Genetic characterization of acquired aplastic anemia by targeted sequencing

Michael Heuser,[1] Carola Schlarmann,[1] Vera Dobbernack,[1] Victoria Panagiota,[1] Lutz Wiehlmann,[2] Carolin Walter,[3] Fabian Beier,[4] Patrick Ziegler,[4] Haiyang Yun,[1] Sofia Kade,[1] Aylin Kirchner,[1] Liu Huang,[1,5] Christian Koenecke,[1] Matthias Eder,[1] Tim H. Brümmendorf,[4] Martin Dugas,[3] Arnold Ganser,[1] and Felicitas Thol[1]

[1]Department of Hematology, Hemostasis, Oncology, and Stem Cell Transplantation, Hannover Medical School, Germany; [2]Clinical Research Group, Department of Pediatric Pneumology, Allergology and Neonatology, Hannover Medical School, Germany; [3]Institute of Medical Informatics, Münster University, Germany; [4]Department of Oncology, Hematology and Stem Cell Transplantation, RWTH Medical School, Aachen, Germany; and [5]Department of Oncology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, China

Correspondence: heuser.michael@mh-hannover.de
doi:10.3324/haematol.2013.101642

Genetic characterization of acquired aplastic anemia by targeted sequencing

Michael Heuser[1], Carola Schlarmann[1], Vera Dobbernack[1], Victoria Panagiota[1], Lutz Wiehlmann[2], Carolin Walter[3], Fabian Beier[4], Patrick Ziegler[4], Haiyang Yun[1], Sofia Kade[1], Aylin Kirchner[1], Liu Huang[1,5], Christian Koenecke[1], Matthias Eder[1], Tim H. Brümmendorf[4], Martin Dugas[3], Arnold Ganser[1], Felicitas Thol[1]

Supplementary Information

Supplementary methods

Thirty eight patients with aplastic anemia who were diagnosed and treated in our institution were included in this study. Fanconi anemia was excluded in only one patient experimentally. However, none of the patients presented with clinical, non-hematologic abnormalities indicating Fanconi anemia (eg. short stature, abnormalities of the skin, bone, eyes, kidneys or ears). Twenty two patients underwent allogeneic hematopoietic stem cell transplantation for whom DNA from peripheral blood mononuclear cells was available within 4 weeks before transplantation. We expanded our cohort with patients not undergoing transplantation (n=16) for whom bone marrow (n=5) or peripheral blood (n=11) samples were available at a time when the patients were cytopenic in at least one blood lineage (time from diagnosis to sample harvest 2 years, range 0 - 21 years). We used DNA from eyebrow hair follicles or sorted CD3+ cells from peripheral blood as germline control. Peripheral blood leukocytes from 20 healthy volunteers were used for telomere length assessment.

For flow cytometric diagnosis of a PNH cell population, at least two cell lineages with significant GPI-deficient populations were required. The following antibodies were used to identify GPI-deficient cells: CD55 and CD59 on erythrocytes, FLAER on CD24 positive granulocytes and FLAER on CD14 positive monocytes. PNH was defined as presence of marker-negative populations exceeding the cut-off value of 1% in at least two lineages. Cytogenetic analysis was performed by G- and R-banding at diagnosis, and chromosomal abnormalities were described according to the International System for Human Cytogenetic Nomenclature. The presence of a PNH clone was evaluated in the neutrophil, monocytic and erythroid lineages by flow cytometric analysis at diagnosis and during the course of the disease.

Mononuclear cells from BM and/or blood specimens were enriched by Ficoll density gradient centrifugation and were stored at -70°C in liquid nitrogen until use. Genomic DNA was extracted from samples using the All Prep DNA/RNA Kit (Qiagen, Hilden, Germany) according to the manufacturer’s recommendations, and was whole genome amplified (GenomePlex whole genome amplification kit, Sigma-Aldrich, Seelze, Germany). The coding exons of 33 MDS-related genes (Table S2) were PCR-amplified with not more than 5 amplicons per well using standard conditions (334 amplicons in total). Primer sequences are available upon request. For library construction, the PCR amplified amplicons from one patient were pooled and randomly ligated (Quick Ligation Kit, New England Biolabs, Ipswich, MA). The long concatenated DNA was then sheared into 100-250 bp fragments using the Covaris® System (Covaris, Woburn, MA) to obtain randomly fragmented sequences, and was size selected for 200 bp fragments using Agencourt AMPure® XP reagent (Agencourt Bioscience Corporation, Beverly, MA). Patient-specific barcodes and sequencing primers P1 and P2 were ligated to these fragments. Fragments were loaded and amplified during emulsion PCR. The beads were then added to the Flow Chip for sequencing using the SOLiD™ system (Life Technologie, Darmstadt, Germany), and sequenced according to the manufacturers protocol (Life Technologies) using the ECC (exact call chemistry) module for higher accuracy. Individual reads were 75 basepairs long.

Sequence reads were assigned to their patient-specific barcode, and sequences were analyzed twice separately using the DNAnexus software (DNAnexus.com) and an in-house pipeline of bioinformatics software. The color-space reads were aligned with NovoAlignCS and genotyped with GATK's Unified Genotyper.[1] SNV discovery was performed across all samples using standard parameters and a maximum coverage of 10,000. The resulting list of candidate single nucleotide variants (SNV) was filtered with R. First, mutations outside the coding region were excluded. Second, known single nucleotide polymorphisms (dbSNP version 137) were removed. Third, variants with a quality score of less than 8000 or a frequency of less than 15% were excluded. The 15% cutoff was chosen as we wanted to consider only variants that can be validated by Sanger sequencing. To determine the coverage of genomic intervals, the data was processed with BEDTools[2] and analysed in R. All candidate SNVs were validated by Sanger sequencing and only those confirmed either in genomic DNA or in an independently whole genome-amplified DNA sample are

reported. Exon 13 of *ASXL1* and exons 3 to 11 of *TET2* were also sequenced by Sanger sequencing in addition to 9 other genes (Table S2), which were only sequenced by Sanger sequencing or fragment analysis using a 3130 Genetic Analyzer (Applied Biosystems, Darmstadt, Germany). Confirmed mutations were also sequenced in germline DNA.

Telomere length was evaluated by monochrome multiplex quantitative polymerase chain reaction (PCR) using primers for telomeres and the reference gene beta-globin (Table S4) as described before[3] in the same samples that were used for molecular analysis (if sufficient material was available). DNA was assayed in duplicate in two independent runs on a MyIQ2 two-color real-time PCR detection system (Bio-Rad Laboratories, München, Germany). Amplification efficiency was 90 to 100%. T/S ratio was calculated by dividing the copy number of the telomere template T by the copy number of the beta-globin template S. The average T/S ratio was calculated from two independent runs.

Pair-wise comparisons of variables were performed for exploratory purposes using Student's t-test for continuous variables and Chi-squared test for categorical variables. The Kaplan-Meier method and log-rank test were used to estimate the distribution of OS, and to compare differences between survival curves, respectively. Statistical analyses were performed with the statistical software package SPSS 20.0 (IBM Corporation, Armonk, NY) and GraphPad Prism 5 (Statcon, Witzenhausen, Germany).

Supplementary References

1. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43(5):491-8.

2. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841-2.

3. Ziegler P, Schrezenmeier H, Akkad J, Brassat U, Vankann L, Panse J, et al. Telomere elongation and clinical response to androgen treatment in a patient with aplastic anemia and a heterozygous hTERT gene mutation. Ann Hematol. 2012;91(7):1115-20.

4. Walter MJ, Shen D, Shao J, Ding L, White BS, Kandoth C, et al. Clonal diversity of recurrently mutated genes in myelodysplastic syndromes. Leukemia. 2013;27(6):1275-82.

Supplementary Tables

Supplementary Table S1. Clinical characteristics of aplastic anemia patients.

Characteristic	All patients
	(n =38)
Age at diagnosis, years	
median	30
range	9-79
Sex	
male - no. (%)	18 (47)
female - no. (%)	20 (53)
Severity of aplastic anemia	
Moderate – no. (%)	11 (29)
Severe – no. (%)	12 (32)
Very severe – no. (%)	15 (39)
WBC count at time of sample harvest	
median - ($x10^9$/l)	2.3
range - ($x10^9$/l)	0.3 - 7.6
Neutrophils at time of sample harvest	
median - ($x10^9$/L)	1.2
range - ($x10^9$/L)	0 - 6.5
Hemoglobin at time of sample harvest	
median – g/dL	9.5
range – g/dL	6 -14.6
Reticulocytes at time of sample harvest	
median - ($x10^9$/L)	35
range - ($x10^9$/L)	3 - 235
Platelet count at time of sample Harvest	
median - ($x10^9$/L)	28.5
range - ($x10^9$/L)	1 - 267
Source of genomic DNA	
Peripheral blood - no. (%)	33 (87)
Bone marrow - no. (%)	5 (13)
Karyotype	
normal – no. (%)	22 (58)
Aberrant[§] – no. (%)	3 (8)
Missing – no. (%)	13 (34)
PNH clone present	

yes – no. (%)	17 (44.7)
clone size (%)	1.5 - 98
missing data – no. (%)	3 (8)
Median follow-up from diagnosis (years)	7.0
10-year overall survival (%)	90

Abbreviations: n.a., not applicable; alloHSCT, allogeneic hematopoietic stem cell transplantation; PNH, paroxysmal nocturnal hemoglobinuria.

[§] One patient developed trisomy 8 eleven years after first diagnosis; one patient had Robertsonian translocation rob(13;14)(q10;q10); and one patient had t(6,13)(p22;q13). Two of these patients received an allogeneic transplant (except the patient who developed trisomy 8).

Supplementary Table S2. Genes analyzed by next generation or Sanger sequencing.

Analyzed genes[§]	Analyzed regions of genes
Next generation sequencing	
ASXL1*	exon 13
BCL11B	exon 1-4
BCOR*	exon 2-15
CBL*	exon 8-9
cKIT	exon 8, 11, 17
CXXC4	exon 2, 3
CXXC5	exon 2, 3
DAXX	exon 2-8
DNM2	exon 1-21
DNMT3A*	exon 2-23
ETV6*	exon 2-8
EZH2*	exon 2-20
GATA2*	exon 2-6
IDH1*	exon 4
IDH2*	exon 4
JAK2*	exon 12, 14
KRAS	exon 1, 2
MYBL2	exon 2-14
PHF6*	exon 2-10
PIK3CA	exon 9, 20
PTPN11*	exon 3, 13
RAD21	exon 2-14
RUNX1*	exon 3-8
SF3B1*	exon 13-16
SMC1A	exon 1-25
SMC3*	exon 1-29

STAG1	exon 2-34
STAG2*	exon 3-35
TET2*	exon 3-11
TP53*	exon 5-8
U2AF1*	exon 1-8
WT1*	exon 7, 9
ZRSR2*	exon 1-11
Sanger sequencing	
ASXL1*	exon 13
MPL	exon 8, 10 (W515 hotspot)
RIT1	exon 5
SETPB1*	D868 and G870 mutation hotspot in exon 4
SLIT1	exon 13
SRSF2*	P95 mutation hotspot in exon 1
STAT3	exon 21
TET2*	exon 3-11
TERT	Exon 4-16
Fragment analysis[#]	
FLT3*	exon 14, 15
NPM1*	exon 11

[§] based on genome assembly GRCh37, Ensembl human release 72

* denotes genes that were also sequenced by Walter et al.[4]

[#] Fragment analysis of PCR amplicons obtained from genomic DNA using a 3130 Genetic Analyzer (Applied Biosystems, Darmstadt, Germany)

Supplementary Table S3. Comparison of patient characteristics between 38 aplastic anemia patients of the present study and 150 MDS patients from the study by Walter et al.[4]

Characteristic	Aplastic anemia patients	MDS patients (Walter et al.)[4]
	(n =38)	(n =150)
Age at diagnosis, years		
<60	32 (84)	57 (38)
≥60	6 (16)	93 (62)
Sex		
male - no. (%)	18 (47)	92 (61)
female - no. (%)	20 (53)	58 (39)
Severity of aplastic anemia		
Moderate – no. (%)	11 (29)	
Severe – no. (%)	12 (32)	
Very severe – no. (%)	15 (39)	

IPSS-R of MDS patients		
Very low - no. (%)		12 (8)
Low - no. (%)		34 (23)
Intermediate - no. (%)		37 (25)
High - no. (%)		29 (19)
Very high - no. (%)		33 (22)
missing - no. (%)		5 (3)

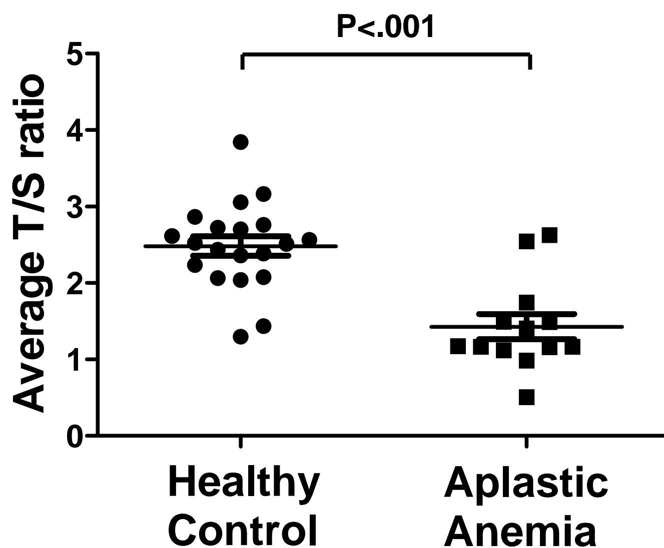
Abbreviation: IPSS-R, international prognostic scoring system, revised.

Supplementary Table S4. Primer sequences for telomere length assessment.

Name	5’ to 3’ sequence
telg	CACTAAGGTTTGGGTTTGGGTTTGGGTTTGGGTTAGTGT
telc	TGTTAGGTATCCCTATCC CTATCCCTATCCCTATCCCTAACA
hbgu	CGGCGGCGGGCGGCGCGGGCTGGG CGGCTTCATCCACGTTCACCTTG
hbgd	GCCCGGCCCGCCGCGCCCGTCCCGCCGGAGGA GAAGTCTGCCGTT

Supplementary Figures

Supplementary Figure S1: Telomere length in peripheral blood leukocytes of healthy volunteers (n=20) and aplastic anemia patients (n=13).



Supplementary Figure S2. Sequencing traces of mutated genes in aplastic anemia patients. Sequence numbering is according to the DNA coding sequence of ENST00000266058 (SLIT1), ENST00000282030 (SETBP1) and ENST00000375687 (ASXL1), Ensembl transcript ENST00000217026 (MYBL1), ENST00000540549 (TET2) of genome build GRCh37, release 73.