Aus der Klinik für Hämatologie, Onkologie und Tumorimmunologie
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

New Insights into the Dynamics of Clonal Hematopoiesis from Targeted Sequencing Studies

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Dipl.-Phys. Christopher Maximilian Arends

aus Bonn

Datum der Promotion: 18.12.2020

Table of Contents

Abstract 5

Abstract in German 7

1 Introduction 9

2 Methods 12

2.1 Patient Collectives 12

2.2 Sample Processing of Fresh Peripheral Blood and Bone Marrow 13

2.3 Targeted Sequencing 13

2.4 Sequencing Data Analysis 14

2.5 Validation and Longitudinal Quantification 15

2.6 Flow-Sorting of Peripheral Blood and Bone Marrow Cells 16

2.7 Neutrophil Activation Assays and mRNA Expression 16

2.8 Statistical Analysis 17

3 Results 18

3.1 Sequencing Results in the Non-Cancer and Cancer Cohort 18

3.2 Hematopoietic Lineage Distribution 18

3.3 Evolution of Mutated Clones under Cytotoxic Treatment 19

3.4 Clonal Hematopoiesis in Allogeneic Stem Cell Transplantation 20

3.5 Clonal Hematopoiesis in Patients with ANCA-Associated Vasculitis 21

4 Discussion 23

4.1 Lineage distribution analysis 23

4.2 Clonal Hematopoiesis in Patients Receiving Chemotherapy 24

4.3 Clonal Hematopoiesis in Allogeneic Stem Cell Transplantation 25

4.4 Clonal Hematopoiesis in Patients with ANCA-Associated Vasculitis 25

4.5 Summary and Outlook 27

Figures 28

References 32

Eidesstattliche Versicherung 37

Anteilserklärung an den erfolgten Publikationen..... 38

Publication 1: Hematopoietic Lineage Distribution and Evolutionary Dynamics of Clonal Hematopoiesis..... 40

Publication 2: Role of Donor Clonal Hematopoiesis in Allogeneic Hematopoietic Stem-Cell Transplantation 52

Publication 3: Clonal Hematopoiesis in Patients with Anti-Neutrophil Cytoplasmatic Antibody-associated Vasculitis 64

Curriculum Vitae 68

List of Publications..... 70

Acknowledgements 71

Abstract

The presence of a clonal population of blood cells harboring somatic mutations in individuals without further clinical manifestations is termed “clonal hematopoiesis of indeterminate potential” (CHIP). It occurs frequently in the elderly, predisposes for the development of hematologic malignancies and cardiovascular diseases and is associated with inferior overall survival.

Using a combination of flow-sorting and targeted deep sequencing, we studied CHIP in an extensive patient collective, including 365 patients with cardiovascular disease, 72 patients with solid tumors, 112 patients with anti-neutrophil cytoplasmatic antibody (ANCA)-associated vasculitis (AAV) and 500 elderly healthy stem cell donors. CHIP status was correlated with clinical outcomes such as overall survival, relapse risk, therapy-related complications and paraclinical parameters. We performed lineage distribution analyses and longitudinal quantification of clonal evolution in a subset of 63 patients, each.

CHIP prevalence varied between 16.0 % in the stem cell donor cohort and 30.6 % in the cancer cohort. Lineage distribution analyses revealed a myeloid-biased differentiation pattern with only few exceptions. Longitudinal analyses revealed diverse patterns of clonal evolution under different stress scenarios such as cytotoxic treatment, allogeneic hematopoietic stem cell transplantation (HSCT), and immunosuppressive treatment. A subgroup of CHIP-positive patients with solid cancer receiving chemotherapy more frequently required erythrocyte transfusions and dose reductions due to hematotoxicity. Donor CHIP did not affect recipients’ overall survival after HSCT but had impact on the cumulative incidences of chronic graft versus host disease (cGVHD) and relapse/progression. In patients with AAV we detected a higher CHIP prevalence as compared to an age- and sex-matched control cohort. We found evidence for potential disease-modifying effects as shown for disease manifestation patterns, *in vitro* measurement of neutrophil activation and ANCA-antigen transcription regulation.

In conclusion, we provide novel insights into CHIP and its implication in cancer therapy, HSCT and autoimmune diseases. Our results might spark future translational studies that will help to further decipher the clinical significance of CHIP in these contexts.

Abstract in German

Klonale Hämatopoese unbestimmten Potentials (engl. *clonal hematopoiesis of indeterminate potential*, CHIP) bezeichnet das Vorhandensein einer klonalen Population von Blutzellen mit somatischen Mutationen in Individuen ohne hämatologische Erkrankung. Sie tritt mit zunehmendem Alter gehäuft auf, prädisponiert zur Entwicklung hämatologischer Neoplasien und kardiovaskulärer Erkrankungen und ist mit einem reduzierten Gesamtüberleben assoziiert.

Mithilfe einer Kombination aus *FACS-sorting* und *targeted deep sequencing* wurde CHIP in einem umfangreichen Patientenkollektiv untersucht, darunter 365 Patienten mit kardiovaskulären Erkrankungen, 72 Patienten mit soliden Tumoren, 112 Patienten mit Anti-Neutrophile cytoplasmatische Antikörper (ANCA)-assoziierten Vaskulitiden (AAV) und 500 gesunde ältere Stammzellspender. Der CHIP Status wurde mit dem Gesamtüberleben, dem Rückfallrisiko, dem Auftreten von therapieassoziierten Komplikationen und paraklinischen Parametern korreliert. Linienrestriktionsanalysen und longitudinale Quantifizierung der klonalen Evolution wurden in je 63 Patienten durchgeführt.

Die CHIP Prävalenz lag zwischen 16.0 % in der Stammzellspenderkohorte und 30.6 % in der Tumorkohorte. Linienrestriktionsanalysen ergaben mit wenigen Ausnahmen eine eindeutige Verlagerung der Differenzierung zugunsten der myeloischen Reihe. In der longitudinalen Analyse zeigten sich diverse klonale Evolutionsmuster unter Chemotherapie, allogener Stammzelltransplantation und immunsuppressiver Therapie. Eine Subgruppe der CHIP-positiven Patienten mit soliden Tumoren benötigte häufiger Dosisreduktionen und Transfusionen von Erythrozytenkonzentraten unter chemotherapeutischer Behandlung. CHIP in Stammzellspendern hatte keinen Einfluss auf das Gesamtüberleben der Empfänger, jedoch auf die kumulative Inzidenz von chronischen Graft versus Host Erkrankungen und Rückfall/Progress der primären malignen Erkrankung. In Patienten mit AAV zeigte sich eine signifikant höhere CHIP Prävalenz im Vergleich zu einer Kontrollkohorte von Gesunden. Weiterhin zeigten sich Hinweise für potenziell krankheitsmodifizierende Effekte von CHIP bezüglich der Organmanifestation, sowie bei der *in vitro* Neutrophilen-Aktivierung und der Transkriptionsregulation von ANCA-Antigenen.

Zusammenfassend liefert diese Arbeit neue Einsichten in CHIP und ihre Implikationen in der Tumortherapie, allogenen Stammzelltransplantation und bei Autoimmun-erkrankungen. Unsere Resultate bilden eine gute Grundlage für künftige translationale Studien zur weiteren Aufschlüsselung der klinischen Signifikanz von CHIP.

1 Introduction

The origin of cancer is an evolutionary process involving stepwise acquisition of independent genetic and/or epigenetic alterations in stem cells. In many cases, a pre-malignant clonal expansion of one cell population precedes the transformation to overt malignant growth. Mutation acquisition leading to acute myeloid leukemia (AML) often shows a particular pattern: an initial "founding" mutation in hematopoietic stem cells (HSC), often affecting epigenetic regulator genes such as *DNMT3A*, *TET2* or *ASXL1*, leads to the expansion of a premalignant founding clone. This population of cells exhibits a fitness advantage compared to their wild type counterparts and acquires further genetic alterations over time before certain "proliferative" mutations in genes such as *NPM1*, *FLT3,* or *NRAS/KRAS* trigger progression to full-blown acute leukemia.[1,2] It has been shown that the founding clone can persist after chemotherapy and can cause relapse years after successful treatment of the initial leukemia.[3]

In 2014, three large sequencing studies including 32,290 individuals reported that mutations in leukemia-associated genes frequently occur in elderly healthy individuals: more than 10% of the population of age greater than 60 harbor a mutated clone without showing any signs of aberrant bone marrow function. This phenomenon was termed clonal hematopoiesis (CH). It is strongly correlated with age, affects primarily epigenetic regulator genes such as *DNMT3A*, *TET2* and *ASXL1* and is associated with increased risk for hematologic cancer and reduced overall survival (OS, hazard ratio (HR) for death 1.4).[4-6] To distinguish this condition from myelodysplastic syndromes (MDS), the entity "clonal hematopoiesis of indeterminate potential" (CHIP) was proposed and defined as the presence of a somatic mutation associated with hematologic neoplasia at a variant allele frequency (VAF) of ≥ 2 % in the peripheral blood (PB) of individuals without morphological evidence for a hematologic neoplasm.[7] The overlap with idiopathic cytopenia of unknown significance (ICUS) is termed clonal cytopenia of unknown significance (CCUS). The cut-off for the VAF in this definition is arbitrary and to some extent owed to the limits of commercially available sequencing techniques at the time of definition. As it turns out, somatic mutations in blood cancer associated genes are almost ubiquitous if sequencing depth is increased and sequencing errors are minimized appropriately.[8] However, it seems reasonable to define a cut-off level of 2 % VAF, as these clones seem to have a manifest selective advantage and disproportionally contribute to the production of PB cells. Since its

molecular description in 2014, a lot of research has been devoted to the investigation of the impact of CHIP on outcomes in different clinical settings, as will be briefly summarized in the following.

CHIP is a precursor condition for hematologic malignancies. In fact, CHIP increases the risk for future hematologic malignancy by a factor of ten. The overall risk for progression to hematologic cancer is estimated 0.5-1 % per year, rendering CHIP a pre-malignant state comparable to monoclonal gammopathy of unknown significance (MGUS) for multiple myeloma or monoclonal B-cell lymphocytosis for chronic lymphatic leukemia. However, this rate is low and most individuals with CHIP never develop any hematologic disorder. Moreover, the increased risk for hematologic disease does not completely account for the reduced OS of individuals with CHIP.[5,6] Surprisingly, it was pointed out by Jaiswal and colleagues that the reduced OS of individuals with CHIP is mainly caused by death due to cardiovascular disease such as coronary heart disease and ischemic stroke.[5] This observation was confirmed by an independent retrospective case-control study, where CHIP was associated with a doubled risk for coronary heart disease.[9] One mechanism linking CHIP to cardiovascular disease was found to be an altered monocyte/macrophage function leading to dysregulation of the IL-1β/NLRP3 inflammasome in a model of *Tet2*-deficient mice fueling the progression of atherosclerosis[10] and heart failure.[11]

Pre-existing CHIP strongly raises the risk for therapy-related myeloid malignancies. A small fraction of patients surviving radio-/chemotherapeutic treatment develops a therapy-related myeloid neoplasm (t-MN) such as t-MDS or t-AML as a late complication of the treatment of their primary malignant disease. T-MNs constitute a distinct entity, as they are often chemotherapy-resistant, show a distinct mutational landscape and are associated with a dismal overall prognosis. The mutational spectrum of t-AML is enriched for *TP53* mutations compared to *de novo* AML. It was shown that in many cases of t-MN, a premalignant *TP53*-mutated clone is already present by the time of primary chemotherapy administration, indicating a selection process of the *TP53*-mutated clone during treatment which several years later can result in t-MN.[12] More recent studies demonstrated that CHIP increases the risk for t-MNs in patients with non-hematologic cancer undergoing chemotherapy[13,14] as well as in patients with lymphoma undergoing autologous stem cell transplantation.[15] In this context, mutations in genes involved in the DNA-damage repair machinery (in particular *TP53* and *PPM1D)* seem to play a particular role.

In this work, we aimed to study CHIP in different patient collectives and clinical contexts. First, we characterize the lineage distribution of CHIP mutations within the hematopoietic differentiation tree as well as the cell of origin using targeted deep sequencing of six flow-sorted PB and five bone marrow (BM) cell fractions in a subset of 365 elderly patients. While these experiments contribute to a more thorough static picture of CHIP, we next aimed to describe the dynamic evolution of CHIP clones over time. As argued above, CHIP clones might exhibit a selective advantage under the evolutional pressure of myelotoxic chemotherapy. With this in mind, we longitudinally studied the evolution of CHIP clones in patients with solid cancer undergoing (radio-)chemotherapy and determined the influence of CHIP status on chemotherapy-associated complications.

In a second study, we assessed the role of donor CHIP in the context of allogeneic stem cell transplantation (HSCT). In this scenario it is interesting to determine the dynamics of CHIP clones when being transplanted to a recipient having received conditioning therapy and to study the impact of donor CHIP with respect to OS, non-relapse mortality (NRM), incidence of relapse/progression and the incidence of transplantation-associated complications such as Graft-versus-host disease (GVHD) and cytomegalovirus reactivation.

While there is increasing evidence that CHIP confers an alteration of the inflammatory function in the context of cardiovascular disease, little is known about the role of CHIP in the context of autoimmune diseases. The last part of this work is devoted to the impact of CHIP in ANCA-associated vasculitides (AAV) as a model for autoimmune diseases and inflammatory conditions (ADIC). AAV are a group of necrotizing vasculitides of the small vessels including granulomatosis with polyangiitis (GPA), eosinophilic granulomatosis with polyangiitis (EGPA) and microscopic polyangiitis (MPA). Upon binding to the auto-antigens PR3 or MPO expressed on the cell surface, ANCA (antineutrophil cytoplasmatic antibodies) trigger the activation of neutrophils leading to severe endothelial damage via various inflammatory mechanisms. As CHIP mutations are mainly found in the myeloid lineage and monocytes and neutrophils are the main ANCA-autoantigen expressing cells, a potential impact of CHIP on the pathogenesis and clinical manifestation of AAV is conceivable. Here, we paradigmatically studied CHIP in 112 patients suffering from AAV with respect to prevalence, clinical outcome and *in vitro* measurement of ANCA-induced neutrophil activation and ANCA-antigen messenger ribonucleic acid (mRNA) expression.

2 Methods

2.1 Patient Collectives

For the first part of this work, PB samples were collected from 437 patients hospitalized in the Campus Virchow Klinikum, Charité, Berlin, from June 2015 to October 2017. 365 patients, in the following referred to as the “non-cancer cohort”, were hospitalized either in the Department of Nephrology or Cardiology of Charité (335 patients) or at the Charité Centre for Musculoskeletal Surgery (30 patients). For the latter 30 patients, PB and bone marrow (BM) samples as leftovers from hip replacement were available. Inclusion criteria were (a) age ≥ 55 years, (b) no known hematologic disease, (c) no active infectious disease and (d) no history of malignant disease.

The remaining 72 patients, in the following referred to as the “cancer cohort”, were recruited on the oncology ward or at the oncology outpatient clinic of the Charité Campus Virchow Klinikum. Inclusion criteria (a)-(c) were identical to the non-cancer cohort, however, these patients had been recently diagnosed with a solid tumor (gastroesophageal, biliopancreatic, sarcoma, colorectal or neuroendocrine) requiring (radio-)chemotherapy. PB samples in the cancer cohort were collected before the initiation of chemotherapy and at regular intervals (every 2-3 cycles) directly before the administration of chemotherapy. Blood counts were required to be sufficiently regenerated for chemotherapy administration at all sampling points.

Patient data including demographic and clinical data as well as blood counts at the sampling time points were extracted from the (electronic) patient records. For the cancer cohort, also the occurrence of chemotherapy-related complications (neutropenic fever, dose reduction due to hematotoxicity, unscheduled G-CSF-administration, delay of chemotherapy administration) were extracted. Transfusion history data was extracted from the local electronic database of the Institut für Transfusionsmedizin at the Charité Universtitätsmedizin Berlin. Transfusions in the context of surgery were not counted as chemotherapy-related complications. All patients gave their written informed consent according to the Declaration of Helsinki. The study was approved by the local ethics committees.[16]

In the second retrospective multicenter study, a cohort of healthy, older donors (≥ 55 years) for allogenic HSCT with available PB or BM samples at the time of stem cell donation was investigated, in the following referred to as the “donor cohort”. Samples from 500 related donor-recipient couples with sufficient patient data were available.

They were included from ten transplantation centers in Germany and France. Samples were collected between 1993 and 2017. The study was conducted in accordance with the Declaration of Helsinki and ethical approval was obtained from the local ethics committees.[17]

For the third part of this work, PB samples of 112 patients with diagnosed AAV treated at the Charité Universitätsmedizin Berlin Campus Mitte or at the HELIOS Klinikum Buch between 04/2005 and 10/2018 were analyzed (in the following called the "AAV cohort"). There was no lower age limit for the inclusion of AAV patients. Demographic and clinical data were extracted from written and/or electronic patient records. All patients gave their informed consent and the study was conducted in accordance with the Declaration of Helsinki. Ethical approval was obtained from the local ethics committees.[18]

2.2 Sample Processing of Fresh Peripheral Blood and Bone Marrow

10 - 20 ml of PB were taken during routine diagnostics and further processed as follows. 1 - 2 ml PB were used for DNA extraction: the sample was treated with red blood cell lysis buffer at 0 °C for 20 min before DNA was extracted with Qiagen's QIAamp DNA Minikit according to the manufacturer's recommendations (Qiagen, Hilden, Germany). DNA concentrations were measured with a Promega QUANTUS™ fluorometer (Promega, Madison, Wisconsin, USA) and stored at -20 °C.

From the remaining PB, mononuclear cells (MN) were enriched via Ficoll density centrifugation, washed twice with phosphate buffered saline (PBS) and stored at -196°C in liquid nitrogen. The pellet residing on the ground after Ficoll density centrifugation (containing erythrocytes, thrombocytes and granulocytes) was treated with red blood cell lysis buffer at 0 °C for 20 min, washed twice with PBS and stored in liquid nitrogen at -196 °C.[16]

2.3 Targeted Sequencing

Libraries of whole-blood DNA samples were created using Illumina's Trusight Myeloid Panel Kit or a customized version thereof according to the TruSight DNA Amplicon Sequencing Panel Guide (Part # 15054779 Rev. B, Illumina, San Diego, California, USA). A detailed list of the target regions of the panels used can be found in the supplementary material of the respective references.[16,17] Hybridization of oligos with

50 ng sample DNA, removal of unbound oligos, extension-ligation of bound oligos and polymerase chain reaction (PCR) amplification were performed according to the manufacturer's protocol for batches of 4-16 libraries. In the amplification process, libraries were tagged with different combinations of Illumina's i5/i7 multiplexing index sequences. Successful library amplification was confirmed for every library by gel electrophoresis (2 µl library aliquot on 4 % agarose gel or 1 µl library aliquot on the Agilent Bioanalyzer System (Agilent, Santa Clara, California, USA)). Libraries were stored at a safe stopping point at -20 °C before further processing. In a next step, PCR products were cleaned using AMPure XP beads (Beckman Coulter, Brea, California, USA), libraries were normalized before being pooled run-wise using 14 µl of the pooled amplicon library together with 586 µl of Illumina's HT1-buffer. Pools were paired-end sequenced on a MiSeq or HiSeq sequencer using the MiSeq 300 cycle v2 Reagent Kit or the HiSeq Rapid Duo cBot v2 Sample Loading Kit, respectively. In order to guarantee for sufficient sequencing depth, at most 32 samples were pooled per MiSeq run.[16-18]

2.4 Sequencing Data Analysis

Demultiplexing and FastQ-file generation were performed using Illumina's standard software. FastQ files were further processed as follows: genetic sequences were aligned to the human genome (GRCh37) with additional Epstein-Barr virus and decoy sequences (hs37d5) using the Burrows-Wheeler algorithm (BWA-MEM version 0.7.15-r1140)[19] and locally realigned around insertions/deletions (indels) with the GATK IndelRealigner (3.6-0-g89b7209).[20] For the identification of genetic variants, a combination of two well-established variant calling tools, SAMtools (version 1.3.1)[21] and VarScan (version 2.4.2)[22], was used. The resulting variants were annotated with data from the Single Nucleotide Polymorphism Database (dbSNP) (b146/GRCh37p13) of single nucleotide polymorphisms (SNPs) and the COSMIC database (COSMIC v76) of reported somatic mutations using Jannovar (version 0.16)[23]. The list of called variants was further processed in order to filter for those mutations of relevance in CHIP. To this end, we developed an R-based algorithm excluding irrelevant variant calls: (1) intronic and synonymous variants, (2) variants with a VAF below 2 %, (3) variants which appear in more than 10 % of the samples (as they are most likely sequencing errors), (4) variants with a VAF between 45 % and 55 % or VAF > 95 % and reported as SNPs in dbSNP were excluded, if not part of the following list of

recurrent mutations in myeloid malignancies: *DNMT3A* R882H/C, *JAK2* V617F, *SF3B1* K700E or K666N, *SRSF2* P95L, *U2AF1* S34F or R156H. The remaining set of potential variants was manually assessed in the Integrated Genome Viewer (IGV, Broad Institute, version 2.3) and selected for validation.[16-18]

2.5 Validation and Longitudinal Quantification

Given the set of potential variants resulting from the procedure in the previous paragraph, a second step for identification of false positive variants was performed. This was done either with an amplicon-based deep sequencing approach or, in case of specific mutations (*DNMT3A* R882C and R882H, *JAK2* V617F and *U2AF1* S34F), with digital droplet PCR. Here, only the former approach is described, the latter is outlined in detail in the respective publication.[16] In a first step, primers for the variant candidates were designed using Primer3web (version 4.1.0)[24,25], in a way that the resulting amplicon had a length of 150 - 200 basepairs (bp) and the suspected mutation had a distance of more than 30 bp to the amplicon's ends. Primers were synthesized by Eurofins (Eurofins Scientific, Luxembourg, Belgium). The amplification PCR was established using a standard touch-down protocol along the lines of reference[26] on a Bio-Rad MyCycler (Bio-Rad laboratories, Berkeley, California, USA). Amplicons were generated from patient sample DNA and successful amplification was confirmed via gel electrophoresis. Each amplification process was controlled for unspecific amplification with H_2O instead of sample-DNA for all primers. Amplicons were then distributed to different pools under avoidance of overlap between two or more amplicons in the same pool using an R-based algorithm. The pools were cleaned up using AMPure XP beads, tagged with i5/i7 indexes from the NEBNext Ultra DNA Library Prep Kit (New England Biolabs, Massachusetts, USA) and paired-end sequenced on a MiSeq System with the MiSeq Reagent Kit v2 (300 cycles, Illumina). The bioinformatic processing of the validation data was to most extent analogous to the screening. Since variant candidate positions were known, base counts (and the resulting VAFs) of only those positions were selectively extracted and manually reevaluated in IGV. A mutation was classified as validated if the sequencing depth was higher than 100 reads and the VAF in the validation process was found to be above 1 %. Candidates with insufficient sequencing depth were repeated. A visualization of the complete sequencing process is shown in Figure 1.

2.6 Flow-Sorting of Peripheral Blood and Bone Marrow Cells

A detailed description of the methods and materials for flow sorting can be found in the methods section, Supplementary Table S2 and Supplementary Figure S2 of Arends et al. 2018.[16] In summary, frozen mononuclear cells from PB/BM were thawed and incubated overnight in RPMI medium (Thermo Fisher Scientific, Waltham, Massachusetts, USA) at a temperature of 37 °C and 5 % CO_2. PB cells were stained with anti-CD14 (APC; eBioscience, San Diego, California, USA), anti-CD3 (FITC; Becton Dickinson, Franklin Lakes, New Jersey USA), anti-CD56 (PerCP-Cy5.5; Becton Dickinson), anti-CD19 (PE-Cy7; Becton Dickinson), anti-CD66b (PE; Becton Dickinson) and anti-CD34 (PE; Becton Dickinson) or anti-CD45 (PerCP-Cy5.5; eBioscience) and anti-CD66b (PE; Becton Dickinson) in the case of granulocytes. Dead cells were labeled via DAPI-staining. The cells were sorted into CD34+ progenitor cells, CD19+ B-cells, CD3+ T-cells, CD14+ monocytes, CD56+ natural killer cells (NK-cells) and CD66b+ granulocytes on a BD Aria II SORP cell sorter (Becton Dickinson). BM cells were lineage-depleted using the Human Hematopoietic Lineage Antibody (eBioscience) before being sorted into hematopoietic stem cells (HSC), multipotent progenitors (MPP), lymphoid-primed multipotent progenitors (LMPP), granulocyte-macrophage progenitors (GMP), common myeloid progenitors (CMP, for a detailed survey of antibodies and gating strategy, see Supplementary Material of the respective publication).[16]

2.7 Neutrophil Activation Assays and mRNA Expression

RNA was extracted from stored PAXgene blood RNA tubes using the PAXgene Blood RNA Kit (Qiagen). Reverse transcription of 1 µg RNA was conducted using the Thermo Scientific RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). Real-time quantitative PCR (qPCR) with specific primers (Biotez, Berlin, Germany) for *CD177, PR3, MPO, RUNX3, JMJD3* and the house keeping gene *18S* was performed on a QuantStudio3 Real-Time PCR System (Applied Biosystems, Foster City, California, USA) using Fast SYBER Green Master Mix (Applied Biosystems) or Taqman Fast Universal PCR Master Mix (Applied Biosystems). After normalization to *18S*, gene expressions were analyzed with the $2^{-\Delta\Delta Ct}$ method.[27] Expression data of healthy controls were taken as a reference and normalized to a mean of 1.[18]

ANCA autoantigen PR3 and CD177 expression was flow-cytometrically measured as described in Schreiber et al. 2010.[28] The membrane expression of PR3 or CD177 is

given by the expression index EI = ($MFI_{stimulated\ cells}$-$MFI_{unstimulated\ cells}$)/$MFI_{unstimulated\ cells}$, where MFI is the mean fluorescence intensity, or as the percentage of membrane-expressing cells. Neutrophil reactive oxygen generation was measured using the dihydrorhodamine (DHR) oxidation assay as detailed in Arends et al. 2019.[18] The result is expressed as stimulation index SI = $MFI_{stimulated\ cells}$/$MFI_{unstimulated\ non\text{-}primed\ cells}$.

2.8 Statistical Analysis

The statistical analysis was performed using IBM SPSS statistics (version 23), R (version 3.3.1), and RStudio (version 0.99.463). If not stated otherwise, groupwise comparison of categorical variables were carried out using Fisher's exact test and groupwise comparison of continuous variables were carried out using Mann-Whitney tests without corrections for multiple hypothesis testing. The two-sided level of significance was set to $P < 0.05$.

In the analysis of CHIP as a potential predictor for complications occurring during follow-up, baseline parameters with unadjusted univariate P-values < 0.1 with respect to the outcome of interest were accounted for as covariates in a binary logistic regression analysis.

Survival analysis of OS was carried out using the Kaplan-Meier method and hazards were compared between groups using the log-rank test. For the cumulative incidence of relapse/progression (CIR/P, time from diagnosis to relapse or progression of the disease), NRM and occurrence of acute or chronic GVHD, competing risk analyses were conducted using Gray's test.[29] Possible confounding baseline parameters were identified as described above and accounted for in a multivariate analysis according to Fine and Gray.[30]

3 Results

3.1 Sequencing Results in the Non-Cancer and Cancer Cohort

A total of 437 patients (cancer and non-cancer cohort) of median age 75 (range 55 – 98) were screened for the presence of CHIP. In the screening process, 201 potential somatic variants were found, 168 of which were confirmed in a second independent validation experiment (168/201 = 83.6 %) affecting 121 patients (121/437 = 27.7 %; 99/365 = 27.1 % in the non-cancer cohort and 22/72 = 30.6 % in the cancer cohort, see Figure 2A). VAFs from the validation sequencing process strongly correlated with screening VAFs (Pearson coefficient R = 0.93, see Supplementary Figure S3 in Arends et al 2018).[16] CHIP prevalence and number of mutations significantly increased with age (Figure 2B). While 87 patients harbored a single mutation, multiple mutations occurred in 34 patients: 23 patients were identified with two concurrent mutations, nine patients with three mutations and two patients with four mutations. In accordance with previous CHIP studies[4-6], the most frequently mutated genes were the epigenetic regulator genes *DNMT3A* (59/437 = 13.5 %), *TET2* (34/437 = 7.8 %) and *ASXL1* (11/437 = 2.5 %, Figure 2C), accounting for 70.2 % of all mutations, as well as *SF3B1* (9/437 = 2.1 %) and *TP53* (8/437 = 1.8 %). Of note, *TET2* mutations were more often truncating than *DNMT3A* mutations ($P < 0.001$) and were more frequently than *DNMT3A* mutations accompanied by a second somatic mutation (47.1 % vs. 30.5 %; $P < 0.001$). No significant differences regarding CHIP prevalence or mutational spectrum were found between the cancer and non-cancer cohort.[16]

CHIP-positive patients were significantly older than CHIP-negative patients (median age 76.4 years vs. 73.5 years; $P < 0.001$). While CHIP-positive patients had higher numbers of neutrophils (6.0×10^9/l vs. 5.2×10^9/l; $P = 0.022$), no significant differences were found in other blood count values.[16]

3.2 Hematopoietic Lineage Distribution

To gain insight into the lineage distribution within the hematopoietic tree we determined the VAFs of 91 mutations in 63 patients in flow-sorted CD34+ progenitor cells, monocytes, granulocytes, NK-cells, B-cells and T-cells. With one exception, all clones showed a predominant expansion in the myeloid compartment over the lymphoid compartment: B-cells and T-cells showed a significantly lower VAF compared to any other cell fraction ($P < 0.001$ for every comparison, Figure 3). NK-cells, however, which

also derive from lymphoid progenitor cells, had a median VAF comparable to cells from the myeloid compartment. Interestingly, compared to other genes, median VAF in T-cells of *DNMT3A*-mutated patients was higher than in T-cells of patients with other mutations ($P < 0.001$). Apart from that, no gene-specific differences in repartition patterns were observed. For nine CHIP-positive patients, BM samples were available and, in all of these cases, mutations were detected in the Lin-CD34+CD38 HSC fraction.

3.3 Evolution of Mutated Clones under Cytotoxic Treatment

We longitudinally quantified the VAFs of 32 mutations in 22 cancer patients over a median follow-up time of 9 months (range 2 - 20 months). Clonal behaviour was stratified according to three categories into (1) rising clones, that show an increasing relative VAF change of greater than 50 % from the initial value, (2) decreasing clones that show a negative relative VAF change of more than 50 % compared to the initial value, and (3) stable clones that do not meet the criteria for (1) or (2). According to this classification, increasing VAFs were found for *TET2* (n = 2/5), *RAD21* (n = 2/2), *PPM1D* (n = 1/2), and *EZH2* (n = 1/2) mutations. Decreasing VAFs were seen in *SF3B1* (n = 3/3), *JAK2* (n = 1/1), *CBLB* (n = 1/1), *DNMT3A* (n = 1/13), and *TET2* (n = 1/5). Interestingly, only one out of 13 *DNMT3A*-mutated clones showed a non-stable behaviour, while 13/19 clones with mutations in other genes showed major increases or decreases (7.7 % vs. 68.4 %; $P < 0.001$).[16]

Clinical data were available for 70 patients in the cancer cohort with a median follow-up time of 10 months. Among chemotherapy-related complications (erythrocyte transfusion necessity, neutropenic fever, dose reductions and delay of chemotherapy administration due to hematotoxicity), dose reductions occurred more frequently in the CHIP-positive group compared to the CHIP-negative group (7/20 vs. 4/46; $P = 0.035$ in a binary regression analysis after correction for gender and administration of radiotherapy as potential confounders). In the group of patients with a mutation other than *DNMT3A,* more erythrocyte transfusions were administered along the course of the treatment (10/15 vs. 13/46; $P = 0.014$ in a binary regression after correction for administration of radiotherapy, topoisomerase inhibitors or alkylating agents) compared to the group of CHIP-negative patients.[16]

3.4 Clonal Hematopoiesis in Allogeneic Stem Cell Transplantation

Among 500 donors (median age 64.5, range 55 - 79, see Figure 2A), 92 mutations were found in 80 patients (80/500 = 16.0 %). *DNMT3A*, *TET2,* and *ASXL1* were the most frequently mutated genes with a prevalence of 8.0 %, 2.2 % and 1.4 %, respectively (Figure 2C). 70 donors had a single mutation, nine donors had two and one donor had four mutations. CHIP prevalence was found to increase with age from 10.3 % in the group of 60 - 64 up to 28.6 % in the group 75 - 79 years old donors.[17]

Next, we analyzed the impact of CHIP on OS and other clinical outcomes after allogeneic HSCT. No significant differences in OS between recipients transplanted from CHIP-positive or CHIP-negative donors were found in the Kaplan-Meier survival analysis (HR 0.88, 95%-CI 0.65 - 1.32). However, cumulative incidence of chronic GVHD was higher in the group of recipients with transplants from CHIP-positive donors (5-year cumulative incidence 52.9 % vs. 35.7 %; P = 0.008). In a multivariate analysis, donor CHIP was identified as an independent risk factor for the development of chronic GVHD (HR 1.73, 95%-CI 1.21 - 2.49; P = 0.003) after correcting for ATG application and donor age as potential confounders. By restricting the analysis to individual genes, *DNMT3A* mutations in the graft were identified as the main CHIP factor for the development of chronic GVHD (HR 1.99; 95%-CI, 1.28 - 3.12; P = 0.002 in a multivariate analysis). CIR/P was lower in the group of recipients with CHIP-positive transplants (HR 0.62, 95%-CI 0.41 - 0.98; P = 0.048 in a multivariate analysis according to Fine and Gray). No significant difference in non-relapse mortality was found between CHIP-positive and CHIP-negative donors.[17]

For 22 recipients the further evolution of transplanted CHIP clones was quantified during follow-up using targeted deep sequencing (Figure 4). With one exception (a clone with a *SF3B1* K700E mutation) all CHIP clones successfully engrafted. Figure 4B shows three examples of clonal evolution after transplantation: recipient UPN027 was chosen as example for disproportional expansion with the *DNMT3A*-mutated clone expanding to 23.7 % from 4.3 % in the donor. In UPN181, no such disproportional expansion occurs, the expansion pattern was therefore described as linear. In UPN443, the initial expansion of the clone gets opposed by the relapsing disease as can be seen from the parallel decrease of the clone size and chimerism. Interestingly, eleven clones showed a disproportional expansion after transplantation, at least doubling the VAFs that were measured in the donor hematopoietic system. Among these, seven were *DNMT3A*-mutated clones in contrast to the scenario of

cytotoxic treatment, where none of the increasing clones was *DNMT3A*-mutated (7/11 vs. 0/6; P = 0.018).

3.5 Clonal Hematopoiesis in Patients with ANCA-Associated Vasculitis

Investigation of 112 AAV patients for CHIP identified 46 mutations in 34 patients, corresponding to a prevalence of 30.4% in this rare autoimmune disease. As in the other cohorts, *DNMT3A*, *TET2,* and *ASXL1* were the most frequently mutated genes (Figure 2C). Even though median age was ten years below the cancer and non-cancer cohorts, CHIP prevalence was of comparable magnitude. This led us to compare the prevalence in our cohort to healthy cohorts in the literature of similar age and sequencing technology[17,31-33], the results are displayed in Figure 2A. The prevalence of CHIP in the AAV cohort was significantly higher than in each individual cohort and in the pooled healthy cohort (30.4 % vs. 13.5 %; P < 0.001; see also Supplementary Table 3 in Arends et al. 2019).[18] To substantiate this finding, we constructed an age- and sex-matched control cohort including 79 individuals of the donor cohort and 33 additional healthy individuals of age below 55 years, sequenced with identical sequencing technology. In the control cohort, 22 somatic mutations in 20 patients were found, corresponding to a CHIP prevalence of 17.9 % (17.9 % vs. 30.4 %; P = 0.042 in a two-sided Fisher test).

Follow-up samples were available for 19 patients. In longitudinally quantifying 23 clones in these patients over a median follow-up time of 2.3 years (range 0.3 - 10.9 years), we found increasing clone sizes in five patients, decreasing clone sizes in two patients and stable clone sizes in 12 patients. Three exemplary patients are shown in Figure 4C. Of note, the two mutations of UPN079 affecting *SF3B1* and *RAD21* showed differential dynamics, one increasing over time, the other remaining stable, indicating that the two mutations constitute independent clones. To investigate whether patients, which were initially CHIP-negative, develop CHIP along the course of their disease, we screened follow-up samples of 20 initially CHIP-negative patients two to ten years after the initial sampling. In none of them, a newly acquired somatic mutation was found.[18]

Next, we assessed the impact of CHIP on clinical outcomes. Positive CHIP status was associated with higher age (median 70.5 vs. 63.0; P = 0.017). No significant differences between the group of CHIP-positive and CHIP-negative patients was found with respect to diagnosis (GPA or MPA), disease activity status, previous treatment

regimens, blood counts, creatinine levels, comorbidities or relapse risk. In the group of CHIP-negative GPA patients, however, more patients showed renal involvement (88.5 % vs. 68.2 %; P = 0.049) and involvement of the nervous system (19.0 % vs. 0.0 %; P = 0.028).[18]

Neutrophil activation and ANCA-antigen mRNA expression was measured *in vitro* for a subset of CHIP-positive and -negative AAV patients and healthy controls (HCs). CHIP-positive AAV patients showed a reduced neutrophil activation upon stimulation with monoclonal MPO/PR3 antibodies compared to CHIP-negative AAV patients (median stimulation index: anti-MPO 6.29 vs. 13.01; P = 0.057; anti-PR3: 7.72 vs. 13.00; P = 0.026). No difference in membrane expression index or percentage of positive cells between the groups was observed. We further measured mRNA expression levels of *PR3*, *MPO*, *CD177*, *RUNX3*, and *JMJD3* using qPCR. Here, CHIP-negative AAV patients showed an increased expression of *MPO* and *PR3* (*MPO*: 1.94 vs. 0.86; P = 0.026; *PR3*: 2.02 vs. 0.58; P = 0.057) compared to HCs, which was less evident in CHIP-positive AAV patients (*MPO*: 1.27 vs. 0.86; P = 0.391, *PR3*: 0.92 vs. 0.58; P = 0.297).[18]

4 Discussion

We characterized CHIP in four cohorts of patients with cardiovascular disease, solid cancer, AAV, and healthy stem cell donors using a combination of targeted deep sequencing, flow-sorting and evaluation of clinical parameters. The CHIP prevalence we report in the cancer, non-cancer and AAV cohort were higher than in healthy cohorts of the recent literature (Figure 2A). However, differences in age distribution and sequencing technology, in particular in sequencing depth, extent of the target region and VAF threshold, have to be taken into account when comparing the respective results. The mutation spectrum of our cohorts was consistent with the mutational landscape of CHIP in healthy individuals as previously reported.[5,6] It differs from the spectrum in patients who have previously been treated with cytotoxic agents, which was shown to be enriched for *TP53* and *PPM1D* mutations.[15]

4.1 Lineage distribution analysis

Our lineage distribution analysis showed a myeloid-pronounced differentiation pattern with very few exceptions: we found higher VAFs in myeloid and NK-cells than in lymphoid cells. However, a myeloid skew is a known phenomenon of the aging hematopoietic system as described by Pang and colleagues[34] and is not necessarily related to CHIP. Our data does not permit conclusions about a possible causal relation between myeloid differentiation bias and CHIP. This could in principle be further investigated by additionally determining the lineage distribution of clones with synonymous SNVs and compare them to those of CHIP clones. Similar results have been reported by Buscarlet and colleagues, who categorized lineage restriction patterns into myeloid, myeloid-lympho-B and multipotent, depending on the involvement of myeloid cells, B-cells and T-cells, respectively. 38 % of their patients were categorized myeloid and 40 % myeloid-lympho-B.[35]

Interestingly, VAFs in T-cells of *DNMT3A*-mutated patients were significantly higher compared to those of *TET2*-mutated patients. This was also noted by Buscarlet and colleagues, who found that none of their *TET2*-mutated patients was classified into the multipotent group in contrast to 36 % of the *DNMT3A*-mutated patients. It is important to note, however, that mutations in the T-cell compartment were only considered for analysis if the respective VAF was > 2 %. With our method, we were able to detect VAFs far below 2 %, showing that T-cells are affected also in *TET2*-mutated patients,

though with lower mutational burden (median VAF 0.3 %). A simple explanation would be that *DNMT3A* occur in less committed HSCs than *TET2* mutations. Alternatively, it is conceivable that *TET2* mutations and *DNMT3A* mutations both occur in multipotent HSCs but the myeloid bias conferred by *TET2* mutations is more pronounced.[35] Future preclinical models using genome editing might provide further information about the underlying mechanisms.

4.2 Clonal Hematopoiesis in Patients Receiving Chemotherapy

In the study of CHIP in patients with solid tumors receiving cytotoxic chemotherapy we found different dynamics of CHIP clones. In fact, in four out of 22 patients, clone size increased to more than 1.5-fold of the initial size within months as we initially hypothesized. In the remaining 18 patients, no relevant change or decrease in clone size was measured. Interestingly, 12/13 *DNMT3A*-mutated clones remained stable during chemotherapy. Most likely, gene-specific as well as therapy-specific effects determine the evolutionary process of CHIP clones during cytotoxic treatment. It was shown by Gibson and colleagues that patients after cytotoxic treatment more frequently harbor *TP53* or *PPM1D* mutations compared to chemotherapy-naïve patients. *PPM1D*-mutated patients had received higher cumulative doses of doxorubicin compared to patients with other mutations.[15] Moreover, the mutational spectrum of t-AML differs from the spectrum of *de novo* AML[36-40], indicating that previous cytotoxic treatment induces and/or selects mutated clones with specific gene mutations. Regardless of the clonal evolution, our data provide evidence that CHIP-positive patients are more likely to require dose reductions and erythrocyte transfusions during chemotherapy. A possibly related effect was reported recently by Murphy and colleagues who showed that the time to platelet and neutrophil recovery was longer in AML patients in complete remission with persisting mutations in *DNMT3A*, *TET2*, *ASXL1,* or *SRSF2*, indicating a compromised capacity of repopulating the peripheral blood by mutated HSCs after exposure to cytotoxic chemotherapy.[41] It is important to keep in mind that the robustness of our results is limited by the small patient number, diversity of tumor entities, diversity of treatment regimens and the diversity of mutated genes. Our findings will have to be reevaluated in larger and more homogeneous cohorts before being able to draw firm conclusions.

4.3 Clonal Hematopoiesis in Allogeneic Stem Cell Transplantation

In the second part we studied CHIP in the setting of allogeneic HSCT. The prevalence of CHIP in the donor cohort was lower than in the other cohorts, most likely due to lower age and the absence of severe disease that would preclude individuals from stem cell donation. In contrast to the setting of autologous HSCT in which CHIP was found to be associated with inferior OS[15], CHIP status of the donor did not affect the OS of recipients in our cohort. The difference in OS in the study of Gibson and colleagues was mainly driven by NRM (in particular t-MN and cardiovascular events).[15] The reason we do not observe differences in OS in our cohort might be, that the median OS was much shorter – as expected for patients after allogeneic HSCT compared to autologous HSCT – and cardiovascular events play a subordinate role within this time horizon. On the other hand, CHIP and in particular *DNMT3A* mutations had an impact on CIR/P and chronic GVHD. While the risk for CIR/P was lower in the presence of donor CHIP, the risk for chronic GVHD was increased. This highly interesting finding is in line with previous data, linking CHIP with aberrant inflammatory response of blood cells.[9,10]

Concerning the fate of CHIP clones in the recipients' hematopoietic systems, it is striking that seven out of twelve *DNMT3A*-mutated clones showed a disproportional expansion after allogeneic HSCT, whereas none of the *DNMT3A*-mutated clones of patients receiving cytotoxic treatment expanded (7/12 vs. 0/13; P = 0.002). It is tempting to speculate that *DNMT3A* mutations might confer a competitive advantage under the selective pressure of stem cell transplantation, but not in response to cytotoxic treatment.

In summary, allogeneic HSCT from elderly donors with CHIP does not affect the OS of recipients but has implications on CIR/P and incidence of chronic GVHD. Further prospective studies including also younger and unrelated donors are necessary to derive reliable clinical implications.

4.4 Clonal Hematopoiesis in Patients with ANCA-Associated Vasculitis

Our last study examined the role of CHIP in 112 patients with AAV. We found a surprisingly high prevalence of CHIP compared to recently published healthy cohorts. We were able to confirm this observation by constructing a 1:1 age- and sex-matched cohort sequenced with identical technology (30.4% vs. 17.9%, P = 0.042), revealing a new association between CHIP and AAV. To our knowledge, no association of CHIP

with autoimmune diseases has been reported in the previous literature. A study of 54 patients with rheumatoid arthritis by Savola and colleagues did not reveal any unexpected findings concerning frequency, mutational spectrum, or clinical manifestations.[42] Potential explanations for the high prevalence in our study are diverse. One directly arising question is whether there is a causal relation between CHIP and AAV and if so, whether CHIP-induced aberrant cellular immune response might facilitate the outbreak of AAV or whether the chronic inflammatory state in AAV patients fuels the progression of CHIP. Of note, an association with autoimmune diseases has long been known in patients with MDS.[43-45] Patients with a history of autoimmune disease (in particular systemic vasculitides) have a higher risk of developing MDS and/or AML.[46,47] It seems plausible that a history of chronic inflammation triggers MDS, however, MDS is often preceded by CHIP/CCUS[48,49] and it is also conceivable that CHIP leads to an aberrant inflammatory state favoring the development of autoimmune diseases. This issue can only be resolved in large longitudinal cohort studies (which are difficult due to the very low incidence of AAV) in combination with mechanistic studies in cell culture and animal models.

Clinically, CHIP-positive AAV patients showed less renal and less nervous system involvement, sparking the idea of CHIP as a disease-modifying factor. *In vitro*, we found reduced ANCA-induced activation in neutrophils from CHIP-positive compared to neutrophils from CHIP-negative AAV patients and HCs. It was recently shown by Schreiber and colleagues that the production of reactive oxygen species (ROS) downregulates inflammasome activation by inhibiting the inflammasome-caspase-1-IL1β-cascade.[50] Hence, a reduced neutrophil activation might lead to an overactivation of the inflammasome and thereby potentially constitutes an additional pathogenic mechanism in the development of AAV in CHIP-positive patients. It should be noted though, that our *in vitro* measurements are impacted by differing baseline characteristics of the examined groups (e.g. age) as well as differing blood counts. Age- and sex-matched groups of patients at an equivalent point of their disease history (e.g. diagnosis before treatment) should be investigated to challenge these results. Nevertheless, our data provides novel evidence for an association and potentially disease-modifying effects of CHIP in AAV.

4.5 Summary and Outlook

In summary, CHIP is a highly interesting phenomenon with implications in malignant, cardiovascular and autoimmune diseases and is worth further investigating in the context of personalized medicine. Early risk stratification, closer monitoring and targeted treatment options for patients at risk are just some of numerous conceivable translational applications. Detailed prospective and mechanistic *in vitro* studies are warranted to further decipher the influence of CHIP mutations on proliferative behavior and immune function and to differentially correlate individual gene mutations with adverse outcomes. Promising new methods such as single cell DNA sequencing[51] and error corrected sequencing[52-54] will help to more accurately study the clonal complexity of the aging hematopoietic system and the mechanisms of leukemogenesis.

Figures

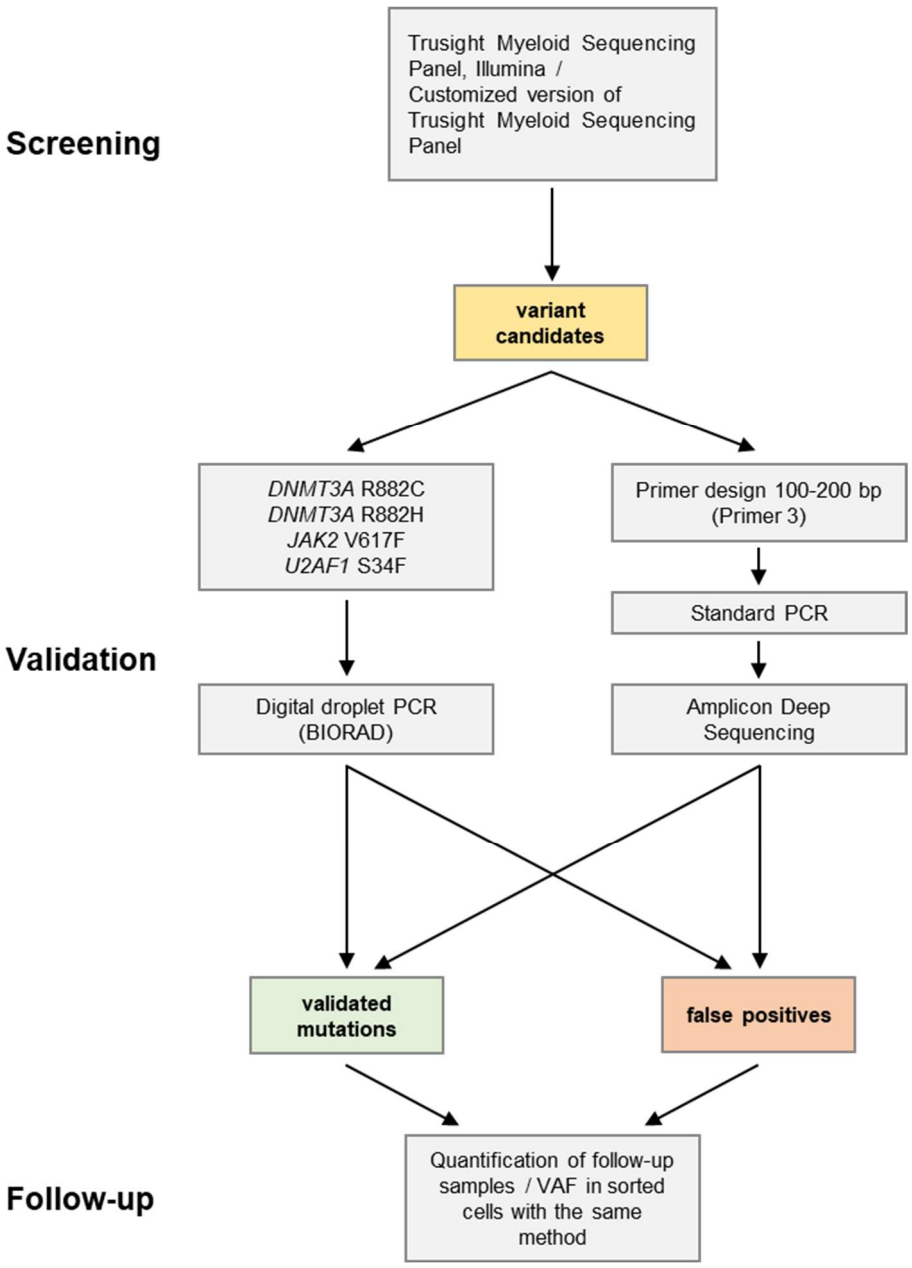

Figure 1. *Visualization of the multi-step sequencing process*. Sequencing was performed in three steps, including (1) a screening phase, where samples with potential CHIP mutations were identified, (2) a validation phase, where variant candidates were validated in an independent deep sequencing experiment and (3) the quantification of mutational burden in the follow-up samples/different cell lines.[16]

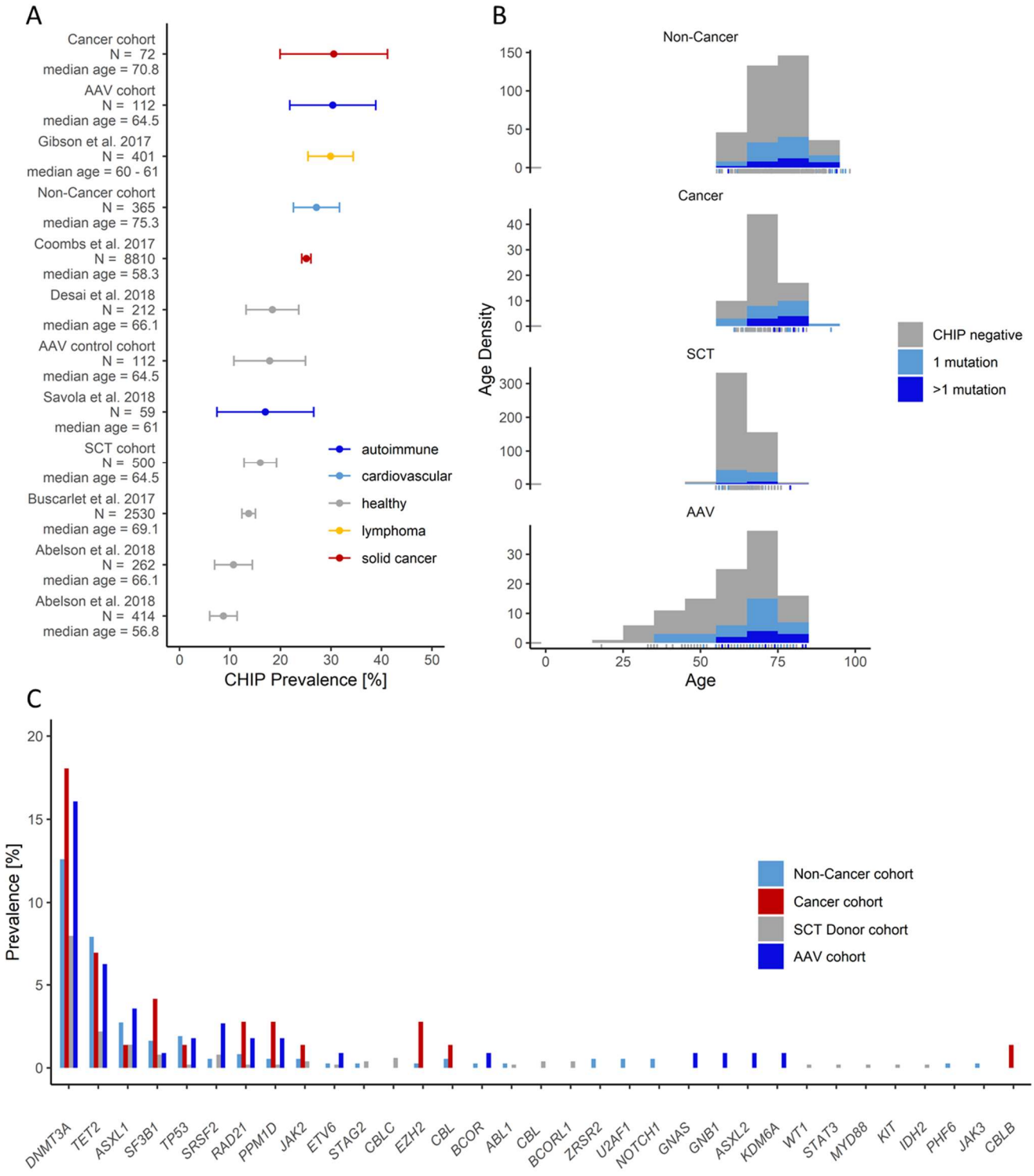

Figure 2. *Overview of sequencing analysis.* (A) CHIP prevalence reported in different patient cohorts of this work and cohorts described in recent publications.[15-18,31-33,42] Error bars depict 95%-confidence intervals. (B) Age distributions of the four cohorts of this work. The fraction of patients with one mutation is coloured in light blue, the fraction of patients with multiple mutations in dark blue. (C) Mutational spectrum of CHIP found in the cohorts of this work.[16-18]

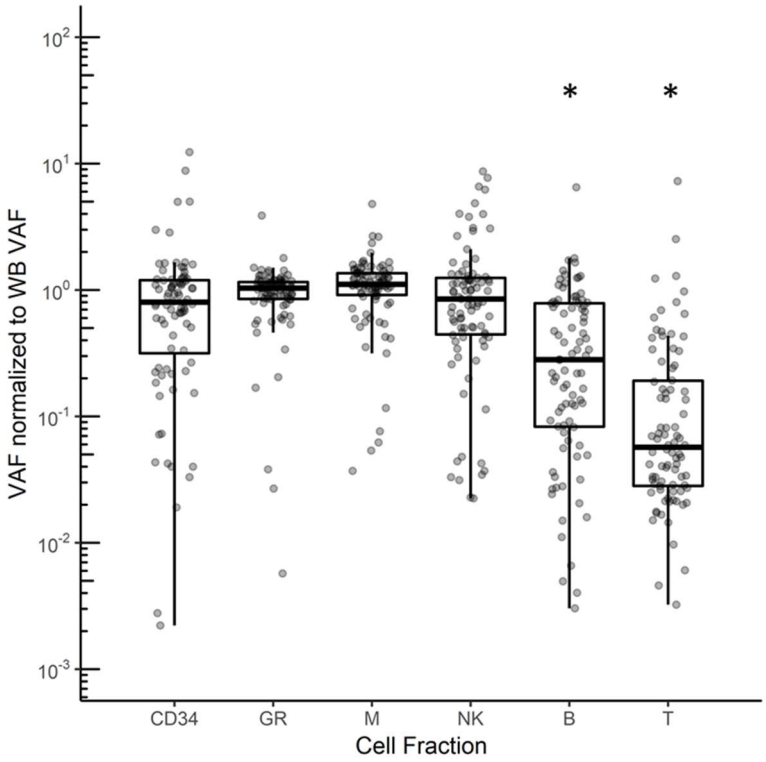

Figure 3. *Variant allele frequencies in sorted cell fractions*. VAFs of 91 mutations in 63 patients are depicted as scatter plots and summarized in box plots. VAFs were normalized to the value for the whole blood VAF and are shown on a logarithmic scale. Normalized VAFs in B- and T-cells were significantly lower than in the other four cell lines ($P < 0.001$ for each comparison) as indicated by the Asterisk.[16] Abbreviations: VAF = variant allele frequency, WB = whole blood, CD34 = CD34-positive cells, GR = granulocytes, M = monocytes, NK = natural killer cells, B = B-cells, T = T-cells.

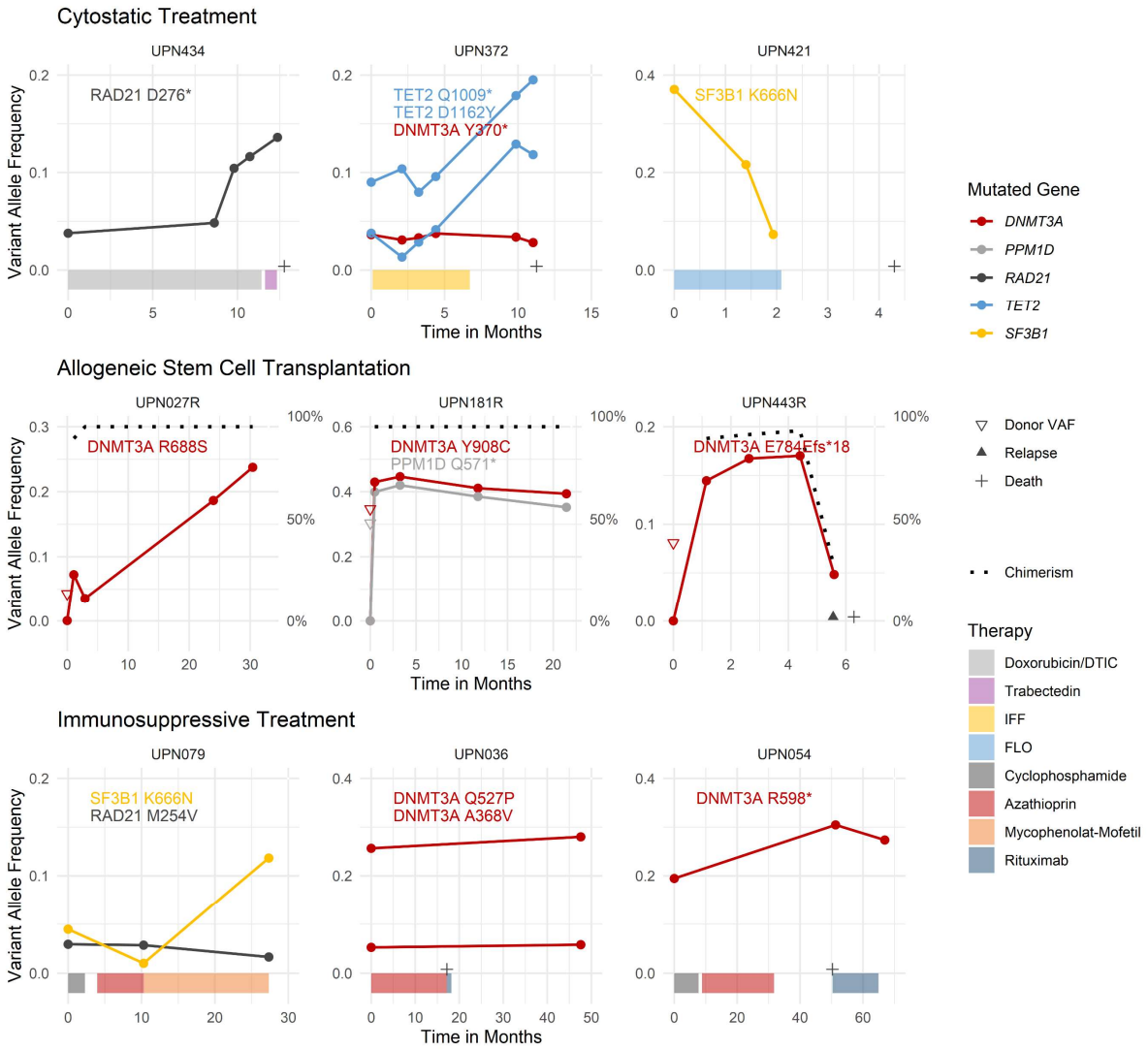

Figure 4. *Clonal evolution in nine exemplary patients*. The time since therapy start/transplantation in months is shown on the x-axis, whereas clone size in terms of VAF is shown on the y-axis. Lines are colored according to mutated gene. Different treatment regimens are depicted as colored bars below the x-axis. Relapse and death are indicated by upright filled triangles and crosses on the x-axis, respectively. (A) Three exemplary patients of the cancer cohort.[16] (B) Three exemplary patients undergoing allogeneic HSCT from a CHIP-positive donor.[17] The VAF measured in the donor hematopoietic system is depicted by an empty reversed triangle at time 0. (C) Three exemplary patients of the AAV cohort.[18] Abbreviations: VAF = variant allele frequency, FLO = Folinic acid / 5-FU / Oxaliplatin, IFF = Folinic acid / 5-FU / Irinotecan, DTIC = Dacarbazin.

References

1. Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C., Wartman, L.D., Lamprecht, T.L., Liu, F., Xia, J., Kandoth, C., Fulton, R.S., McLellan, M.D., Dooling, D.J., Wallis, J.W., Chen, K., Harris, C.C., Schmidt, H.K., Kalicki-Veizer, J.M., Lu, C., Zhang, Q., Lin, L., O'Laughlin, M.D., McMichael, J.F., Delehaunty, K.D., Fulton, L.A., Magrini, V.J., McGrath, S.D., Demeter, R.T., Vickery, T.L., Hundal, J., Cook, L.L., Swift, G.W., Reed, J.P., Alldredge, P.A., Wylie, T.N., Walker, J.R., Watson, M.A., Heath, S.E., Shannon, W.D., Varghese, N., Nagarajan, R., Payton, J.E., Baty, J.D., Kulkarni, S., Klco, J.M., Tomasson, M.H., Westervelt, P., Walter, M.J., Graubert, T.A., DiPersio, J.F., Ding, L., Mardis, E.R. & Wilson, R.K. The origin and evolution of mutations in acute myeloid leukemia. *Cell* **150**, 264-278 (2012).

2. Jan, M., Snyder, T.M., Corces-Zimmerman, M.R., Vyas, P., Weissman, I.L., Quake, S.R. & Majeti, R. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Sci Transl Med* **4**, 149ra118 (2012).

3. Corces-Zimmerman, M.R., Hong, W.J., Weissman, I.L., Medeiros, B.C. & Majeti, R. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. *Proc Natl Acad Sci U S A* **111**, 2548-2553 (2014).

4. Xie, M., Lu, C., Wang, J., McLellan, M.D., Johnson, K.J., Wendl, M.C., McMichael, J.F., Schmidt, H.K., Yellapantula, V., Miller, C.A., Ozenberger, B.A., Welch, J.S., Link, D.C., Walter, M.J., Mardis, E.R., Dipersio, J.F., Chen, F., Wilson, R.K., Ley, T.J. & Ding, L. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat Med* **20**, 1472-1478 (2014).

5. Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P.V., Mar, B.G., Lindsley, R.C., Mermel, C.H., Burtt, N., Chavez, A., Higgins, J.M., Moltchanov, V., Kuo, F.C., Kluk, M.J., Henderson, B., Kinnunen, L., Koistinen, H.A., Ladenvall, C., Getz, G., Correa, A., Banahan, B.F., Gabriel, S., Kathiresan, S., Stringham, H.M., McCarthy, M.I., Boehnke, M., Tuomilehto, J., Haiman, C., Groop, L., Atzmon, G., Wilson, J.G., Neuberg, D., Altshuler, D. & Ebert, B.L. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med* **371**, 2488-2498 (2014).

6. Genovese, G., Kahler, A.K., Handsaker, R.E., Lindberg, J., Rose, S.A., Bakhoum, S.F., Chambert, K., Mick, E., Neale, B.M., Fromer, M., Purcell, S.M., Svantesson, O., Landen, M., Hoglund, M., Lehmann, S., Gabriel, S.B., Moran, J.L., Lander, E.S., Sullivan, P.F., Sklar, P., Gronberg, H., Hultman, C.M. & McCarroll, S.A. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med* **371**, 2477-2487 (2014).

7. Steensma, D.P., Bejar, R., Jaiswal, S., Lindsley, R.C., Sekeres, M.A., Hasserjian, R.P. & Ebert, B.L. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood* **126**, 9-16 (2015).

8. Young, A.L., Challen, G.A., Birmann, B.M. & Druley, T.E. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. *Nat Commun* **7**, 12484 (2016).

9. Jaiswal, S., Natarajan, P., Silver, A.J., Gibson, C.J., Bick, A.G., Shvartz, E., McConkey, M., Gupta, N., Gabriel, S., Ardissino, D., Baber, U., Mehran, R., Fuster, V., Danesh, J., Frossard, P., Saleheen, D., Melander, O., Sukhova, G.K., Neuberg, D., Libby, P., Kathiresan, S. & Ebert, B.L. Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *N Engl J Med* **377**, 111-121 (2017).

10. Fuster, J.J., MacLauchlan, S., Zuriaga, M.A., Polackal, M.N., Ostriker, A.C., Chakraborty, R., Wu, C.L., Sano, S., Muralidharan, S., Rius, C., Vuong, J., Jacob, S.,

Muralidhar, V., Robertson, A.A., Cooper, M.A., Andres, V., Hirschi, K.K., Martin, K.A. & Walsh, K. Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. *Science* **355**, 842-847 (2017).

11. Sano, S., Oshima, K., Wang, Y., MacLauchlan, S., Katanasaka, Y., Sano, M., Zuriaga, M.A., Yoshiyama, M., Goukassian, D., Cooper, M.A., Fuster, J.J. & Walsh, K. Tet2-Mediated Clonal Hematopoiesis Accelerates Heart Failure Through a Mechanism Involving the IL-1beta/NLRP3 Inflammasome. *J Am Coll Cardiol* **71**, 875-886 (2018).

12. Wong, T.N., Ramsingh, G., Young, A.L., Miller, C.A., Touma, W., Welch, J.S., Lamprecht, T.L., Shen, D., Hundal, J., Fulton, R.S., Heath, S., Baty, J.D., Klco, J.M., Ding, L., Mardis, E.R., Westervelt, P., DiPersio, J.F., Walter, M.J., Graubert, T.A., Ley, T.J., Druley, T.E., Link, D.C. & Wilson, R.K. Role of TP53 mutations in the origin and evolution of therapy-related acute myeloid leukaemia. *Nature* **518**, 552-555 (2015).

13. Takahashi, K., Wang, F., Kantarjian, H., Doss, D., Khanna, K., Thompson, E., Zhao, L., Patel, K., Neelapu, S., Gumbs, C., Bueso-Ramos, C., DiNardo, C.D., Colla, S., Ravandi, F., Zhang, J., Huang, X., Wu, X., Samaniego, F., Garcia-Manero, G. & Futreal, P.A. Preleukaemic clonal haemopoiesis and risk of therapy-related myeloid neoplasms: a case-control study. *Lancet Oncol* **18**, 100-111 (2017).

14. Gillis, N.K., Ball, M., Zhang, Q., Ma, Z., Zhao, Y., Yoder, S.J., Balasis, M.E., Mesa, T.E., Sallman, D.A., Lancet, J.E., Komrokji, R.S., List, A.F., McLeod, H.L., Alsina, M., Baz, R., Shain, K.H., Rollison, D.E. & Padron, E. Clonal haemopoiesis and therapy-related myeloid malignancies in elderly patients: a proof-of-concept, case-control study. *Lancet Oncol* **18**, 112-121 (2017).

15. Gibson, C.J., Lindsley, R.C., Tchekmedyian, V., Mar, B.G., Shi, J., Jaiswal, S., Bosworth, A., Francisco, L., He, J., Bansal, A., Morgan, E.A., Lacasce, A.S., Freedman, A.S., Fisher, D.C., Jacobsen, E., Armand, P., Alyea, E.P., Koreth, J., Ho, V., Soiffer, R.J., Antin, J.H., Ritz, J., Nikiforow, S., Forman, S.J., Michor, F., Neuberg, D., Bhatia, R., Bhatia, S. & Ebert, B.L. Clonal Hematopoiesis Associated With Adverse Outcomes After Autologous Stem-Cell Transplantation for Lymphoma. *J Clin Oncol* **35**, 1598-1605 (2017).

16. Arends, C.M., Galan-Sousa, J., Hoyer, K., Chan, W., Jager, M., Yoshida, K., Seemann, R., Noerenberg, D., Waldhueter, N., Fleischer-Notter, H., Christen, F., Schmitt, C.A., Dorken, B., Pelzer, U., Sinn, M., Zemojtel, T., Ogawa, S., Mardian, S., Schreiber, A., Kunitz, A., Kruger, U., Bullinger, L., Mylonas, E., Frick, M. & Damm, F. Hematopoietic lineage distribution and evolutionary dynamics of clonal hematopoiesis. *Leukemia* **32**, 1908-1919 (2018).

17. Frick, M., Chan, W., Arends, C.M., Hablesreiter, R., Halik, A., Heuser, M., Michonneau, D., Blau, O., Hoyer, K., Christen, F., Galan-Sousa, J., Noerenberg, D., Wais, V., Stadler, M., Yoshida, K., Schetelig, J., Schuler, E., Thol, F., Clappier, E., Christopeit, M., Ayuk, F., Bornhauser, M., Blau, I.W., Ogawa, S., Zemojtel, T., Gerbitz, A., Wagner, E.M., Spriewald, B.M., Schrezenmeier, H., Kuchenbauer, F., Kobbe, G., Wiesneth, M., Koldehoff, M., Socie, G., Kroeger, N., Bullinger, L., Thiede, C. & Damm, F. Role of Donor Clonal Hematopoiesis in Allogeneic Hematopoietic Stem-Cell Transplantation. *J Clin Oncol* **37**, 375-385 (2019).

18. Arends, C.M., Weiss, M., Christen, F., Eulenberg-Gustavus, C., Rousselle, A., Kettritz, R., Eckardt, K.U., Chan, W., Hoyer, K., Frick, M., Bullinger, L., Bieringer, M., Schreiber, A. & Damm, F. Clonal hematopoiesis in patients with ANCA-associated vasculitis. *Haematologica* (2019).

19. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-595 (2010).

20. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. & DePristo, M.A. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303 (2010).

21. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & Genome Project Data Processing, S. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).

22. Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L. & Wilson, R.K. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* **22**, 568-576 (2012).

23. Jager, M., Wang, K., Bauer, S., Smedley, D., Krawitz, P. & Robinson, P.N. Jannovar: a java library for exome annotation. *Hum Mutat* **35**, 548-555 (2014).

24. Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. & Rozen, S.G. Primer3--new capabilities and interfaces. *Nucleic Acids Res* **40**, e115 (2012).

25. Koressaar, T. & Remm, M. Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23**, 1289-1291 (2007).

26. Korbie, D.J. & Mattick, J.S. Touchdown PCR for increased specificity and sensitivity in PCR amplification. *Nat Protoc* **3**, 1452-1456 (2008).

27. Livak, K.J. & Schmittgen, T.D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408 (2001).

28. Schreiber, A., Rolle, S., Peripelittchenko, L., Rademann, J., Schneider, W., Luft, F.C. & Kettritz, R. Phosphoinositol 3-kinase-gamma mediates antineutrophil cytoplasmic autoantibody-induced glomerulonephritis. *Kidney Int* **77**, 118-128 (2010).

29. Gray, R.J. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat*, 1141-1154 (1988).

30. Jason P. Fine, R.J.G. A proportional hazards model for the subdistribution of a competing risk. *J Am Stat Assoc*, 496-509 (1999).

31. Abelson, S., Collord, G., Ng, S.W.K., Weissbrod, O., Mendelson Cohen, N., Niemeyer, E., Barda, N., Zuzarte, P.C., Heisler, L., Sundaravadanam, Y., Luben, R., Hayat, S., Wang, T.T., Zhao, Z., Cirlan, I., Pugh, T.J., Soave, D., Ng, K., Latimer, C., Hardy, C., Raine, K., Jones, D., Hoult, D., Britten, A., McPherson, J.D., Johansson, M., Mbabaali, F., Eagles, J., Miller, J.K., Pasternack, D., Timms, L., Krzyzanowski, P., Awadalla, P., Costa, R., Segal, E., Bratman, S.V., Beer, P., Behjati, S., Martincorena, I., Wang, J.C.Y., Bowles, K.M., Quiros, J.R., Karakatsani, A., La Vecchia, C., Trichopoulou, A., Salamanca-Fernandez, E., Huerta, J.M., Barricarte, A., Travis, R.C., Tumino, R., Masala, G., Boeing, H., Panico, S., Kaaks, R., Kramer, A., Sieri, S., Riboli, E., Vineis, P., Foll, M., McKay, J., Polidoro, S., Sala, N., Khaw, K.T., Vermeulen, R., Campbell, P.J., Papaemmanuil, E., Minden, M.D., Tanay, A., Balicer, R.D., Wareham, N.J., Gerstung, M., Dick, J.E., Brennan, P., Vassiliou, G.S. & Shlush, L.I. Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature* **559**, 400-404 (2018).

32. Buscarlet, M., Provost, S., Zada, Y.F., Barhdadi, A., Bourgoin, V., Lepine, G., Mollica, L., Szuber, N., Dube, M.P. & Busque, L. DNMT3A and TET2 dominate clonal

hematopoiesis and demonstrate benign phenotypes and different genetic predispositions. *Blood* **130**, 753-762 (2017).

33. Desai, P., Mencia-Trinchant, N., Savenkov, O., Simon, M.S., Cheang, G., Lee, S., Samuel, M., Ritchie, E.K., Guzman, M.L., Ballman, K.V., Roboz, G.J. & Hassane, D.C. Somatic mutations precede acute myeloid leukemia years before diagnosis. *Nat Med* **24**, 1015-1023 (2018).

34. Pang, W.W., Price, E.A., Sahoo, D., Beerman, I., Maloney, W.J., Rossi, D.J., Schrier, S.L. & Weissman, I.L. Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proc Natl Acad Sci U S A* **108**, 20012-20017 (2011).

35. Buscarlet, M., Provost, S., Zada, Y.F., Bourgoin, V., Mollica, L., Dube, M.P. & Busque, L. Lineage restriction analyses in CHIP indicate myeloid bias for TET2 and multipotent stem cell origin for DNMT3A. *Blood* **132**, 277-280 (2018).

36. Bowen, D., Groves, M.J., Burnett, A.K., Patel, Y., Allen, C., Green, C., Gale, R.E., Hills, R. & Linch, D.C. TP53 gene mutation is frequent in patients with acute myeloid leukemia and complex karyotype, and is associated with very poor prognosis. *Leukemia* **23**, 203-206 (2009).

37. Christiansen, D.H., Andersen, M.K. & Pedersen-Bjergaard, J. Mutations with loss of heterozygosity of p53 are common in therapy-related myelodysplasia and acute myeloid leukemia after exposure to alkylating agents and significantly associated with deletion or loss of 5q, a complex karyotype, and a poor prognosis. *J Clin Oncol* **19**, 1405-1413 (2001).

38. Haferlach, C., Dicker, F., Herholz, H., Schnittger, S., Kern, W. & Haferlach, T. Mutations of the TP53 gene in acute myeloid leukemia are strongly associated with a complex aberrant karyotype. *Leukemia* **22**, 1539-1541 (2008).

39. Kayser, S., Dohner, K., Krauter, J., Kohne, C.H., Horst, H.A., Held, G., von Lilienfeld-Toal, M., Wilhelm, S., Kundgen, A., Gotze, K., Rummel, M., Nachbaur, D., Schlegelberger, B., Gohring, G., Spath, D., Morlok, C., Zucknick, M., Ganser, A., Dohner, H., Schlenk, R.F. & German-Austrian, A. The impact of therapy-related acute myeloid leukemia (AML) on outcome in 2853 adult patients with newly diagnosed AML. *Blood* **117**, 2137-2145 (2011).

40. Rucker, F.G., Schlenk, R.F., Bullinger, L., Kayser, S., Teleanu, V., Kett, H., Habdank, M., Kugler, C.M., Holzmann, K., Gaidzik, V.I., Paschka, P., Held, G., von Lilienfeld-Toal, M., Lubbert, M., Frohling, S., Zenz, T., Krauter, J., Schlegelberger, B., Ganser, A., Lichter, P., Dohner, K. & Dohner, H. TP53 alterations in acute myeloid leukemia with complex karyotype correlate with specific copy number alterations, monosomal karyotype, and dismal outcome. *Blood* **119**, 2114-2121 (2012).

41. Murphy, T., Zou, J., Daher-Reyes, G.S., Arruda, A., Gupta, V., McNamara, C.J., Minden, M.D., Schimmer, A.D., Sibai, H., Yee, K.W.L., Korulla, M., Stockley, T., Kamel-Reid, S., Maze, D., Tierens, A., Bratman, S.V., Schuh, A.C. & Chan, S.M. Impact of preleukemic mutations and their persistence on hematologic recovery after induction chemotherapy for AML. *Blood Adv* **3**, 2307-2311 (2019).

42. Savola, P., Lundgren, S., Keranen, M.A.I., Almusa, H., Ellonen, P., Leirisalo-Repo, M., Kelkka, T. & Mustjoki, S. Clonal hematopoiesis in patients with rheumatoid arthritis. *Blood Cancer J* **8**, 69 (2018).

43. Castro, M., Conn, D.L., Su, W.P. & Garton, J.P. Rheumatic manifestations in myelodysplastic syndromes. *J Rheumatol* **18**, 721-727 (1991).

44. Kuzmich, P.V., Ecker, G.A. & Karsh, J. Rheumatic manifestations in patients with myelodysplastic and myeloproliferative diseases. *J Rheumatol* **21**, 1649-1654 (1994).

45. Komrokji, R.S., Kulasekararaj, A., Al Ali, N.H., Kordasti, S., Bart-Smith, E., Craig, B.M., Padron, E., Zhang, L., Lancet, J.E., Pinilla-Ibarz, J., List, A.F., Mufti, G.J. & Epling-Burnette, P.K. Autoimmune diseases and myelodysplastic syndromes. *Am J Hematol* **91**, E280-283 (2016).

46. Wilson, A.B., Neogi, T., Prout, M. & Jick, S. Relative risk of myelodysplastic syndromes in patients with autoimmune disorders in the General Practice Research Database. *Cancer Epidemiol* **38**, 544-549 (2014).

47. Anderson, L.A., Pfeiffer, R.M., Landgren, O., Gadalla, S., Berndt, S.I. & Engels, E.A. Risks of myeloid malignancies in patients with autoimmune conditions. *Br J Cancer* **100**, 822-828 (2009).

48. Cargo, C.A., Rowbotham, N., Evans, P.A., Barrans, S.L., Bowen, D.T., Crouch, S. & Jack, A.S. Targeted sequencing identifies patients with pre-clinical MDS at high risk of disease progression. *Blood* (2015).

49. Kwok, B., Hall, J.M., Witte, J.S., Xu, Y., Reddy, P., Lin, K., Flamholz, R., Dabbas, B., Yung, A., Al-Hafidh, J., Balmert, E., Vaupel, C., El Hader, C., McGinniss, M.J., Nahas, S.A., Kines, J. & Bejar, R. MDS-associated somatic mutations and clonal hematopoiesis are common in idiopathic cytopenias of undetermined significance. *Blood* **126**, 2355-2361 (2015).

50. Schreiber, A., Luft, F.C. & Kettritz, R. Phagocyte NADPH oxidase restrains the inflammasome in ANCA-induced GN. *J Am Soc Nephrol* **26**, 411-424 (2015).

51. Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., McIndoo, J., Cook, K., Stepansky, A., Levy, D., Esposito, D., Muthuswamy, L., Krasnitz, A., McCombie, W.R., Hicks, J. & Wigler, M. Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90-94 (2011).

52. Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res* **27**, 491-499 (2017).

53. Wong, W.H., Tong, R.S., Young, A.L. & Druley, T.E. Rare Event Detection Using Error-corrected DNA and RNA Sequencing. *J Vis Exp* (2018).

54. Young, A.L., Wong, T.N., Hughes, A.E., Heath, S.E., Ley, T.J., Link, D.C. & Druley, T.E. Quantifying ultra-rare pre-leukemic clones via targeted error-corrected sequencing. *Leukemia* **29**, 1608-1611 (2015).

Eidesstattliche Versicherung

„Ich, Christopher Maximilian Arends, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema „New Insights into the Dynamics of Clonal Hematopoiesis from Targeted Sequencing Studies“ selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.

Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren beruhen, sind als solche in korrekter Zitierung (siehe „Uniform Requirements for Manuscripts (URM)“ des ICMJE -*www.**icmje**.org*) kenntlich gemacht.

Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) entsprechen den URM (s.o.) und werden von mir verantwortet.

Meine Anteile an den ausgewählten Publikationen entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Betreuer/in, angegeben sind.

Sämtliche Publikationen, die aus dieser Dissertation hervorgegangen sind und bei denen ich Autor bin, entsprechen den URM (s.o.) und werden von mir verantwortet.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§156,161 des Strafgesetzbuches) sind mir bekannt und bewusst.“

_____

Datum Unterschrift

Anteilserklärung an den erfolgten Publikationen

Christopher Maximilian Arends hatte folgenden Anteil an den folgenden Publikationen:

Publikation 1: **Arends CM***, Galan-Sousa J*, Hoyer K*, Chan W, Jäger M, Yoshida K, Seemann R, Noerenberg D, Waldhueter N, Fleischer-Notter H, Christen F, Schmitt CA, Dörken B, Pelzer U, Sinn M, Zemojtel T, Ogawa S, Märdian S, Schreiber A, Kunitz A, Krüger U, Bullinger L, Mylonas E, Frick M, Damm F, *Hematopoietic lineage distribution and evolutionary dynamics of clonal hematopoiesis*, Leukemia, 2018

* Diese Autoren haben in gleichen Teilen zur Arbeit als Erstautoren beigetragen.

Die vorliegende Arbeit entstand im international hochkompetitiven Forschungsfeld der Hämatoonkologie und wurde hochrangig publiziert. Die Arbeit verlangte die Rekrutierung zweier umfangreicher Patientenkohorten, die Sammlung zahlreicher Proben sowie die Etablierung, Durchführung und Auswertung vielfältiger und komplexer Experimente. Unter der Betreuung der Letztautoren hat jeder Erstautor im gleichen Umfang, aber mit verschiedenen Schwerpunkten an dieser Arbeit mitgewirkt und somit ganz entscheidend zur Gewinnung neuer wissenschaftlicher Erkenntnisse beigetragen. Die nachfolgende Anteilserklärung stellt allein die Schwerpunkte des Erstautors Christopher Maximilian Arends dar. Sie enthält keine Aussagen zu den spezifischen Beiträgen der anderen (Erst- und Letzt-)Autoren.

Beitrag im Einzelnen:
Die Schwerpunkte meiner Arbeit liegen vor allem in der bioinformatischen und statistischen Auswertung der Targeted Sequencing Daten, sowie in der Akquirierung der Proben und Daten der Tumorpatientenkohorte und deren Datenauswertung.
Beitrag im Einzelnen:

- Patientenrekrutierung, Proben- und Datenakquirierung der Tumorpatientenkohorte (n = 72)
    - Akquirierung und Verarbeitung der Verlaufsproben (n ≈ 220)
    - Verarbeitung der Patientenproben inklusive Ficoll-Dichtezentrifugation
    - DNA Extraktion
    - Erfassung und Verwaltung der klinischen Daten der Tumorpatientenkohorte
- Erstellung von Sequencing-Libraries, sowie deren Sequenzierung auf dem MiSeq für die Tumorkohorte (n = 72) und einen Teil der Kohorte ohne maligne Erkrankungen (n ≈ 100)
- Literaturrecherche bezüglich der häufig mutierten Gene in CHIP und Design des Custom Sequencing Panels
- Bioinformatische Auswertung der Targeted Sequencing Daten
    - Mitarbeit bei der Erstellung und Optimierung einer bioinformatischen Auswertungspipeline (in Zusammenarbeit mit Marten Jäger)
    - Eigenständige Programmierung eines bioinformatischen Tools zum Filtern der relevanten genetischen Varianten in R
    - Auswertung der Sequenzierdaten aller Patienten (n = 437) im Screeningprozess
- Etablierung des bioinformatischen Rahmens (automatisiertes Pooling, automatisierte Primerselektion und automatisierte Auswertung) für die Validierung der Mutationskandidaten mittels Targeted Deep Sequencing
- Validierung der Mutationskandidaten der Tumorpatientenkohorte sowie die Quantifizierung der follow-up Proben (Primerdesign, PCR, Pooling, bioinformatische Auswertung, n = 175). Diese Ergebnisse sind in Figure 5 dargestellt.
- Deskriptive statistische Auswertung der Mutationsdaten für beide Kohorten, u. a. Komutationsanalyse mit R und CIRCOS. Diese Daten sind in Figure 1 dargestellt.

- Statistische Auswertung der klinischen Daten der Tumorpatientenkohorte mit SPSS. Diese Daten sind in Supplemental Tables S7, S8, S9, S10 dargestellt.
- Mitarbeit bei der Erstellung des Manuskripts mit besonderem Schwerpunkt auf den die Tumorpatientenkohorte betreffenden Teil.
- Erstellung von Abbildungen mit R (auf Grundlage der gemeinsam mit anderen Coautoren gewonnenen Daten): Figure 1A, B, D, E, Figure 4 und Figure 5, Supplemental Figures S4, S7, S8, S9, S10, S11, S12, S15)

Publikation 2: Mareike Frick*, Willy Chan*, **Christopher Maximilian Arends**, Raphael Hablesreiter, Adriane Halik, Michael Heuser, David Michonneau, Olga Blau, Kaja Hoyer, Friederike Christen, Joel Galan-Sousa, Daniel Noerenberg, Verena Wais, Michael Stadler, Kenichi Yoshida, Johannes Schetelig, Esther Schuler, Felicitas Thol, Emmanuelle Clappier, Maximilian Christopeit, Francis Ayuk, Martin Bornhäuser, Igor Wolfgang Blau, Seishi Ogawa, Tomasz Zemojtel, Armin Gerbitz, Eva M. Wagner, Bernd M. Spriewald, Hubert Schrezenmeier, Florian Kuchenbauer, Guido Kobbe, Markus Wiesneth, Michael Koldehoff, Gérard Socié, Nicolaus Kroeger, Lars Bullinger, Christian Thiede, Frederik Damm, *Role of Donor Clonal Hematopoiesis in Allogeneic Hematopoietic Stem-Cell Transplantation*, Journal of Clinical Oncology, 2019

* Diese Autoren haben in gleichen Teilen zur Arbeit als Erstautoren beigetragen.

Beitrag im Einzelnen:
- Mitarbeit bei der Library Präparation und Sequenzierung für einen Teil der DNA Proben
- Mitarbeit bei der bioinformatischen Auswertung der Sequenzierdaten von Screening und Validierung
- Erstellung der Abbildungen Figure 5, A2 und A11 für das Manuskript mit R

Publikation 3: **Christopher Maximilian Arends***, Marlene Weiss*, Friederike Christen, Claudia Eulenberg-Gustavus, Anthony Rousselle, Ralph Kettritz, Kai-Uwe Eckardt, Willy Chan, Kaja Hoyer, Mareike Frick, Lars Bullinger, Markus Bieringer, Adrian Schreiber, and Frederik Damm, *Clonal Hematopoiesis in Patients with ANCA-associated Vasculitis*, Haematologica, 2019

* Diese Autoren haben in gleichen Teilen zur Arbeit als Erstautoren beigetragen.

Beitrag im Einzelnen:
- Supervision der bioinformatischen Auswertung der Sequenzierdaten.
- Sequenzierung der 40 follow-up samples.
- Statistische Auswertung der Sequencing Daten und der klinischen Daten mit R.
- Verfassung des ersten Entwurfs des Manuskripts und Mitarbeit bei dessen Überarbeitung.
- Erstellung aller Abbildungen und des gesamten Supplements.
- Mitarbeit bei den Revisions: Sequenzierung der zusätzlichen Individuen für die Kontrollkohorte und 20 weiterer follow-up samples sowie entsprechende Überarbeitung des Manuskripts, der Abbildungen und des Supplements.

_____
Unterschrift des Doktoranden

_____
Unterschrift, Datum und Stempel des betreuenden Hochschullehrers/der betreuenden Hochschullehrerin

Publication 1

Arends, C.M., Galan-Sousa, J., Hoyer, K., Chan, W., Jager, M., Yoshida, K., Seemann, R., Noerenberg, D., Waldhueter, N., Fleischer-Notter, H., Christen, F., Schmitt, C.A., Dorken, B., Pelzer, U., Sinn, M., Zemojtel, T., Ogawa, S., Mardian, S., Schreiber, A., Kunitz, A., Kruger, U., Bullinger, L., Mylonas, E., Frick, M. & Damm, F. *Hematopoietic lineage distribution and evolutionary dynamics of clonal hematopoiesis.* Leukemia 32, 1908-1919 (2018).

https://doi.org/10.1038/s41375-018-0047-7

Publication 2

Frick, M., Chan, W., Arends, C.M., Hablesreiter, R., Halik, A., Heuser, M., Michonneau, D., Blau, O., Hoyer, K., Christen, F., Galan-Sousa, J., Noerenberg, D., Wais, V., Stadler, M., Yoshida, K., Schetelig, J., Schuler, E., Thol, F., Clappier, E., Christopeit, M., Ayuk, F., Bornhauser, M., Blau, I.W., Ogawa, S., Zemojtel, T., Gerbitz, A., Wagner, E.M., Spriewald, B.M., Schrezenmeier, H., Kuchenbauer, F., Kobbe, G., Wiesneth, M., Koldehoff, M., Socie, G., Kroeger, N., Bullinger, L., Thiede, C. & Damm, F. *Role of Donor Clonal Hematopoiesis in Allogeneic Hematopoietic Stem-Cell Transplantation.* J Clin Oncol 37, 375-385 (2019).

https://doi.org/10.1200/JCO.2018.79.2184

Publication 3

Arends, C.M., Weiss, M., Christen, F., Eulenberg-Gustavus, C., Rousselle, A., Kettritz, R., Eckardt, K.U., Chan, W., Hoyer, K., Frick, M., Bullinger, L., Bieringer, M., Schreiber, A. & Damm, F. *Clonal hematopoiesis in patients with ANCA-associated vasculitis.* Haematologica (2019).

https://doi.org/10.3324/haematol.2019.223305

Curriculum Vitae

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

List of Publications

1. Arends, C.M., Weiss, M., Christen, F., Eulenberg-Gustavus, C., Rousselle, A., Kettritz, R., Eckardt, K.U., Chan, W., Hoyer, K., Frick, M., Bullinger, L., Bieringer, M., Schreiber, A. & Damm, F. *Clonal hematopoiesis in patients with ANCA-associated vasculitis.* Haematologica (2019). IF = 9.090

2. Frick, M., Chan, W., Arends, C.M., Hablesreiter, R., Halik, A., Heuser, M., Michonneau, D., Blau, O., Hoyer, K., Christen, F., Galan-Sousa, J., Noerenberg, D., Wais, V., Stadler, M., Yoshida, K., Schetelig, J., Schuler, E., Thol, F., Clappier, E., Christopeit, M., Ayuk, F., Bornhauser, M., Blau, I.W., Ogawa, S., Zemojtel, T., Gerbitz, A., Wagner, E.M., Spriewald, B.M., Schrezenmeier, H., Kuchenbauer, F., Kobbe, G., Wiesneth, M., Koldehoff, M., Socie, G., Kroeger, N., Bullinger, L., Thiede, C. & Damm, F. *Role of Donor Clonal Hematopoiesis in Allogeneic Hematopoietic Stem-Cell Transplantation.* J Clin Oncol 37, 375-385 (2019). IF = 26.303

3. Arends, C.M., Galan-Sousa, J., Hoyer, K., Chan, W., Jager, M., Yoshida, K., Seemann, R., Noerenberg, D., Waldhueter, N., Fleischer-Notter, H., Christen, F., Schmitt, C.A., Dorken, B., Pelzer, U., Sinn, M., Zemojtel, T., Ogawa, S., Mardian, S., Schreiber, A., Kunitz, A., Kruger, U., Bullinger, L., Mylonas, E., Frick, M. & Damm, F. *Hematopoietic lineage distribution and evolutionary dynamics of clonal hematopoiesis.* Leukemia 32, 1908-1919 (2018). IF = 10.023

4. Mansouri, L., Noerenberg, D., Young, E., Mylonas, E., Abdulla, M., Frick, M., Asmar, F., Ljungstrom, V., Schneider, M., Yoshida, K., Skaftason, A., Pandzic, T., Gonzalez, B., Tasidou, A., Waldhueter, N., Rivas-Delgado, A., Angelopoulou, M., Ziepert, M., Arends, C.M., Couronne, L., Lenze, D., Baldus, C.D., Bastard, C., Okosun, J., Fitzgibbon, J., Dorken, B., Drexler, H.G., Roos-Weil, D., Schmitt, C.A., Munch-Petersen, H.D., Zenz, T., Hansmann, M.L., Strefford, J.C., Enblad, G., Bernard, O.A., Ralfkiaer, E., Erlanson, M., Korkolopoulou, P., Hultdin, M., Papadaki, T., Gronbaek, K., Lopez-Guillermo, A., Ogawa, S., Kuppers, R., Stamatopoulos, K., Stavroyianni, N., Kanellis, G., Rosenwald, A., Campo, E., Amini, R.M., Ott, G., Vassilakopoulos, T.P., Hummel, M., Rosenquist, R. & Damm, F. *Frequent NFKBIE deletions are associated with poor outcome in primary mediastinal B-cell lymphoma.* Blood 128, 2666-2670 (2016). IF = 15.132

5. Arends, C.M., Hebecker, A., Heimpel, K., Kraus, S.C., Lust, D., Mayrhofer, C., Schick, C. & Weigand, T. *D7-Brane Moduli Space in Axion Monodromy and Fluxbrane Inflation.* Fortschr. Phys. 62, No. 8, 647-702 (2014). IF = 2.705

C. M. Arends

Berlin, den 31. Oktober 2020

Acknowledgements

First and most of all, I would like to express my sincere gratitude to my supervisors Frederik Damm and Mareike Frick for giving me the opportunity to carry out these projects, for their assistance and for their guidance throughout and beyond this work.
I would like to thank everyone who has contributed to this work, in particular my co-doctoral students Joel Galan-Sousa, Kaja Hoyer, Friederike Christen and Willy Chan for their contributions; Helga Fleischer-Notter and Antje Maluck for technical assistance; Marten Jäger and Ulrike Krüger for sharing their expertise in sequencing and bioinformatics; Marianne Sinn, Jana Striefler, Uwe Pelzer, Annegret Kunitz, Daniel Noerenberg and Nils Waldhüter for their help in recruiting patients; Axel Pruß for giving me access to transfusion data.
I am grateful to the Berlin Institute of Health for granting me a one-year doctoral scholarship as well as to Team Deutschlandstipendium and id Berlin for granting me the „Deutschlandstipendium“.
Finally, I am deeply grateful to Yona, my friends and my parents for their continuous support.