ARTICLE

Somatic mutations and T-cell clonality in patients with immunodeficiency

Paula Savola,[1,2] Timi Martelius,[3] Matti Kankainen,[1,2,4,5] Jani Huuhtanen,[1,2] Sofie Lundgren,[1,2] Yrjö Koski,[1,2] Samuli Eldfors,[4] Tiina Kelkka,[1,2] Mikko A.I. Keränen,[1,2] Pekka Ellonen,[4] Panu E. Kovanen,[6] Soili Kytölä,[7] Janna Saarela,[4] Harri Lähdesmäki,[8] Mikko R.J. Seppänen[2,3,9] and Satu Mustjoki[1,2,10]

[1]Hematology Research Unit Helsinki, University of Helsinki and Department of Hematology, HUS Helsinki University Hospital Comprehensive Cancer Center, Helsinki; [2]Translational Immunology Research Program, University of Helsinki, Helsinki; [3]Adult Immunodeficiency Unit, Infectious Diseases, Inflammation Center, University of Helsinki, HUS Helsinki University Hospital, Helsinki; [4]Institute for Molecular Medicine Finland (FIMM), HILIFE, University of Helsinki, Helsinki; [5]Medical and Clinical Genetics, University of Helsinki and HUS Helsinki University Hospital, Helsinki; [6]Department of Pathology, University of Helsinki and HUSLAB, HUS Helsinki University Hospital, Helsinki; [7]Laboratory of Genetics, HUSLAB, HUS Helsinki University Hospital, Helsinki; [8]Department of Computer Science, Aalto University School of Science, Espoo; [9]Rare Diseases Center and Pediatric Research Center, Children and Adolescents, University of Helsinki and HUS Helsinki University Hospital, Helsinki and [10]Department of Clinical Chemistry and Hematology, University of Helsinki, Helsinki, Finland

Ferrata Storti Foundation

Haematologica 2020
Volume 105(12):2757-2768

ABSTRACT

Common variable immunodeficiency (CVID) and other late-onset immunodeficiencies often co-manifest with autoimmunity and lymphoproliferation. The pathogenesis of most cases is elusive, as only a minor subset harbors known monogenic germline causes. The involvement of both B and T cells is, however, implicated. To study whether somatic mutations in CD4$^+$ and CD8$^+$ T cells associate with immunodeficiency, we recruited 17 patients and 21 healthy controls. Eight patients had late-onset CVID and nine patients other immunodeficiency and/or severe autoimmunity. In total, autoimmunity occurred in 94% and lymphoproliferation in 65%. We performed deep sequencing of 2,533 immune-associated genes from CD4$^+$ and CD8$^+$ cells. Deep T-cell receptor β-sequencing was used to characterize CD4$^+$ and CD8$^+$ T-cell receptor repertoires. The prevalence of somatic mutations was 65% in all immunodeficiency patients, 75% in CVID, and 48% in controls. Clonal hematopoiesis-associated variants in both CD4$^+$and CD8$^+$ cells occurred in 24% of immunodeficiency patients. Results demonstrated mutations in known tumor suppressors, oncogenes, and genes that are critical for immune- and proliferative functions, such as *STAT5B* (2 patients), *C5AR1* (2 patients), *KRAS* (one patient), and *NOD2* (one patient). Additionally, as a marker of T-cell receptor repertoire perturbation, CVID patients harbored increased frequencies of clones with identical complementarity determining region 3 sequences despite unique nucleotide sequences when compared to controls. In conclusion, somatic mutations in genes implicated for autoimmunity and lymphoproliferation are common in CD4$^+$ and CD8$^+$ cells of patients with immunodeficiency. They may contribute to immune dysregulation in a subset of immunodeficiency patients.

Introduction

Immunodeficiencies often manifest with autoimmune disease, as shown in the most common primary immunodeficiency in the adult population: common variable immunodeficiency (CVID).[1] Although the main manifestations of CVID are recurrent infections and low levels of plasma immunoglobulins, patients often present with immune-mediated blood-cell cytopenias, enteropathy, arthritis, lymphoproliferation, and/or granulomatous disease.[1] Currently known monogenic causes account for 2-10% of CVID cases.[2] In the absence of disease-associated

Correspondence:

SATU MUSTJOKI
satu.mustjoki@helsinki.fi

Received: March 14, 2019.
Accepted: December 18, 2019.
Pre-published: December 19, 2019.

doi:10.3324/haematol.2019.220889

©2020 Ferrata Storti Foundation

Material published in Haematologica is covered by copyright. All rights are reserved to the Ferrata Storti Foundation. Use of published material is allowed under the following terms and conditions: https://creativecommons.org/licenses/by-nc/4.0/legalcode. Copies of published material are allowed for personal or internal use. Sharing published material for non-commercial purposes is subject to the following conditions: https://creativecommons.org/licenses/by-nc/4.0/legalcode, sect. 3. Reproducing and sharing published material for commercial purposes is not allowed without permission in writing from the publisher.

germline variants, the etiology of CVID and other delayed-onset immunodeficiency remains largely elusive.

Although a B-cell defect apparently causes hypogammaglobulinemia, T-cell abnormalities also occur in CVID patients. Decreased levels of regulatory T cells and naïve $CD4^+$ T cells, an increase in $CD8^+$ expansion, disturbed cytokine secretion, and large granular lymphocyte expansions have all been described.[1,3,4] Altered immune homeostasis may promote both immunodeficiency and infections. In CVID, the prevalence of autoimmune disease is 30%,[1,3,5] lymphoproliferation 50%,[3] and patients also have an increased risk for malignant diseases. The standardized incidence ratio for lymphomas and stomach cancer is 10-12.[1,6]

Somatic mutations play a key role in malignant transformation, while recent discoveries have highlighted associations between somatic mutations and non-malignant disease.[7,8] Somatic *STAT3* mutations occur in diseases with autoimmune manifestations, such as large granular lymphocyte (LGL) leukemia and aplastic anemia.[9-12] In LGL leukemia, patients present with immune-mediated cytopenias and rheumatoid arthritis, and LGL leukemia patients with multiple *STAT3* mutations more often have RA (46%) than patients without mutations (6%).[13] Somatic, activating *STAT3* mutations have also been discovered in intraepithelial lymphocytes of refractory celiac disease patients[14] and in patients with Felty syndrome (long-lasting RA and neutropenia).[15] Recent reports by us and others have shown that somatic mutations in genes other than *STAT3* occur in mature T cells.[16,17] For example, patients with untreated RA have somatic mutations in mature $CD8^+$ T cells in 20% of cases.[16]

In addition to mutations in mature T cells, hematopoietic progenitor cells may also harbor mutations: somatic mutations originating from hematopoietic progenitor cells occur in 10% of elderly individuals, and these mutations confer a risk for blood cancer and cardiovascular disease.[18-20] These stem-cell-derived mutations have effects beyond malignant transformation. For example, somatic loss-of-function *Tet2* mutations in myeloid cells lead to proinflammatory cytokine production and increased atherosclerosis in mice.[21] As another example, somatic mutations in the FAS gene in hematopoietic precursors can lead to autoimmune lymphoproliferative syndrome (ALPS).[22]

Thus, somatic events in blood cells have the potential to modulate immune-mediated diseases. To study whether patients with late-onset CVID and other immunodeficiency and/or severe autoimmunity have somatic mutations in $CD4^+$ and $CD8^+$ cells, and whether they associate with disease phenotype, we used a customized deep sequencing panel covering 2,533 genes. In addition, to characterize the $CD4^+$and $CD8^+$ T-cell repertoire in more detail, we investigated the T-cell clone size, clonality, and characteristics by deep T-cell receptor β-chain (TCRB) sequencing.

Methods

Patients

We recruited eight patients with late-onset CVID and nine patients with other types of immunodeficiency or severe autoimmunity from the Helsinki University Hospital infectious disease clinic (Table 1). CVID patients were included in a previously described CVID cohort.[3] Five of the patient cases with other types of immune deficiency or severe autoimmunity have been previously described: patient 17 has a germline mutation in *ADA2* (p.Arg169Gln/p.Arg169Gln) and somatic *STAT3* D661V and N647I mutations in T cells;[23] patients 9-11 have germline gain-of-function mutations in *STAT3*;[24] and patient 12 has a compound immunodeficiency and arr11q24.2q25 (126,074,297-134,927,114) x 1 (Jacobsen syndrome).[25] Blood donor buffy coat samples from the Red Cross Blood Service (*Online Supplementary Table S1*) served as healthy controls. This study was approved by the local ethics committee and was conducted according to the principles of the Declaration of Helsinki. All patients signed informed consent.

Sample preparation

$CD4^+$ and $CD8^+$ cells were purified *via* positive selection with magnetic beads (Miltenyi Biotech) from peripheral blood mononuclear cells. Sample material for patient 9 was bone marrow rather than peripheral blood.

Identification of somatic mutations

To discover somatic variants, we designed a custom sequencing panel that comprised the coding areas of 2,533 genes. The gene selection was based on the InnateDB database (*http://www.innatedb.ca/*) and on other genes important for hematopoietic cells, adaptive immune responses, and autoimmunity (*Online Supplementary Table S2*). All samples ($CD4^+$ and $CD8^+$ cells of 17 immunodeficiency patients and 21 healthy controls) were sequenced with HiSeq2500 (Illumina).

Somatic variants were identified with the GATK toolkit and MuTect2 using a panel of 21 healthy controls (*Online Supplementary Methods*). To discover variants occurring only in $CD4^+$ or $CD8^+$ cells, variant calling was performed as a paired-sample analysis, using $CD4^+$ cells as a germline control and $CD8^+$ as a 'tumor' sample, and vice versa. In addition, we performed variant calling without a germline control sample from the same patient to discover variants that occur in both $CD4^+$ and $CD8^+$ cells. Variants identified by single-sample calling method were required to occur in both $CD4^+$ and $CD8^+$ cells in a known hematopoietic tumor suppressor or oncogene.[18-20] All variants were required to pass multiple quality filters (*Online Supplementary Methods*). Only mutations that altered the amino-acid coding sequence were reported, but non-coding mutations were included in mutation signature analyses. Signaling pathways were annotated using the Reactome public database and gene with Gene Ontology (GO) Consortium biological process terms (*Online Supplementary Table S3*).

T-cell receptor β-chain sequencing

T-cell receptor β-chain (TCRB) sequencing was performed from $CD8^+$ and $CD4^+$ cells' genomic DNA with the human TCRB immunoSEQ assay (Adaptive Biotechnologies) according to the manufacturer's instructions. The TCRB repertoires of seven CVID patients and 27 healthy controls were compared. Only productive T-cell receptor (TCR) sequences were analyzed. To search for public, previously reported TCR sequences associated with pathogens, autoimmunity, or cancer, we queried TCR against the manually curated McPAS database. "Convergent TCR" were defined as clones that share an amino-acid CDR3 sequence although they have distinct nucleotide CDR3 sequences. A T-cell clone was defined by a unique nucleotide sequence (*Online Supplementary Figure S1*).

Data sharing

Original sequencing data are available from the corresponding author upon reasonable request due to ethical permit constraints.

Table 1. Patient phenotypes.

ID	Germline variant?	Disease	Age at sampling	Age at dg	Sex	Phenotype	Details
1	No	CVID	61	59	M	Infections, autoimmunity, enteropathy, lymphoproliferation	Atrophic gastritis, colitis, GLILD, bronchiectasis, liver cirrhosis, splenomegaly, lymphadenopathy
2	No	CVID	67	65	M	Infections, autoimmunity, granulomatous disease, lymphoroliferation	ITP, GLILD, splenomegaly, lymphadenopathy
3	*TACI* S68fsX11	CVID	54	42	M	Infections, autoimmunity, granulomatous disease, lymphoproliferation	GLILD, bronchiectasis, cholestasis due to hepatic LGL infiltration, chronic gastritis, sicca, psoriasis, portal hypertension, LGL lymphoproliferation, thrombocytopenia, splenomegaly, lymphadenopathy
4	No exome seq	CVID	70	59	M	Infections, autoimmunity	Hyperthyroidism, asthma, chronic gastritis
5	No	CVID	70	68	F	Infections, autoimmunity, lymphoproliferation	Atrophic gastritis, psoriasis, asthma, splenomegaly
6	No exome seq	CVID	67	64	F	infections, autoimmunity	Hypothyroidism, vitamin D malabsorption
7	No	CVID	35	32	F	Infections, autoimmunity, granulomatous disease, lymphoproliferation	Neutropenia, anterior uveitis, sicca, GLILD, lymphadenopathy
8	No	CVID	37	30	M	Infections, autoimmunity, granulomatous disease, lymphoproliferation	ITP, AIHA, hyperthyroidism, sialadenitis, GLILD, splenomegaly, lymphadenopathy
9	*STAT3* K658N	STAT3 GOF	17	17	F	Autoimmunity, enteropathy, lymphoproliferation, infections	Autoimmune enteropathy, bronchiectasis, asthma, cryptogenic organizing pneumonia, dermatitis, lymphadenopathy, splenomegaly, AIHA, sicca, uveitis, sterile pleuritis, short stature, bone-marrow eosinophilia, hypogammaglobulinemia
10	*STAT3* K392R	STAT3 GOF	15	15	F	Autoimmunity, infections, lymphoproliferation	Neonatal diabetes, celiac disease, rudimentary pancreas, T-cell LGL leukemia, desquamative interstitial pneumonitis, AIHA, bone-marrow eosinophilia, severe allergy, short stature, hypogammaglobulinemia
11	*STAT3* M394T	STAT3 GOF	22	22	F	Autoimmunity, lymphoproliferation, infections	ITP, lymphocytic colitis, splenomegaly, lymphadenopathy, hypogammaglobulinemia
12	No exome seq, arr11q24.2q25 (126,074,297-134,927,114) x 1	Jacobsen sdr	46	45	F	Infections, autoimmunity, lymphoproliferation	ITP, neutropenia, hypothyroidism, condylomas, angioedema, asthma, splenomegaly, lymphadenopathy, hypogammaglobulinemia
13	No	Other	73	67	F		Hypogammaglobulinemia, recurrent stomatitis
14	No	Other	70	NA	F	Infections, autoimmunity	Recurrent EBV meningoencephalitis, palindromic seronegative arthritis
15	No	Other	60	55	M	Autoimmnunity	Periorbital adul-onset xantogranulomatosis, psoriatic arthritis, ulcerative colitis, asthma, alopecia, stroke
16	No exome seq	Good sdr	70	59	M	Infections, autoimmunity	LGL lymphoproliferation, bronchiectasis, sicca, seronegative rheumatoid arthritis/tenosynovitis, hypogammaglobulinemia
17	*ADA2* R169Q/R169Q	ADA2 def	41	41	F	Infections, autoimmunity, lymphoproliferation	ITP, AIHA, neutropenia, LGL lymphoproliferation, hypogammaglobulinemia, pulmonary hypertension, splenomegaly

Phenotypes of the patients included in the study. Eight were late-onset common variable immunodeficiency (CVID) patients, other patients had another type of immunodeficiency or severe autoimmunity. The column "Germline variant" shows the results of clinical exome sequencing: if a pathogenic variant was identified, it is shown. ID: identifier; dg: diagnosis; M: male; F: female; ITP: immune thrombocytopenic purpura; GLILD: granulomatous-lymphocytic interstitial lung disease; LGL: large granular lymphocyte; AIHA: autoimmune hemolytic anemia; EBV: Epstein-Barr virus; seq: sequencing; GOF: gain-of function; sdr: syndrome; def: deficiency; COSMIC: Catalogue of Somatic Mutations in Cancer.

Results

Patients' characteristics

All patients but one included in the study had autoimmune disease manifestations, and 11 of 17 (65%) had lymphoproliferation (Table 1). All CVID patients (patients 1-8) suffered from recurrent infections. Lymphoproliferation, as defined by lymphadenopathy and/or splenomegaly, occurred in 75% of CVID patients, granulomatous disease in 12.5%, and enteropathy in 12.5%. Because we included only late-onset CVID in the study, the median age at diagnosis was 59 years for CVID patients.

According to the EUROClass classification system, 25% (2 of 8) of CVID patients did not have B cells (B-) (*Online Supplementary Table S4*). In the remaining B^+ patients, switched memory-B cells ($CD19^+CD27^+IgD^{neg}IgM^{neg}$) were low in 50% (3 of 6) of cases. None had high frequencies of transitional B cells ($CD19^+CD38^{++}IgM^{++}$), and 67% (4 of 6) had increased $CD21^{low}$ cells ($CD19^+CD21^{low}CD38^{low}$) according to the EUROClass system. Other patients were similarly classified, although they did not have CVID (*Online Supplementary Table S4*).

Somatic mutations in $CD4^+$ and $CD8^+$ cells

The custom gene panel covered the coding areas of 2,533 genes, spanning over approximately 5.2 million base-pairs. Immunodeficiency patients harbored 45 and healthy controls 28 somatic mutations that existed in either $CD4^+$ or $CD8^+$ cells, identified by paired-sample variant calling (Table 2, Figure 1 and *Online Supplementary Table S5*). In two cases, we had access to non-hematopoietic-tissue-derived DNA; the mutations did not exist in these samples (*Online Supplementary Table S6*). The prevalence of somatic mutations was 65% (11 of 17) in immunodeficiency patients, 75% (6 of 8) in CVID, and 48% (10 of 21) in healthy controls. The median ages of patients and controls were similar (Figure 2A and *Online Supplementary Table S1*). There was no statistically significant difference in the prevalence of mutations between patients and controls (Figure 2B) or in the number of identified mutations (*Online Supplementary Figure S2*). The median number of mapped bases was higher in immunodeficiency samples, but the number of mutations did not correlate with mapped bases in $CD8^+$ cells, in which most mutations were discovered (*Online Supplementary*

Table 2. Selected somatic mutations identified in $CD4^+$ and $CD8^+$ cells.

Pt. ID	Disease	Cells	HGVS	AA change	COSMIC identifier	Gene	VAF	SIFT	Polyphen2	ExAC All
1	CVID	$CD8^+$	1:g.1804503C>G	NM_001282539:exon6:c.G346C:p.G116R		*GNB1*	0.07	D	D	NA
2	CVID	$CD4^+$	4:g.105235713delC	NM_001127208:exon3:c.1771delC:p.Q591fs		*TET2*	0.056	NA	NA	NA
		$CD4^+$	4:g.105243618G>T	NM_001127208:exon6:c.G3643T:p.E1215X	COSM3719016	*TET2*	0.053	NA	NA	NA
		$CD4^+$	16:g.50699619G>T	NM_022162:exon2c.G205T:p.E69X		*NOD2*	0.06	NA	NA	NA
		$CD4^+$	4:g.105269703C>T	NM_001127208:exon9:c.C4138T:p.H1380Y	COSM87161	*TET2*	0.027	D	D	0.000092
		$CD4^+$	17:g.42217380A>T	NM_012448:exon10:c.T1254A:p.N418K		*STAT5B*	0.036	D	D	NA
3	CVID	$CD8^+$	7:g.100677602C>G	NM_005273:exon6:c.C372G:p.Y124X		*GNB2*	0.051	NA	NA	NA
		$CD8^+$	11:g.75290010T>G	NM_004041:exon2:c.A50C:p.K17T		*ARRB1*	0.025	D	D	NA
4	CVID	$CD4^+$	9:g.5073770G>T	NM_001322195:exon13:c.G1849T:p.V617F	COSM12600	*JAK2*	0.017	D	D	0.0007
12	Jacobsen sdr	$CD4^+$	22:g.36929726C>T	NM_000395:exon6:c.C637T:p.R213W		*CSF2RB*	0.026	D	D	NA
		$CD4^+$	19:g.47319962C>T	NM_001736:exon2:c.C185T:p.T62M	COSM400357	*C5AR1*	0.02	T	D	0.000049
		$CD8^+$	X:g.21432736G>A	NM_014927:exon3:c.G353A:p.R118Q		*CNKSR2*	0.03	T	D	NA
13	Other	$CD4^+$	17:g.42210194G>C	NM_012448:exon15:c.C1883G:p.T628S	COSM6022929	*STAT5B*	0.032	D	D	0.000015
15	Other	$CD8^+$	22:g.25667739C>G	NM_005160:exon6:c.C442G:p.P148A		*GRK3*	0.034	P	P	NA
16	Good sdr	$CD8^+$	1:g.56692436C>T	NM_006252:exon4:c.C409T:p.H137Y		*PRKAA2*	0.032	D	D	NA
		$CD8^+$	12:g.25227351G>A	NM_004985:exon3:c.C173T:p.T58I	COSM87288, COSM5490513	*KRAS*	0.079	D	D	NA
		$CD8^+$	19:g.47320366C>T	NM_001736:exon2:c.C589T:p.R197W	COSM4667798	*C5AR1*	0.13	T	P	0.000033
17	*ADA2* def	$CD8^+$	5:g.138466921G>A	NM_001964:exon2:c.G472A:p.V158I		*EGR1*	0.023	D	P	NA
HC1	Healthy	$CD8^+$	9:g.35707732T>C	NM_006289:exon35:c.A4631G:p.K1544R		TLN1	0.021	T	P	NA
HC3	Healthy	$CD8^+$	15:g.41056702G>A	NM_017553:exon17:c.C1990T:p.R664C	COSM5383245, COSM5383244	*INO80*	0.068	D	D	NA
HC11	Healthy	$CD8^+$	2:g.25240306A>G	NM_175629:exon19:c.T2318C:p.L773P	COSM1583115	*DNMT3A*	0.043	D	D	NA
		$CD8^+$	9:g.130880097T>C	NM_007313:exon9:c.T1510C:p.S504P		*ABL1*	0.059	D	P	NA
HC16	Healthy	$CD8^+$	3:g.184365148T>G	NM_001278698:exon3:c.T173G:p.L58W		*POLR2H*	0.032	D	D	NA
		$CD8^+$	4:g.38797118T>C	NM_003263:exon4:c.A1714G:p.M572V		*TLR1*	0.021	T	B	NA
HC17	Healthy	$CD8^+$	5:g.148827319T>G	NM_000024:exon1:c.T488G:p.L163R		*ADRB2*	0.057	D	D	NA
		$CD8^+$	15:g.74036026C>G	NM_033247:exon5:c.C1274G:p.A425G		*PML*	0.025	T	B	NA
		$CD8^+$	16:g.31266064A>T	NM_001145808:exon5:c.A344T:p.E115V		*ITGAM*	0.05	D	P	NA
		$CD8^+$	21:g.31251838C>G	NM_003253:exon6:c.G1315C:p.A439P		*TIAM1*	0.029	T	D	NA
		$CD8^+$	21:g.31266560T>A	NM_003253:exon5:c.A413T:p.D138V		*TIAM1*	0.047	D	P	NA
HC20	Healthy	$CD8^+$	11:g.129870009T>C	NM_006165:exon22:c.A3091G:p.I1031V		*NFRKB*	0.054	T	P	NA
HC21	Healthy	$CD4^+$	1:g.114716123C>T	NM_002524:exon2:c.G38A:p.G13D	COSM573	*NRAS*	0.018	D	B	0.0000082

Selected mutations identified in healthy controls and immunodeficiency patients in $CD4^+$ and $CD8^+$ cells. The mutations were identified in paired-sample analyses in either $CD4^+$ or $CD8^+$ cells. Patient (Pt.) 17 had also somatic *STAT3* mutations in T cells (D661V and N647I), but they did not pass all MuTect2 filters due to clustered events and are thus not listed in the table. CVID: common variable immunodeficiency; ID: identifier; AA: amino acid; VAF: variant allele frequency; SIFT: SIFT prediction; PolyPhen2: Polyphen2 HDIV prediction; T: tolerated; D: deleterious (SIFT)/probably damaging (PolyPhen2); B: benign; P: possibly damaging; ExAC all: the variant frequency in the general Exome Aggregation Consortium; sdr: syndrome; def: deficiency; COSMIC: Catalogue of Somatic Mutations in Cancer.

Table S7 and *Online Supplementary Figure S3*). The vertical and horizontal coverage of the assay is shown in *Online Supplementary Figure S4*.

In immunodeficiency patients, the identified mutations comprised of 38 missense, 5 nonsense, and 2 frameshift mutations (Figure 2C). Immunodeficiency patients harbored a larger proportion of damaging mutations than healthy controls (56% *vs.* 21%, Fisher's exact test P=0.0069) (Figure 2D). For CVID patients, the proportion of damaging mutations was even higher (70%; P=0.0007 when compared with healthy controls). Damaging mutations were defined as frameshift, nonsense and as muta-

Figure 1. Discovered somatic mutations in CD4⁺ and CD8⁺ cells in patients and controls. Selected somatic mutations in patients and healthy controls identified in paired-sample variant calling analyses are shown, with different colors representing different mutation types (missense, nonsense, frameshift, splicing; see color code). Mutations were grouped according to Reactome pathways, but one gene was grouped only to one pathway for visual clarity. For immunodeficiency patients, the patient phenotypes are also shown at the top of the figure. Patients 1-8 had common variable immunodeficiency (CVID). Oncogenes and tumor suppressors relevant for hematopoietic tissue were derived from previous publications.[18-20] The mean RNA expression of the mutated genes in healthy individuals' CD4⁺ (n=3) and CD8⁺ (n=5) is shown as log2-transformed counts per million (cpm) values, and error bars represent standard deviation. B2M and ACTB are shown as highly-expressed housekeeping gene references for expression levels. HC: healthy control; ACTB: actin beta; B2M: beta-2-microglobulin; cpm: counts per million.

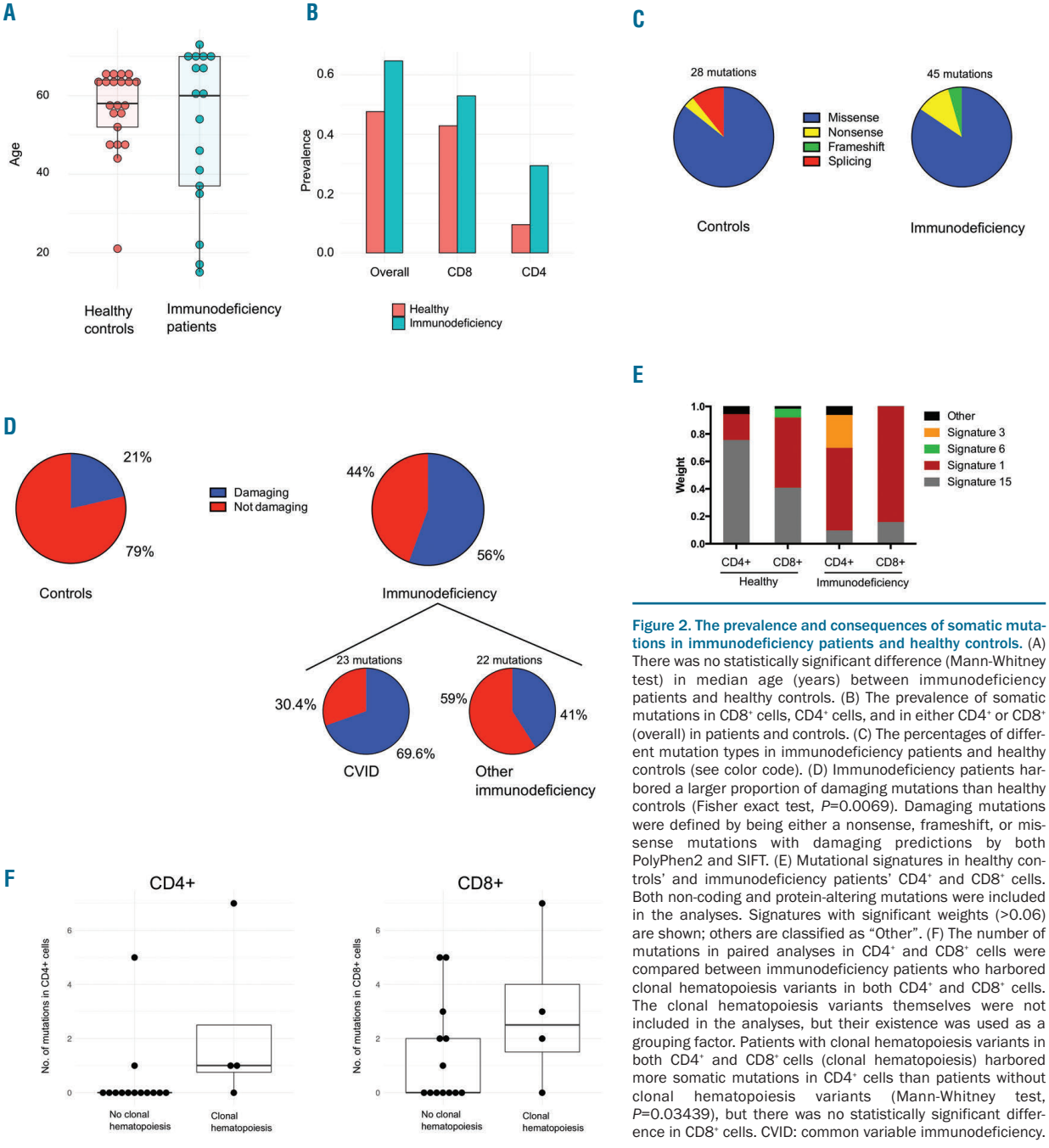

Figure 2. The prevalence and consequences of somatic mutations in immunodeficiency patients and healthy controls. (A) There was no statistically significant difference (Mann-Whitney test) in median age (years) between immunodeficiency patients and healthy controls. (B) The prevalence of somatic mutations in CD8⁺ cells, CD4⁺ cells, and in either CD4⁺ or CD8⁺ (overall) in patients and controls. (C) The percentages of different mutation types in immunodeficiency patients and healthy controls (see color code). (D) Immunodeficiency patients harbored a larger proportion of damaging mutations than healthy controls (Fisher exact test, P=0.0069). Damaging mutations were defined by being either a nonsense, frameshift, or missense mutations with damaging predictions by both PolyPhen2 and SIFT. (E) Mutational signatures in healthy controls' and immunodeficiency patients' CD4⁺ and CD8⁺ cells. Both non-coding and protein-altering mutations were included in the analyses. Signatures with significant weights (>0.06) are shown; others are classified as "Other". (F) The number of mutations in paired analyses in CD4⁺ and CD8⁺ cells were compared between immunodeficiency patients who harbored clonal hematopoiesis variants in both CD4⁺ and CD8⁺ cells. The clonal hematopoiesis variants themselves were not included in the analyses, but their existence was used as a grouping factor. Patients with clonal hematopoiesis variants in both CD4⁺ and CD8⁺ cells (clonal hematopoiesis) harbored more somatic mutations in CD4⁺ cells than patients without clonal hematopoiesis variants (Mann-Whitney test, P=0.03439), but there was no statistically significant difference in CD8⁺ cells. CVID: common variable immunodeficiency.

tions with damaging predictions by both Polyphen-2 and SIFT. Mutations in known tumor suppressors or oncogenes[18–20] occurred in six patients (35%) and two controls (9.5%) (Figure 1). Gene Ontology (GO) term annotations revealed that patients harbored variants in genes that associate with lymphocyte functions (5 patients, 29%), inflammation (7 patients, 41%), and proliferation (8 patients, 47%). Most genes that harbored mutations were expressed in healthy CD4⁺ and CD8⁺ cells (Figure 1), suggesting functional relevance. Mutational signature analysis revealed that the most dominant signature in immunodeficiency patients' CD4⁺ and CD8⁺ cells was age-related signature 1. Mismatch-repair-associated signature 15 was more dominant in healthy controls' CD4⁺ cells (Figure 2E).

Clonal hematopoiesis variants in CD4⁺ and CD8⁺ cells

To complement the findings in paired-sample analyses by discovering variants that occur in both CD4⁺ and CD8⁺ cells, we performed single-sample variant calling analysis in genes associated with clonal hematopoiesis.[18-20] In the immunodeficiency cohort, 24% of patients (4 of 17) had in total seven clonal hematopoiesis-associated variants while none was observed in healthy controls (Table 3 and Figure 1). Patients who had clonal hematopoiesis variants harbored more mutations in CD4⁺ cells than patients without clonal hematopoiesis (Figure 2F). Deep sequencing from flow cytometry-sorted cells assessed the cellular lineage of these variants. Clonal hematopoiesis-associated variants were found in sorted, pure CD3⁺ T cells, and, in some cases, in CD19⁺

Table 3. Clonal hematopoiesis variants occurring in both CD4+ and CD8+ cells.

Pt. ID	AA change	Consequence	COSMIC identifier	Cell type	VAF
13	*DNMT3A*:NM_022552:exon20:c.2388dupT:p.N797_L798delinsX	Nonsense		CD4+	0.043
				CD8+	0.038
2	*TET2*:NM_001127208:exon3:c.1771delC:p.Q591fs	Frameshift		CD4+	0.056
				CD8+	0.0065
	TET2:NM_001127208:exon6:c.G3643T:p.E1215X	Nonsense	COSM3719016	CD4+	0.053
				CD8+	0.027
	TET2:NM_001127208:exon11:c.C4889G:p.S1630X	Nonsense	COSM5945066	CD4+	0.041
				CD8+	0.032
17	*DNMT3A*: NM_022552:exon14:c.G1591A:p.D531N	Missense	COSM1583077	CD4+	0.013
				CD8+	0.024
16	*DNMT3A*: NM_022552:exon14:c.G1591A:p.D531N	Missense	COSM1583077	CD4+	0.022
				CD8+	0.027
	PRPF8:NM_006445:exon39:c.A6272G:p.Y2091C	Missense		CD4+	0.019
				CD8+	0.023

Variants that occurred in clonal hematopoiesis-associated genes and in both CD4+ and CD8+ cells in single-sample variant calling are shown. Pt. ID: patient identifier; VAF: variant allele frequency; AA: amino acid; COSMIC: Catalogue of Somatic Mutations in Cancer.

B cells. There were no mutations in CD14+ monocytes (*Online Supplementary Table S8*).

Somatic mutations in key immunologic pathways

Overall, mutations occurred in genes that are linked with cell proliferation and hematologic malignancies. Patient 2 harbored the highest number of mutations: seven in CD4+ cells and two in CD8+ cells. This patient had a *STAT5B* (N418K) mutation in CD4+ cells with 3.3% variant allele frequency (VAF), but also a NOD2 and four different *TET2* mutations in CD4+ cells (Tables 2 and 3, and Figure 1). The patient had late-onset CVID with multiple co-morbidities: infections, idiopathic thrombocytopenic purpura (ITP), granulomatous-lymphocytic interstitial lung disease (GLILD), splenomegaly, lymphadenopathy, and granulomas (Table 1).

Patient 13 also harbored a STAT5B T628S mutation with 3.2% variant allele frequency (VAF) in CD4+ cells (Table 2 and Figure 1). In addition, she had a truncating clonal hematopoiesis-associated DNMT3A variant in both CD4+ and CD8+ cells (Table 3). Her disease phenotype comprised of hypogammaglobulinemia and recurrent, non-infectious stomatitis. Patient 16 harbored seven mutations in CD8+ cells and one in CD4+ cells, occurring in genes involved in RAF/MAPK and GPCR signaling, *TP53* regulation and inflammasome activity (*KRAS* T58I, *C5AR1, PRKAA2* and *MEFV*) (Figure 1). In addition, this patient had two clonal hematopoiesis-associated variants in both CD4+ and CD8+ cells (Table 3). This patient's immunodeficiency presented as Good syndrome with LGL lymphoproliferation, bronchiectasis, sicca, seronegative RA/tenosynovitis, hypogammaglobulinemia, and infections (Table 1).

Patient 12 harbored five mutations in CD4+ cells and three in CD8+ cells. Mutations included a C5AR1 mutation, and two mutations in RAF/MAPK signaling cascade genes (*CNKSR2* and *CSF2RB*) (Table 2 and Figure 1). The patient had Jacobsen syndrome with infections, hypogammaglobulinemia, lymphoproliferation, and multiple autoimmune manifestations (Table 1).

To validate a subset of mutation findings, we performed Amplicon re-sequencing of interesting mutations (23 different positions in patients) and negative control positions (*Online Supplementary Methods* and *Online Supplementary Table S9*) with a multiplexed customized amplicon panel. All mutation positions covered by sufficient coverage (20 of 23, 87%) were successfully validated (20 of 20, 100%). As expected, mutations were only detected in the original patient and not in other individuals. All control positions yielded negative results. Mutations presented with similar VAF in re-sequencing as in the original sequencing assay (*Online Supplementary Table S10*).

CD4+ clonality correlates with CD8+ clonality and B-cell phenotypes

There was no difference in clonality indices of either CD4+ or CD8+ cells between patients and controls (Figure 3A). Overall, CD8+ cells showed higher clonality than CD4+ cells ($P<0.0001$ in healthy controls; $P=0.0135$ in CVID) (Figure 3A). The size of the largest T-cell clone (as frequency of all productive rearrangements) and the proportion of productive sequences of all TCRB rearrangements were similar in patients and controls (*Online Supplementary Figure S5*).

CD4+ and CD8+ clonality showed a positive non-linear correlation in both CVID and healthy controls (Figure 3B; Spearman correlation $P=0.0019$ and $P=0.0234$, respectively). An increased proportion of CD8+ cells was associated with higher clonality in CD4+ and CD8+ cells (Figure 3C; Spearman correlation $P=0.0238$ CD4+ clonality and 0.0341 CD8+ clonality). The largest T-cell clone size was greater in individuals with mutations than in individuals without mutations (*Online Supplementary Figure S6*). Due to the importance of helper T cells in B-cell maturation, we correlated CD4+ clonality indices with the frequencies of different B-cell subsets in CVID patients. Higher CD4+ clonality correlated with lower frequencies of memory and switched memory B cells ($P=0.008$ and $P=0.0238$) (Figure 3D and E). CD4+ clonality did not associate with the frequency of total B cells, marginal-zone, activated, or transitional B cells (*Online Supplementary Figure S7*).

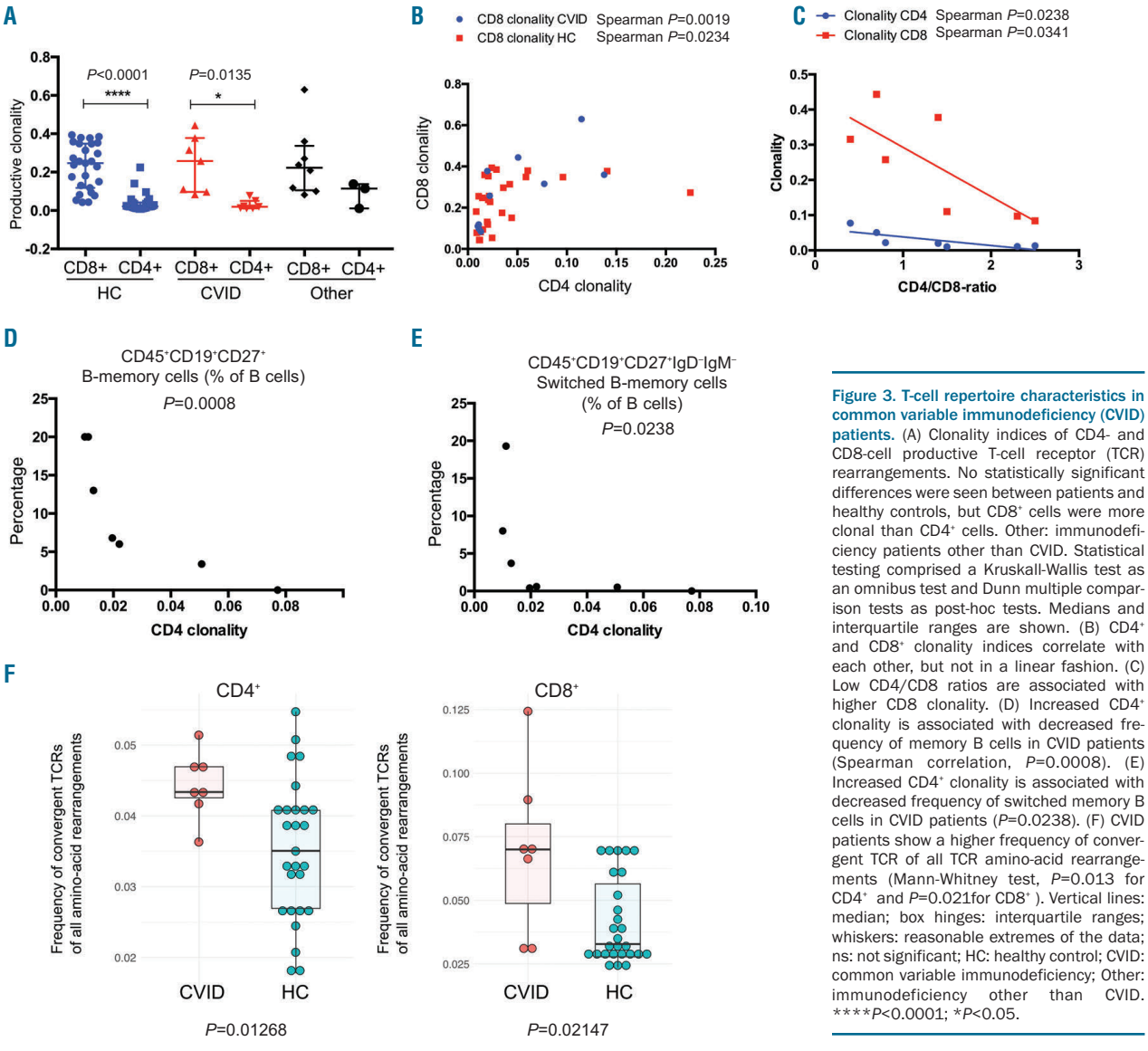

Figure 3. T-cell repertoire characteristics in common variable immunodeficiency (CVID) patients. (A) Clonality indices of CD4- and CD8-cell productive T-cell receptor (TCR) rearrangements. No statistically significant differences were seen between patients and healthy controls, but CD8+ cells were more clonal than CD4+ cells. Other: immunodeficiency patients other than CVID. Statistical testing comprised a Kruskall-Wallis test as an omnibus test and Dunn multiple comparison tests as post-hoc tests. Medians and interquartile ranges are shown. (B) CD4+ and CD8+ clonality indices correlate with each other, but not in a linear fashion. (C) Low CD4/CD8 ratios are associated with higher CD8 clonality. (D) Increased CD4+ clonality is associated with decreased frequency of memory B cells in CVID patients (Spearman correlation, *P*=0.0008). (E) Increased CD4+ clonality is associated with decreased frequency of switched memory B cells in CVID patients (*P*=0.0238). (F) CVID patients show a higher frequency of convergent TCR of all TCR amino-acid rearrangements (Mann-Whitney test, *P*=0.013 for CD4+ and *P*=0.021for CD8+). Vertical lines: median; box hinges: interquartile ranges; whiskers: reasonable extremes of the data; ns: not significant; HC: healthy control; CVID: common variable immunodeficiency; Other: immunodeficiency other than CVID. ****$P<0.0001$; *$P<0.05$.

T-cell receptor β-chain gene segment usage and junctional diversity

T-cell receptor β-chain V- and J-gene usage often serves as a marker for structural similarity of TCR. In total productive CD4+ and CD8+ TCR repertoires, there was no difference in *TCRB* V-gene family and J-gene usage between CVID patients and controls (*Online Supplementary Figure S8*). When comparing junctional diversity, no differences were found between patients and controls in the sums of base deletions and non-templated base insertions in the V-D and D-J junctions (*Online Supplementary Figure S9*). All non-templated base edition comparisons between CVID patients and controls are shown in *Online Supplementary Figure S10*.

Common variable immunodeficiency patients harbor higher frequencies of convergent T-cell receptor clones than healthy controls

Distinct T-cell clones have unique nucleotide-level CDR3 sequences, but different nucleotide sequences can translate into the same functional CDR3 amino-acid sequences (TCR convergence). CVID patients showed more convergence in CD4+ and CD8+ cells than healthy controls (CD4+ *P*=0.0123 and CD8+ *P*=0.0215) (Figure 3F). This was not due to higher frequencies of previously reported pathogen-specific TCR, such as Epstein-Barr virus (EBV) or cytomegalovirus (CMV)-specific TCR, in the CVID convergent TCR sequences (*Online Supplementary Figure S11*). However, convergent clones did not show higher abundance in CVID (abundance being defined as the sum of all convergent clone rearrangement frequencies) (*Online Supplementary Figure S12*).

We also investigated the physico-chemical properties of highly convergent and highly expanded TCR sequences with a generic string kernel algorithm (GSKernel) and unsupervised clustering analyses: CD8+ cells had 275 expanded and 3,738 highly convergent clones, and CD4+ cells had 342 expanded clones and 848 highly convergent clones. In CD8+ cells, these selected TCR formed 24 clusters based on the TCR physico-chemical similarity (Figure 4A). Six clusters were enriched with CVID TCR (Figure 4B). A major determinant driving the clustering was J-gene usage, but not V-gene usage (*Online Supplementary Figure S13*). In CD4+ cells, selected TCR formed 20 clusters with

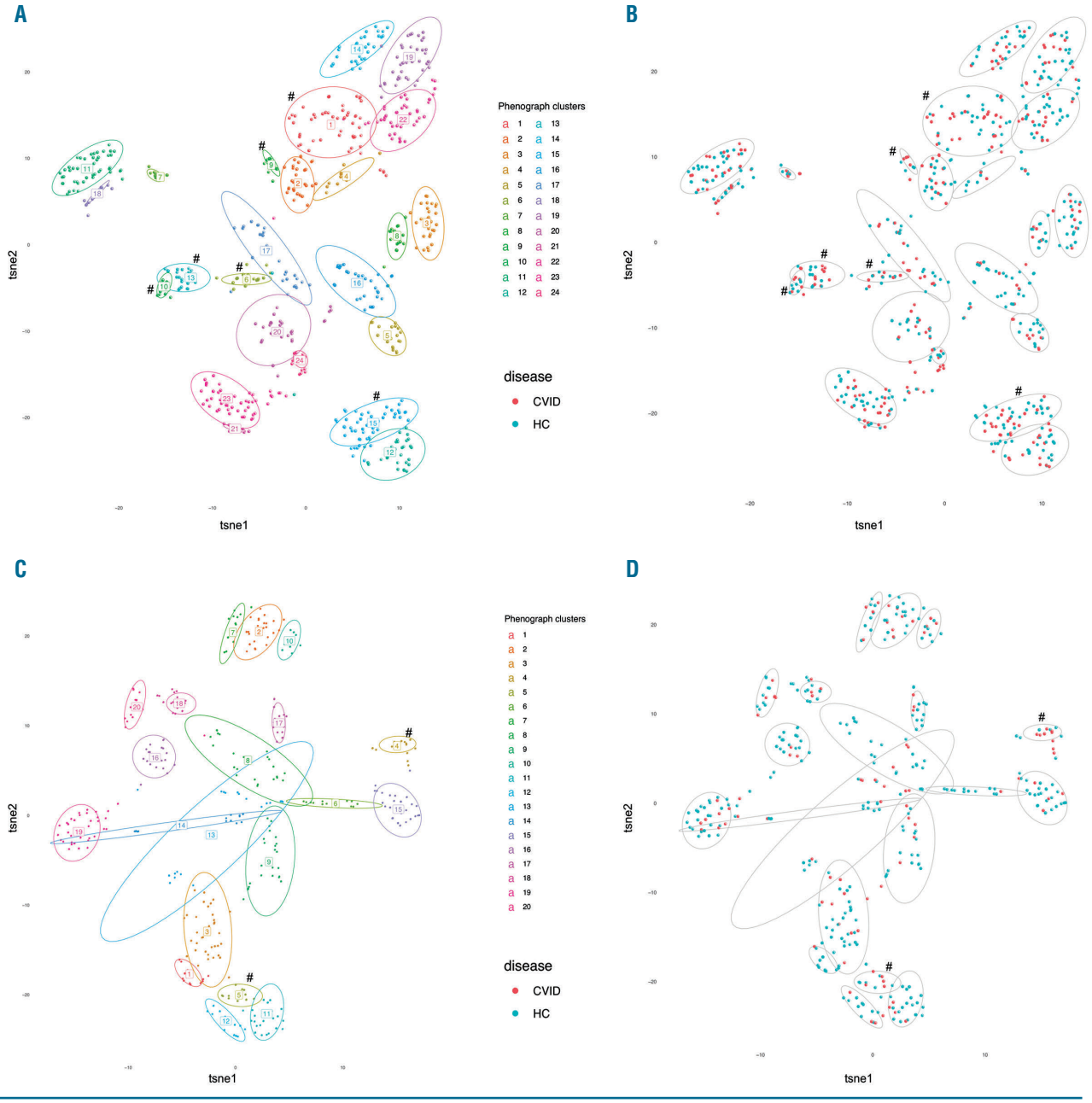

Figure 4. Convergent T-cell receptors (TCR) form structurally similar clusters, and some clusters enrich with common variable immunodeficiency (CVID) patients. Convergent and highly expanded TCR were analyzed for physico-chemical structure similarity and visualized with t-distributed stochastic neighbor embedding (t-SNE). Only highly convergent TCR and large clones (defined in "Methods") were included in the analyses. Each dot represents a single TCR, and each TCR can be mapped to one patient. (A) Analyzed TCR formed 24 clusters in CD8$^+$ cells, of which six were enriched with CVID TCR, as analyzed with Fisher's test (#). (B) The same clustering analysis, with CD8$^+$ CVID or healthy control (HC) TCR with colors. (C) Analyzed TCR formed 20 clusters in CD4$^+$ cells, of which two were enriched with CVID TCR (#). (D) The same clustering analysis, with CD4$^+$ CVID or HC TCR with colors.

similar results (Figure 4C and *Online Supplementary Figure S14*). Interestingly, convergent CD8$^+$ TCR in CVID had slightly different properties when compared to control TCR: CVID TCR showed a difference in polarity and hydrophobicity when compared to controls (*Online Supplementary Table S11*), whereas no differences were seen in CD4$^+$ TCR.

CD8$^+$ cells show higher frequencies of antigen-specific T cells than CD4$^+$ cells

Lastly, we investigated whether CVID patients had higher frequencies or abundances of previously reported TCR, with a focus on pathogen-specific TCR (*see "Methods" section*). There was no difference in frequencies of antigen-specific TCR between CVID patients and controls (*Online Supplementary Figures S15* and *S16*; healthy CD4$^+$ *vs.* CVID CD4$^+$, and healthy CD8$^+$ *vs.* CVID CD8$^+$). However, in healthy controls, the frequencies of pathogen-specific, cancer-specific, and autoimmune-associated TCR were higher in CD8$^+$ cells than CD4$^+$ cells, but these differences were not statistically significant in CVID (*Online Supplementary Figures S15 and S16*). The abun-

dance of pathogen-specific, but not cancer or autoimmune TCR, was also increased in CD8+ when compared to CD4+ cells in controls.

T-cell clone size and somatic mutation variant allele frequency

Individuals with mutations had larger CD8+ T-cell clones, but similar association was not observed among CD4+ T cells (*Online Supplementary Figure S6*). We plotted the frequency of the largest T-cell clone and mutation VAF (*Online Supplementary Figure S17*) to investigate the relationship between the clone size and VAF. If a somatic mutation occurs only in one individual T-cell clone, the mutation VAF should be half the clone size or less (assuming that most somatic mutations are heterozygous). In CD8+ T cells this was true, and mutation VAF correlated well with the clone size (*Online Supplementary Figure S17*). However, in some cases, the mutation VAF in CD4+ cells were higher than expected (*Online Supplementary Table S12* and *Online Supplementary Figure S17*), suggesting that mutations occur also in other cells rather than just in one CD4+ T-cell clone. Using flow cytometry-mediated sorting and sequencing, we confirmed that the mutations were still confined to the CD4+ T-cell fraction as they did not occur in other cell types, such as CD14+ cells or CD3+CD8+ cells (patients with *STAT5B* mutations) (*Online Supplementary Table S13*). However, clonal hematopoiesis-associated variants occurred in both CD4+ and CD8+ cells (Table 3 and *Online Supplementary Table S8*).

Discussion

In our small cohort (n=17) of late-onset CVID and other immunodeficiency patients, we show that somatic mutations occur in CD4+ and CD8+ cells in 75% of the CVID patients and in 65% of all patients. Lymphoproliferation (65%) and autoimmune disease manifestations were highly prevalent (94%) in our study cohort. Healthy controls harbored mutations in 45% of cases. Considering also the data on somatic mutations occurring in either CD4+ or CD8+ cells, we identified clonal hematopoiesis variants in both CD4+ and CD8+ cells in 24% of immunodeficiency patients. CVID patients also harbored higher frequencies of convergent TCR than healthy controls, and clustering analyses on selected TCR identified structurally similar TCR clusters that were enriched with CVID TCR.

Somatic mutations occur throughout life at an estimated rate of $3.5*10^{-9}$–$1.6*10^{-7}$ per base-pair per mitosis in many somatic tissues.[26] Since somatic mutations accumulate with age, the discovery of somatic mutations in CD8+ and CD4+ cells is not unexpected.[27-29] The majority of the identified mutations occurred in CD8+ cells, which could be due to the higher clonality of CD8+ T cells. In agreement with this, our results showed that patients with mutations had larger CD8+ T-cell clones. However, it should be noted that although the overall CD4+ clonality was low, many immunodeficiency patients harbored mutations also in CD4+ cells. Interestingly, the most dominant mutational signatures in T cells were signatures 1 and 15. Signature 1 correlates with age and derives from spontaneous deamination of 5-methylcytosine;[30] this signature was especially dominant in CD8+ cells and in immunodeficiency patients. Signature 15, which associates with defective mismatch repair,[31] was more prominent in healthy control CD4+ cells.

Many of the genes with somatic mutations in immunodeficiency patients link to malignant disorders, autoimmunity, or lymphoproliferative disorders. Somatic mutations in known hematologic tumor suppressors and/or oncogenes[18-20] occurred in 35% of immunodeficiency patients. CVID patients harbored these mutations in 38% (3 of 8) of cases and healthy controls in 9.5% (2 of 21). Two patients harbored somatic *STAT5B* mutations (patient 2, N418K; patient 13, T628S) in CD4+ cells. The T628S mutation in patient 13 results in increased STAT5 phosphorylation, transcriptional activity, and cell proliferation.[32] Activating *STAT5B* mutations in lymphocytes occur especially in CD4+ T-cell large-granular lymphocyte leukemia and T-prolymphocytic leukemia.[32,33] Patient 16 harbored a *KRAS* T58I mutation in CD8+ cells with VAF 7.9%. *KRAS* is mutated in nearly 30% of cancers,[34] and somatic activating *KRAS* mutations in both lymphoid and myeloid cells are observed in Ras-associated autoimmune leukoproliferative disorder.[35] Of note, the patient had a major T-cell clone in CD8+ cells comprising 51% of all CD8+ cells. Two patients had *C5AR1* (complement C5a receptor 1) mutations. Recent data indicate that C5AR1 activation promotes the proinflammatory Th1 responses in CD4+ cells *via* reactive oxygen species and inflammasome activation.[36]

Due to the lack of statistical power and differences in sequencing coverage, we cannot draw conclusions about the differences in mutation prevalence and mutation counts between immunodeficiency patients and controls. Although immunodeficiency patients had potentially pathogenic variants, also healthy controls harbored mutations in tumor suppressors and oncogenes, such as in *NRAS* (G13D) and *DNMT3A* (L773P), which is not unexpected due to the high prevalence of clonal hematopoiesis in 60-69 year olds.[37-39] Nevertheless, immunodeficiency patients had mutations that do not generally occur in clonal hematopoiesis in the general population, such as mutations in *STAT5B*. These mutations, or a subset of them, could contribute to lymphoproliferation in these patients, despite showing low VAF. Although the function of the discovered mutations was not analyzed in our study, previous studies by us and others suggest that somatic mutations may alter the phenotype and function of mutated T cells.[16,40] Interestingly, Fraietta *et al*. has shown that a hypomorphic *TET2* mutation can lead to the expansion and persistence of chimeric antigen receptor (CAR)-T cells.[40] Tet2 inactivation in mice results in altered T- and B-cell differentiation, and *TET2* mutations occur in T-cell lymphomas.[41] *DNMT3A* mutations may also facilitate lymphoproliferation, as shown by Dnmt3a deficiency leading to both myeloid and lymphoid malignancies in mice.[42] *DNMT3A* is also commonly mutated in human acute T-cell lymphoblastic leukemias.[42] The clinical relevance of *DNMT3A* and *TET2* mutations and the possible survival advantage they may confer in immunodeficiency requires more studies, as the prevalence of these mutations is relatively high also in healthy T cells.

The majority of the detected mutations occurred exclusively in CD4+ or CD8+ cells. The origin of these mutations is elusive. Do these somatic mutations occur in one T-cell clone or in multiple clones? Do they occur in every cell of the same clone or only in a subset? Any conclusive answer could only be obtained through in-depth single-cell sequencing studies. In our previous reports, we

showed that mutations are confined to specific T-cell populations in CD8+ cells.[16] Similarly, in immunodeficiency patient mutation, VAF and clone sizes matched well in CD8+ T cells, suggesting that mutations occur late during T-cell maturation. However, the analysis of CD4+ T cells suggested that mutations may also occur in precursor cells, as mutation VAF were higher than individual CD4+ T-cell clone sizes.

In addition to mutations found exclusively in one cell type, we also detected mutations which were found in both CD4+ and CD8+ compartments, suggesting that they could have originated from a hematopoietic stem cell, or from multipotent or lineage-specific progenitors. Our results are in line with previous publications showing that clonal hematopoiesis-associated mutations (such as *DNMT3A* and *TET2*) can occur in T cells.[37,43,44] These mutations likely originate from progenitor cells, because in our cases they were found both in CD4+ and CD8+ T cells, and in CD19+ B cells, but not in myeloid CD14+ cells. However, myeloid cells can harbor the same mutations as B cells (as in monoclonal B-cell lymphocytosis and indolent chronic lymphocytic leukemia) or T cells (in individuals without hematologic disease).[37,44-47]

In the present study, we detected somatic mutations with relatively low VAF by using high sequencing coverage in the assays. To avoid exhaustion of the unique DNA molecules, a sufficient amount of DNA (100-500 ng) was used based on the calculations that one nanogram of DNA comprises approximately 300 copies of unique DNA molecules. Assay sensitivity is also determined by other factors, such as sequencer performance and bioinformatics approaches. In general, Illumina sequencers show a low error rate,[48] and we applied standard variant detection practices with multiple quality filters (such as VAF threshold and the requirement for at least seven supporting reads for the variant), similar to reports by others using capture sequencing technologies.[7,18,19,49,50]

Although a previous study had suggested increased T-cell clonality in CVID patients,[51] we did not observe any difference in clonality between CVID and controls. The heterogenous phenotypes of CVID patients, ages of controls, and sample material may explain the difference in results. Interestingly, however, we noted increased TCR convergence in CVID. Previous studies have shown that TCR convergence associates with clone size[52] and public TCR in mice.[53] We hypothesize that convergence is due to enrichment mediated by antigen stimulation. The latter is further supported by the finding that CVID convergent TCR showed some structural similarity (Figure 4), which could be related to common antigen stimulation.

In conclusion, we discovered somatic mutations and increased TCR convergence in mature T cells in patients with immunodeficiency. TCR convergence may be due to a chronic antigen response, and somatic variants may contribute to immune dysregulation. Somatic mutations may originate either in mature T cells or hematopoietic stem cells. Although healthy controls also harbor some variants with potentially pathogenic consequences, our results demonstrate that mutations in genes that have been associated with autoimmunity and lymphoproliferation (such as *STAT5B*) occur in immunodeficiency patients. Further studies are needed to discover the full spectrum of somatic mutations and their role as possible regulators of immune responses both in normal and diseased T cells.

Acknowledgments

We would like to acknowledge the personnel at Hematology Research Unit Helsinki for their assistance, and the FIMM Technology Centre sequencing unit for their assistance on sequencing and data analysis. The HUSLAB flow cytometry unit is acknowledged for their expertise on the B-cell phenotyping assays. The Biomedicum Flow cytometry core unit (HiLIFE) is acknowledged for their assistance on cell sorting.

Funding

This work was supported by the European Research Council (M-IMM project), the ERAPerMed consortium 'JAKSTAT-TARGET, Academy of Finland, Finnish special governmental subsidy for health sciences, research and training, the Sigrid Juselius Foundation, the Instrumentarium Science Foundation, the Finnish Cancer Societies, the Finnish Cancer Institute, the Biomedicum Helsinki Foundation, the Finnish Medical Foundation, the Orion Research Foundation, the Juhani Aho Foundation, the K. Albin Johansson Foundation, the Paulo Foundation, and the Foundation for Pediatric Research.

References

1. Bonilla FA, Barlan I, Chapel H, et al. International consensus document (ICON): common variable immunodeficiency disorders. J Allergy Clin Immunol Pract. 2016; 4(1):38-59.
2. Bogaert DJA, Dullaers M, Lambrecht BN, Vermaelen KY, De Baere E, Haerynck F. Genes associated with common variable immunodeficiency: one diagnosis to rule them all? J Med Genet. 2016;53(9):575-590.
3. Selenius JS, Martelius T, Pikkarainen S, et al. Unexpectedly high prevalence of common variable immunodeficiency in Finland. Front Immunol. 2017;81190.
4. Wong GK, Huissoon AP. T-cell abnormalities in common variable immunodeficiency: the hidden defect. J Clin Pathol. 2016;69(8):672-676.
5. Gathmann B, Mahlaoui N, Gerard L, et al. Clinical picture and treatment of 2212 patients with common variable immunodeficienc. J Allergy Clin Immunol. 2014; 134(1):116-126.
6. Mellemkjær L, Hammarström L, Andersen V, Yuen J, Heilmann C, Barington T. Cancer risk among patients with IgA deficiency or common variable immunodeficiency and their relatives : a combined Danish and Swedish study. Clin Exp Immunol. 2002; 130(3):495-500.
7. Nikolaev SI, Vetiska S, Bonilla X, et al. Somatic activating KRAS mutations in arteriovenous malformations of the brain. N Engl J Med. 2018;378(3):250-261.
8. Anglesio MS, Papadopoulos N, Ayhan A, et al. Cancer-associated mutations in endometriosis without cancer. N Engl J Med. 2017;376(19):1835-1848.
9. Koskela HLM, Eldfors S, Ellonen P, et al. Somatic STAT3 mutations in large granular lymphocytic leukemia. N Engl J Med. 2012; 366(20):1905-1913.
10. Jerez A, Clemente MJ, Makishima H, et al. STAT3 mutations unify the pathogenesis of chronic lymphoproliferative disorders of NK cells and T cell large granular lymphocyte leukemia. Blood. 2012;120(15):3048-3058.
11. Andersson E, Kuusanmäki H, Bortoluzzi S, et al. Activating somatic mutations outside the SH2-domain of STAT3 in LGL leukemia. Leukemia. 2016;30(5):1204-1208.
12. Jerez A, Clemente MJ, Makishima H, et al. STAT3-mutations indicate the presence of subclinical T cell clones in a subset of aplastic anemia and myelodysplastic syndrome patients. Blood. 2013;122(14):2453-2459.
13. Rajala HLM, Olson T, Clemente MJ, et al. The analysis of clonal diversity and therapy responses using STAT3 mutations as a molecular marker in large granular lymphocytic leukemia. Haematologica. 2015; 100(1):91-99.
14. Ettersperger J, Montcuquet N, Malamut G, et al. Interleukin-15-dependent T-cell-like innate intraepithelial lymphocytes develop in the intestine and transform into lymphomas in celiac disease. Immunity. 2016; 45(3):610-625.

15. Savola P, Brück O, Olson T, et al. Somatic STAT3 mutations in Felty syndrome: an implication for a common pathogenesis with large granular lymphocyte leukemia. Haematologica. 2018;103(2):304-312.
16. Savola P, Kelkka T, Rajala HLM, et al. Somatic mutations in clonally expanded cytotoxic T lymphocytes in patients with newly diagnosed rheumatoid arthritis. Nat Commun. 2018;8:15869.
17. Valori M, Jansson L, Kiviharju A, et al. A novel class of somatic mutations in blood detected preferentially in CD8+ cells. Clin Immunol. 2017;175:75-81.
18. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med. 2014; 371(26):2488-2498.
19. Jaiswal S, Natarajan P, Silver AJ, et al. Clonal hematopoiesis and risk of atherosclerotic cardiovascular disease. N Engl J Med. 2017; 377(2):111-121.
20. Genovese G, Kähler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med. 2014;371(26):2477-2487.
21. Fuster JJ, MacLauchlan S, Zuriaga MA, et al. Clonal hematopoiesis associated with Tet2 deficiency accelerates atherosclerosis development in mice. Science. 2017; 355(6327):842-847.
22. Holzelova E, Vonarbourg C, Stolzenberg M-C, et al. Autoimmune lymphoproliferative syndrome with somatic Fas mutations. N Engl J Med. 2004;351(14):1409-1418.
23. Trotta L, Martelius T, Siitonen T, et al. ADA2 deficiency: clonal lymphoproliferation in a subset of patients. J Allergy Clin Immunol. 2018;141(4):1534-1537.
24. Haapaniemi EM, Kaustio M, Rajala HLM, et al. Autoimmunity, hypogammaglobulinemia, lymphoproliferation, and mycobacterial disease in patients with activating mutations in STAT3. Blood. 2015;125(4):639-649.
25. Seppänen M, Koillinen H, Mustjoki S, Tomi M, Sullivan KE. Terminal deletion of 11q with significant late-onset combined immune deficiency. J Clin Immunol. 2014; 34(1):114-118.
26. Werner B, Sottoriva A. Variation of mutational burden in healthy human tissues suggests non-random strand segregation and allows measuring somatic mutation rates. PLoS Comput Biol. 2018;14(6): e1006233.
27. Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. Cell. 2012;150(2):264-278.
28. Blokzijl F, De Ligt J, Jager M, et al. Tissue-specific mutation accumulation in human adult stem cells during life. Nature. 2016; 538(7624):260-264.
29. Franco I, Johansson A, Olsson K, et al. Somatic mutagenesis in satellite cells associates with human skeletal muscle aging. Nat Commun. 2018;9(1):800.
30. Alexandrov LB, Jones PH, Wedge DC, et al. Clock-like mutational processes in human somatic cells. Nat Genet. 2015;47(12):1402-1407.
31. Petljak M, Alexandrov LB. Understanding mutagenesis through delineation of mutational signatures in human cancer. Carcinogenesis. 2016;37(6):531-540.
32. Kiel MJ, Velusamy T, Rolland D, et al. Integrated genomic sequencing reveals mutational landscape of T-cell prolymphocytic leukemia. Blood. 2014;124(9):1460-1473.
33. Andersson EI, Tanahashi T, Sekiguchi N, et al. High incidence of activating STAT5B mutations in CD4-positive T-cell large granular lymphocyte leukemia. Blood. 2016; 128(20):2465-2468.
34. Hobbs GA, Der CJ, Rossman KL. RAS isoforms and mutations in cancer at a glance. J Cell Sci. 2016;129(7):1287-1292.
35. Niemela JE, Lu L, Fleisher TA, et al. Somatic KRAS mutations associated with a human nonmalignant syndrome of autoimmunity and abnormal leukocyte homeostasis. Blood. 2012;117(10):2883-2886.
36. Arbore G, West EE, Spolski R, et al. T helper 1 immunity requires complement-driven NLRP3 inflammasome activity in CD4+ T cells. Science. 2016;352(6292):aad1210.
37. Arends CM, Galan-Sousa J, Hoyer K, et al. Hematopoietic lineage distribution and evolutionary dynamics of clonal hematopoiesis. Leukemia. 2018;32(9):1908-1919.
38. Acuna-Hidalgo R, Sengul H, Steehouwer M, et al. Ultra-sensitive sequencing identifies high prevalence of clonal hematopoiesis-associated mutations throughout adult life. Am J Hum Genet. 2017;101(1):50-64.
39. Buscarlet M, Provost S, Zada YF, et al. DNMT3A and TET2 dominate clonal hematopoiesis and demonstrate benign phenotypes and different genetic predispositions. Blood. 2017;130(6):753-762.
40. Fraietta JA, Nobles CL, Sammons MA, et al. Disruption of TET2 promotes the therapeutic efficacy of CD19-targeted T cells. Nature. 2018;558(7709):307-312.
41. Quivoron C, Couronné L, Della Valle V, et al. TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. Cancer Cell. 2011;20(1):25-38.
42. Bowman RL, Busque L, Levine RL. Clonal hematopoiesis and evolution to hematopoietic malignancies. Cell Stem Cell. 2018; 22(2):157-170.
43. Buscarlet M, Provost S, Feroz Zada Y, et al. Lineage restriction analyses in CHIP indicate myeloid bias for TET2 and multipotent stem cell origin for DNMT3A. Blood. 2018; 132(3):277-280.
44. Young AL, Challen GA, Birmann BM, Druley TE. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. Nat Commun. 2016; 7:12484.
45. Damm F, Mylonas E, Cosson A, et al. Acquired initiating mutations in early hematopoietic cells of CLL patients. Cancer Discov. 2014;4(9):1088-1101.
46. Condoluci A, Rossi D. Age-related clonal hematopoiesis and monoclonal B-cell lymphocytosis / chronic lymphocytic leukemia: a new association? Haematologica. 2018; 103(5):751-752.
47. Agathangelidis A, Ljungström V, Scarfò L, et al. Highly similar genomic landscapes in monoclonal B-cell lymphocytosis and ultra-stable chronic lymphocytic leukemia with low frequency of driver mutations. Haematologica. 2018;103(5):865-873.
48. Caspar SM, Dubacher N, Kopps AM, Meienberg J, Henggeler C, Matyas G. Clinical sequencing: from raw data to diagnosis with lifetime value. Clin Genet. 2018; 93(3):508-519.
49. Desai P, Mencia-Trinchant N, Savenkov O, et al. Somatic mutations precede acute myeloid leukemia years before diagnosis. Nat Med. 2018;24(7):1015-1023.
50. Abelson S, Collord G, Stanley W, et al. Prediction of acute myeloid leukaemia risk in healthy individuals. Nature. 2018; 559(7714):400-404.
51. Ramesh M, Hamm D, Simchoni N, Cunningham-Rundles C. Clonal and constricted T cell repertoire in common variable immune deficiency. Clin Immunol. 2017; 178:1-9.
52. Venturi V, Quigley MF, Greenaway HY, et al. A mechanism for TCR sharing between T cell subsets and individuals revealed by pyrosequencing. J Immunol. 2011; 186(7):4285-4294.
53. Madi A, Shifrut E, Reich-Zeliger S, et al. T-cell receptor repertoires share a restricted set of public and abundant CDR3 sequences that are associated with self-related immunity. Genome Res. 2014;24(10):1603-1612.