymerases that can replicate through damaged DNA. High-fidelity DNA polymerases are usually blocked by bulky DNA lesions, but cancate through damaged DNA. High-fidelity DNA polymerases 473–870. Although the structure and function of the main isoform of POLK are known, the role of isoform 2, which is affected in our patients, is unknown. POLK is one of four Y-family bypass polymerases that can replicate through damaged DNA. High-fidelity DNA polymerases are usually blocked by bulky DNA lesions, but can replace or replace by one or more bypass polymerases for local translesion synthesis (TLS). POLK is involved in error-free TLS opposite N2-guanyl DNA adducts induced by carcinogens such as polycyclic aromatic hydrocarbons [22]. While TLS by the appropriate bypass polymerase is largely error-free, its absence forces other similar enzymes to take its place to rescue cells from the lethal consequences of arrested DNA replication. However, these enzymes are not optimal, and often generate an increased mutational burden, as is the case if the appropriate enzyme is functionally impaired. Accordingly, mice defective for the Polk gene (Polk−/−) have a spontaneous mutator phenotype [23]. It is thus conceivable that certain germline POLK variants might alter the functionality of the enzyme and modify individual risk of mutation and cancer in the presence of carcinogens. While POLK is an intriguing candidate gene, the variant shared by our patients alters the isoform just three amino acids prematurely. Neither somatic nor germline mutations in POLK have to our knowledge been reported in AITL or other T-cell lymphomas.

The second most significantly enriched variant, present in three patients, was a missense substitution p.Pro588His in the catalytic kinase domain (amino acids 342–600) of a protein kinase C (PKC) family member PRKCB. Chronic lymphocytic leukemia cells have been shown to overexpress PRKCB [24]. Also, more recent research on adult T-cell lymphoma (ATL) found the gene to be somatically mutated in 33% of cases [25]. Ninety-three percent of these mutations were confined to highly conserved regions within the catalytic domain, suggesting gain-of-function and thus an oncogenic nature in ATL. The PKC family has historically been considered to consist of oncoproteins. Recently, however, they have been suggested to have a tumor suppressive rather than oncogenic role [26]. One of the patients harboring the p.Pro588His variant was found to have a PTCL-NOS rather than an AITL diagnosis during the revision process of this manuscript. The variant with the third highest odd ratio was a missense variant causing p.Lys475Arg in ZNF676. Polymorphisms in ZNF676 have been associated with leukocyte telomere length variation in two separate studies [27, 28]. The function of the gene is unknown.

Two genes, GNL3L and RASSF9, with predicted pathogenic germline variants in two AITL patients each, had known functions related to proliferation. GNL3L is a putative nucleolar GTPase required for cell proliferation and normal processing of ribosomal pre-rRNA [29]. It increases homodimerization, protein stability, and telomeric association of TRF1, a protein that negatively controls telomeric length by denying telomerase access to the telomeres [30]. The TRF1-stabilizing function of GNL3L causes a mitotic increase in TRF1, and promotes the transition from metaphase to anaphase during mitosis. GNL3L has also been suggested to participate in maintaining the function of so-called tumor initiating cells [31], and in the regulation of cell proliferation during tumorogenesis by modulation of NF-κB signaling [32]. RASSF9 is associated with endosomes and is linked to vesicle trafficking [33]. It binds the Ras family proteins N-Ras, K-Ras, and R-Ras [34]. RASSF9 null mice exhibit severe growth retardation and short lifespan [35]. RASSF9 is predominantly expressed in epidermal keratinocytes of skin, and null mice also display a change in epidermal organization of skin, with increased proliferation and aberrant differentiation.

OTOG, with shared variants in two AITL patients, encodes otogelin, a glycoprotein specific to acellular membranes of the inner ear [36]. Mutations in the gene have been found to cause autosomal recessive non-syndromic hearing impairment [37, 38]. Based on this function, it appears unlikely that the variant reported here has a role in AITL lymphomagenesis. The function of the remaining genes with shared variants in our sample set: PRRC2B, PCDHGB6, TTC36, and OSGEP1L1, have to date been poorly described.

Of the patients, 16 had been tested for EBV infection. Of these ten were found to be positive, which is a somewhat low fraction compared to previous reports [6]. Considering the role that EBV infection together with malaria plays in Burkitt’s lymphoma, it is intriguing to speculate that infection together with a specific genotype could constitute a risk factor for AITL. However, rather than contributing to the disease itself, it is possible that EBV is reactivated due to AITL induced immunodeficiency [6]. Another possibility is that EBV infection plays an important role in the development of the extensive microenvironment in AITL. To robustly link EBV infection causally to AITL pathology requires future studies with large sample sets.

Based on known function and previous studies, the most intriguing variants described here are the ones found in POLK and PRKCB. As a limitation of the work, we only looked at rare variants and AITL susceptibility could also
emerge through more common variants that would have been missed by our screen. While emanating from a population-based source our final sample set was relatively small, and incidental enrichment of the variants cannot be excluded. If validated in independent sample sets, the findings would provide important new insight into the molecular pathogenesis of AITL, as well as tools for early detection of the condition in susceptible individuals.

Acknowledgements This work was supported by the Academy of Finland, Finnish Center of Excellence Program 2012–2017 [Grant Number 250345]. We wish to thank Heikki Metsola, Simi Niemiinen, Alison Ollikainen, Marjo Rajalaakso, Sirpa Soisalo, Inga-Lill Svedberg, and Lina Vuoristo for their indispensable assistance.

References