

12-2020

Subclonal Evolution Of Chronic Lymphocytic Leukemia After Allogeneic T Cell Therapies

Haven Garber

Follow this and additional works at: https://digitalcommons.library.tmc.edu/utgsbs_dissertations



Part of the [Cancer Biology Commons](#), [Immunotherapy Commons](#), and the [Medicine and Health Sciences Commons](#)

Recommended Citation

Garber, Haven, "Subclonal Evolution Of Chronic Lymphocytic Leukemia After Allogeneic T Cell Therapies" (2020). *Dissertations and Theses (Open Access)*. 1044.

https://digitalcommons.library.tmc.edu/utgsbs_dissertations/1044

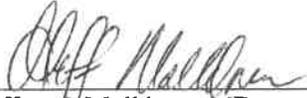
This Dissertation (PhD) is brought to you for free and open access by the MD Anderson UTHealth Houston Graduate School at DigitalCommons@TMC. It has been accepted for inclusion in Dissertations and Theses (Open Access) by an authorized administrator of DigitalCommons@TMC. For more information, please contact digcommons@library.tmc.edu.

SUBCLONAL EVOLUTION OF CHRONIC LYMPHOCYTIC LEUKEMIA AFTER
ALLOGENEIC T CELL THERAPIES

By

Haven Rebecca Garber, MD

APPROVED:



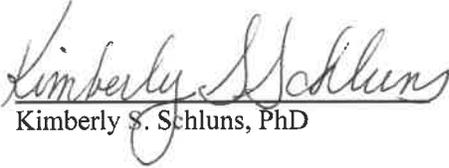
Jeffrey H. Molldrem, MD
Advisory Professor



R. Eric Davis, MD



P. Andrew Futreal, PhD



Kimberly S. Schluns, PhD



William G. Wierda, MD, PhD

APPROVED:

Dean, The University of Texas

MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences

SUBCLONAL EVOLUTION OF CHRONIC LYMPHOCYTIC LEUKEMIA AFTER
ALLOGENEIC T CELL THERAPIES

A

DISSERTATION

Presented to the Faculty of

The University of Texas

MD Anderson Cancer Center UTHealth

Graduate School of Biomedical Sciences

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Haven Rebecca Garber, MD

Houston, Texas

December 2020

Dedication

To my children - Scout, Kai, and James.

It is my life's greatest blessing to help you thrive and grow. Keep exploring! – both inside and out! But humble yourself, the truth is hard to know.

Acknowledgements

I am grateful to the many people who have helped me along the way - in clinical and research training. Fortunately, most of them will be spared the experience of holding this document in their hands, but I will acknowledge some of them here. It's rare to have a moment to reflect, and this reminds me to thank them in person.

First thank you to my mentor, Dr. Jeff Mouldrem, for taking me on as a postdoc and putting up with me and my frequent skepticism. I am lucky to know you. I have profound respect and admiration for you as a doctor, scientist, boss, and person and I look forward to continuing our collaboration.

Next, to Dr. Crystal Mackall, my first immunology mentor. It was sheer serendipity that I landed in your lab at the NCI as a wide-eyed, naive med student. You provided the spark as you have for so many.

To the many thoughtful and careful scientists who helped me with this project - Hannah, Eric, Sahil, Celine, Karen Clise-Dwyer, Jin Im, Lily, Jason R., Jason C., Shoudan, Anne, Priya, Ethan – some of the finest scientists around.

Thank you to my cheerleading section (because everyone needs one!), many who double as my mentors – Drs. Amanda Hermann, Jenn Litton, Stacy Moulder, Haley Peters, Gheath Alatrash, Mariana Chavez MacGregor, Kim Schluns, Patrick Pilie, and Jon Mizrahi.

To the Mouldrem/Futreal/Alatrash/Mittendorf labs - Thank you Lisa, Annalea, Kathryn, Pariya, Helen, Anna, Sijie, Lily, Qing, Alex, Na, Mr. Mao, Tian-Hui, Joy, Shannon, the flow core, and the SMF staff.

To my GSBS committee, thank you for your insight and encouragement.

To the GSBS administration, thank you for taking a chance on me, a nontraditional student. I enrolled to learn about the process of discovery and to gain confidence in my own scientific voice. I think I achieved both goals. I now have a collection of favorite cancer biology/immunology papers, am rarely intimidated by a basic science paper, and understand that most of the interesting data are in the supplement!

To Dr. Tripathy, thank you for giving me some space to wrap up this project.

To Gheath, I'm so lucky to have you as a mentor. Thank you for everything.

To Dr. Nishin Bhadkamkar and Dr. Michael Fisch et al. Thank you for teaching me what it means to be an oncologist.

To my best friend Roxanne, thank you for your friendship, I am lucky every day to have you by my side (even when you're in Nome, AK - move closer!)

To my partner Jen, Thank you for your support these last 13 years. What a joy to raise our awesome children together. In the professional realm, I am astonished by your

endless capacity for team science and your tenacity. You have earned (twice over) every accolade and success— more to come no doubt.

To my kids - If you someday find yourself reading this (!?) please know that I love you beyond measure and that I am so proud of the people you are becoming. Please use your talents to make this world a better place. Make lots of friends, too. From my favorite surgeon-humanitarian in Boston – Dr. AK Goodman (who may have been quoting Mother Teresa but is as good a source) “We can do no great things, only small things with great love.”

To my parents – Thank you Dad (Keith) - a manufacturing engineer, a brilliant man (no hyperbole), and the hardest worker I know. He adroitly navigated an adverse job market and an industry that was shifting abroad to provide us with the stability that he never had. Thank you Mom (Janice) – a sensitive soul who forwent her own teaching career to raise and teach us. Some of my earliest memories are of us taking trips to the Greensboro Public Library to pick out a book, any book that interested us. It is no great mystery how we all landed in academic careers. I am proud to be your daughter. Thanks also to my younger sister, Chelsea, and older brother, Grant, for being awesome and providing endless banter. I admire you both.

Last, thank you to the patients and their families who trust us with their cancer care and cancer research. I've sat in countless exam rooms where a patient with terminal cancer looks me in the eye and tells me they want to enroll onto a protocol or clinical trial not to help themselves but to help whoever comes after them. It's a level of bravery I can't comprehend but I'm so grateful.

Abstract

Subclonal evolution of chronic lymphocytic leukemia after allogeneic T-cell therapies

Haven Garber, MD

Advisory Professor: Jeffrey Molldrem, MD

Intratumoral genetic heterogeneity describes the molecular differences among subclones within a tumor and is a major barrier to effective therapy in many solid and liquid cancers, including chronic lymphocytic leukemia (CLL). Rare, treatment-resistant subclones can expand to compose relapsed disease during tumor evolution. Examination of malignant evolution in the context of specific treatment provides insight into the molecular lesions that mediate therapeutic response and resistance. Both chemotherapy and targeted therapy were shown to precipitate CLL subclonal evolution. We hypothesized that allogeneic T-cell immunotherapies, including allogeneic stem cell transplant (alloSCT) and donor lymphocyte infusion (DLI), would impact malignant evolution through the application of selective immunologic pressure imposed by donor T cells. Here, we tested this prediction in a cohort of 24 CLL patients treated with nonmyeloablative HLA-matched alloSCT and DLI utilizing whole exome sequencing of purified CLL. Our cohort included 11 patients who relapsed after alloSCT, and we studied sequential samples in these patients to examine leukemic evolution. We identified clear patterns of linear and branched evolution in 8/11 patients after alloSCT/DLI that included CLL-specific drivers in every case. In two patients, leukemic evolution was coincident with DLI, suggesting immunoediting of leukemic subclones. To investigate complementary changes in immunity, we analyzed the post-alloSCT T cell repertoires of CLL transplant recipients at multiple time points after engraftment and observed restricted diversity. Last, we adapted and employed a strategy to identify and track candidate graft-versus-leukemia T cells that expanded and contracted coincident with loss of specific tumorigenic lesions. We provide novel evidence of ongoing genetic subclonal evolution of CLL post-alloSCT.

Table of Contents

Dedication.....	3
Acknowledgements.....	4
Abstract.....	6
List of Abbreviations.....	9
Chapter 1: Introduction.....	11
Intratumoral heterogeneity and clonal evolution in cancer	11
Chronic lymphocytic leukemia.....	12
Allogeneic stem cell transplant and Donor lymphocyte infusion.....	14
Hypothesis and Specific Aims	16
Chapter 2: Genomic evolution of CLL after alloSCT	18
CLL subclonal evolution occurs after DLI.....	26
CLL drivers evolve after alloSCT	31
Chapter 3: Identification of a new candidate CLL driver through study of leukemic evolution.....	33
Candidate driver <i>CHEK2</i> c.C275T (p.P92L) affects Chk2 dimerization	35
Chapter 4: Mutation load and neoantigen burden do not predict CLL response to alloSCT.....	39
CLL alloSCT patients.....	40
Structural heterogeneity differs between CLL alloSCT responders and nonresponders.....	45
Chapter 5: Post-alloSCT T cell repertoires	48
Chapter 6: Single cell TCR sequencing of candidate GVL clones	62
Single T cell TCR sequencing protocol.....	64
Single cell analysis of candidate GVL T cells	74
Chapter 7: Discussion and Future Directions	84
Bibliography	104
Vita.....	138

List of Figures

Figure 1. Sort purification of CLL and T cells.....	22
Figure 2. WES of post-alloSCT CLL is feasible despite donor chimerism	26
Figure 3. Branched evolution of CLL subclones occurs after DLI in patients CLL5 and CLL8....	28
Figure 4. AF scatter plots for CLL patients 1-11 showing somatic variants across 2 time points for each patient's course	29
Figure 5. Evolution of CLL drivers and enrichment for novel mutations occurs after alloSCT and DLI	32
Figure 6. Candidate driver <i>CHEK2</i> c.C275T (p.92L) affects Chk2 dimerization.	36
Figure 7. Chk2 P92L mutants form tighter homodimers	38
Figure 8. Somatic variants detected in the alloSCT patient cohort (n=24).....	43
Figure 9. Mutation load and neoantigen burden do not predict response to alloSCT	43
Figure 10. Structural heterogeneity of pre-transplant CLL differs between alloSCT response groups	46
Figure 11. CDR3 length plot for patient CLL 14	52
Figure 12. Top 20 T cell clonotypes	58
Figure 13. TCR repertoire analysis and reproducibility	59
Figure 14. T cell repertoire diversity of post-alloSCT samples	60
Figure 15. AlloSCT donor and host T cell repertoires	61
Figure 16. Example layout of TCR α and β barcoding primers in preparation for NGS of single cell TCR transcripts.....	70
Figure 17. Example gel of the final TCR amplicon at 350 - 380 bp after PCR reaction #3	71
Figure 18. MiSeq output.....	71
Figure 19. Single leukemia cell expression of neoantigens	78
Figure 20. Neoantigen-specific T cells in CLL patient 12	80
Figure 21. Integrating neoantigen tetramer-binding TCR β sequences with longitudinal, bulk repertoire TCR β sequences for CLL patient 12	80
Figure 22. Neoantigen-specific T cells in CLL patient 8.....	83

List of Tables

Table 1. CLL prognostic features and prior treatments for the alloSCT cohort.....	23
Table 2. Patient and donor alloSCT characteristics.....	24
Table 3. Post-alloSCT T cell samples and TCR V β repertoire diversity.....	50
Table 4. Primers used for single T cell TCR sequencing.....	72

List of Abbreviations

AF	Allele fraction
ALL	Acute lymphocytic leukemia
AlloSCT	Allogeneic stem cell transplant
AML	Acute myeloid leukemia
BM	Bone marrow
CDR3	Complementarity determining region 3
CLL	Chronic lymphocytic leukemia
CML	Chronic myeloid leukemia
CMV	Cytomegalovirus
CNA	Copy number alteration
CR	Complete response
DLI	Donor lymphocyte infusion
FCR	Fludarabine, cyclophosphamide, rituximab
FFS	Failure free survival
GEP	Gene expression profiling
GVHD	Graft versus host disease
GVL	Graft versus leukemia
HLA	Human leukocyte antigen
HPE	Homeostatic peripheral expansion
IGHV	Immunoglobulin heavy chain variable gene
ITH	Intratumoral heterogeneity
LAA	Leukemia associated antigen
LOH	Loss of heterozygosity
MDACC	University of Texas MD Anderson Cancer Center

mHA	Minor histocompatibility antigen
MHC	Major histocompatibility complex
MRCA	Most recent common ancestor
MRD	Matched related donor
MUD	Matched unrelated donor
NGS	Next generation sequencing
NMA	Nonmyeloablative
OS	Overall survival
PBMC	Peripheral blood mononuclear cells
RIC	Reduced intensity conditioning
sIndel	Somatic insertion or deletion
SNP	Single nucleotide polymorphism
sSNV	Somatic single nucleotide variant
TCR	T cell receptor
WES	Whole exome sequencing

Chapter 1: Introduction

Intratumoral heterogeneity and clonal evolution in cancer: Intratumoral heterogeneity (ITH) refers to the biologic heterogeneity among cancer cells within a single patient's disease and presents a major barrier to the effective treatment of cancer [1]. Cancer develops from a founder clone that has acquired 'hallmark' neoplastic capabilities, including the ability to sustain chronic proliferation and to resist cell death [2]. Over time, daughter cells of the founder clone blindly acquire new genetic and epigenetic alterations, some that further enhance their biological fitness and some that enable them to evade various cancer treatments. By the time cancer is diagnosed, this process has given rise to billions of tumor cells that can be grouped into related, but heterogeneous population subsets or subclones. Cancer has been compared to the Darwinian model of evolution wherein cancer subclones with variable fitness undergo natural selection resulting in shifts within the tumor population over time [3]. Cancer treatment has been equated to artificial selection, which can serve as a population bottleneck and temporarily reduce ITH [4]. However, population heterogeneity is regenerated through the acquisition of mutations in the daughter cells of the therapy-resistant clones.

There is evidence for ITH in both solid and liquid tumors. In a study of ITH in glioblastoma, 4-6 tumor fragments at least 1 cm apart were subjected to copy number and gene expression profiling (GEP). On average, only 31% of copy number alterations (CNAs) were shared between all fragments from the same tumor. Moreover, in 6 of 10 cases with GEP data, fragments from the same tumor were classified into at least 2 distinct glioblastoma subtypes [5]. In a separate study, ITH in glioblastoma, particularly in *TP53* mutated disease, was identified as the source of treatment-resistant subclones that expanded at the time of tumor recurrence [6]. One of the first studies to evaluate genomic ITH in AML utilized whole genome sequencing and deep targeting sequencing on matched primary tumor and relapse (post-chemotherapy) samples from eight patients [7]. Multiple clusters of mutations were detected in the primary

disease in 5 of 8 patients. For example, in the first patient, four distinct clusters of mutations (though related by founder events) were identified: clone 1 (the founder clone) made up 13% of the leukemia population, clone 2 (53%), clone 3 (29%), and clone 4 (5%). The relapsed disease was comprised entirely of clone 4, which had acquired additional mutations, including in the candidate drivers *ETV6* and *MYO18B*. As illustrated by this patient case, ITH underlies clonal evolution, which is a frequent mechanism of therapy resistance to both cytotoxic and targeted therapies. Examples of ITH and clonal evolution in the context of targeted therapy include the outgrowth of CD19 negative subclones after CAR T cell therapy in B-ALL [8] and the selective expansion of *KRAS* mutant subclones under therapeutic pressure with panitumumab, an anti-EGFR antibody [9].

Chronic lymphocytic leukemia: Chronic lymphocytic leukemia (CLL) is the most common adult leukemia and is characterized by the proliferation of mature B cells in the blood, bone marrow, and lymphoid tissues. CLL is typically a disease of older adults with a median age of diagnosis of 70, though ~ 10% of patients are between ages 45 - 54 when diagnosed. Most CLL cases are preceded by monoclonal B cell lymphocytosis, a common condition marked by the asymptomatic proliferation of clonal CD5+CD19+ B cells that progresses to CLL in ~1% of patients [10]. The cell of origin in CLL is debated but is proposed to be a mature CD5+ B cell [11] or a hematopoietic progenitor cell [12].

Relative to most cancers, the prognosis of CLL is excellent though several adverse features exist. One adverse prognostic feature is the presence of an unmutated immunoglobulin heavy-chain variable region (*IGHV*) gene, which accounts for ~ 50% of cases [13]. Deletion of 17p (including the *TP53* locus) and/or *TP53* mutation are also adverse features in the context of older chemoimmunotherapy and newer targeted therapies [14-16]. Next-generation sequencing (NGS) has been used to identify several recurrent driver alterations in CLL including chromosome losses (13q, 11q, 17p), gains (trisomy 12 and amp 2p), and mutated genes

(*SF3B1*, *ATM*, *TP53*, *NOTCH1*, *POT1*, *CHD2*, *XPO1*), but there is considerable interpatient heterogeneity and the most frequently mutated gene (*SF3B1*) is present in ~ 25% of cases [17].

CLL stage is assessed using the Rai or Binet staging systems, which evaluate for lymphadenopathy, hepatosplenomegaly, and progressive cytopenias including anemia and thrombocytopenia. Early stage CLL is characterized by isolated clonal B cell lymphocytosis and lymphadenopathy. Progressive anemia and thrombocytopenia signify advanced disease. The first rule for clinicians treating early stage CLL is to “watch and wait.” This is because CLL can exhibit indolent behavior for years before causing symptoms and early treatment can actually worsen survival by precipitating more aggressive disease [18]. Treatment is reserved for patients with B symptoms (fever, night sweats, weight loss), threatened end organ function, bulky splenomegaly or lymphadenopathy (> 10 cm), and progressive cytopenias. The “watch and wait” rule still applies despite the growing armamentarium of targeted CLL drugs that have been approved within the past few years. These agents, including ibrutinib and venetoclax, have revolutionized CLL treatment and have improved the survival of CLL patients regardless of age and disease subtype. For example, in the RESONATE-2 trial that compared frontline treatments for CLL patients ages 65 and older, the estimated 5 year overall survival (OS) was 83% for ibrutinib vs. 68% for chlorambucil (HR 0.45; 95% CI 0.266-0.761) [19].

CLL is an ideal model for investigating ITH and subclonal evolution for several reasons: (i) most patients have a relapsing, remitting disease course that spans many years permitting longitudinal study at multiple time points (ii) the leukemia is often accessible in the peripheral blood and (iii) CLL cells have a classic surface immunophenotype (CD5+CD19+, monoclonal kappa or lambda restricted) enabling enrichment or purification from other cell types. Landau et al. studied the subclonal evolution of CLL in the context of chemotherapy and demonstrated that the majority of patients (57 of 59) experienced leukemic subclonal evolution after chemotherapy, with frequent therapeutic selection of *TP53* variant subclones [17]. In contrast, subclonal evolution was observed in only 1 of 6 untreated patients. In addition, treatment of CLL with

ibrutinib was found to trigger clonal shifts in the canonical resistance mutation *BTK* (C481S) and in its protein partner *PLCG2* [20]. The delineation of CLL subclonal evolution caused by therapeutic selection provides the mechanistic basis for the “watch and wait” strategy that CLL clinicians have practiced for decades.

Allogeneic stem cell transplant and Donor lymphocyte infusion: Allogeneic stem cell transplant (alloSCT) is the transplantation of multipotent hematopoietic stem cells and mature leukocytes from donor bone marrow, peripheral blood, or umbilical cord into patients with malignant hematologic disorders or nonmalignant bone marrow failure disorders. The goal is reconstitution of normal donor-derived hematopoietic and immune function in the transplant recipient/host. AlloSCT was originally conceived after World War II as a means to reconstitute hematopoiesis in soldiers suffering from radiation-induced bone marrow failure [21]. Initial attempts resulted in high mortality due to the lack of human leukocyte antigen (HLA) matching; however, in the 1970s, Thomas et al. published a report detailing early success of sibling donor alloSCT for acute leukemia [22]. Further investigation of alloSCT for leukemia revealed the donor graft was doing more than reconstituting hematopoiesis and that transferred donor T cells were participating in the elimination of donor leukemia through a phenomenon known as the graft-versus-leukemia (GVL) effect [23]. Unfortunately, the GVL effect frequently came at the cost of graft-versus-host disease (GVHD). GVHD is divided into acute and chronic disorders that, in the matched HLA setting, are mediated by donor T cell recognition of minor histocompatibility antigens (mHAs) presented on the surface of normal host cells. mHAs are MHC-bound peptides that derive from polymorphisms between the transplant recipient and donor genomes, the majority of which are single nucleotide polymorphisms (SNPs). Donor immune cells recognize mHAs as non-self and cause GVHD, which manifests as multiorgan tissue inflammation/fibrosis including in the GI tract, liver, skin, lungs, and mucosal surfaces.

GVHD is treated with systemic immunosuppression using agents like glucocorticoids and calcineurin-inhibitors (cyclosporine, tacrolimus). These agents are tapered as GVHD

subsides and, in some cases, withdrawal of immunosuppression is accompanied by leukemia remission providing evidence of the GVL effect [24-26]. Studies of T-cell replete versus T-cell deplete alloSCT also highlight the importance of the GVL effect since a higher incidence of disease relapse was observed among recipients of T-cell depleted transplants [27-30]. Finally, trials of donor lymphocyte infusions (DLI) for relapsed leukemia provide the strongest evidence of the GVL effect. DLI is a treatment composed of unselected, polyclonal donor lymphocytes (from the original alloSCT donor) that is usually administered in the absence of conditioning chemotherapy [31, 32]. DLI can induce disease remission in CLL, acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL), and chronic myeloid leukemia (CML), which has the highest rate of DLI response (~70-80%) [33-37]. T cells operative in GVL eliminate leukemia cells through the recognition of leukemia-associated antigens (LAAs), mHAs, and leukemic neoantigens [38-41]. LAAs are nonpolymorphic self antigens that are overexpressed or mislocalized in malignant cells [42]. Leukemic neoantigens are tumor-specific antigens that result from exonic somatic alterations [43, 44].

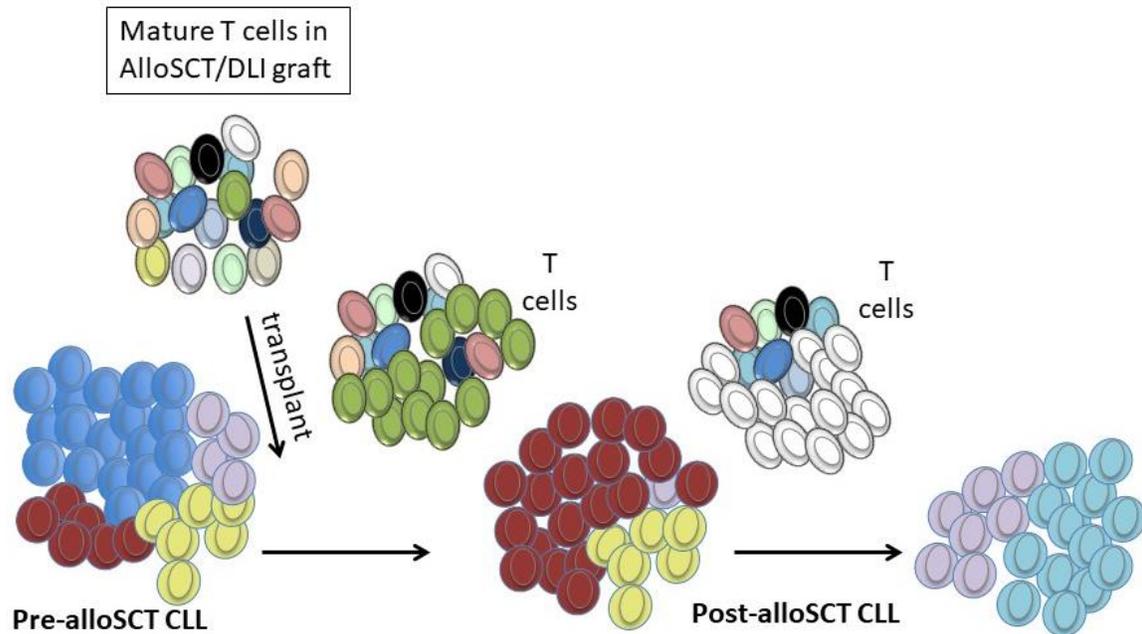
AlloSCT and DLI are original forms of cancer immunotherapy as they rely on donor T cells for their anti-tumor effect. Cancer immunotherapy no longer needs an introduction and now represents an integral part of the standard treatment for multiple tumor types [45]. The remarkable and durable cancer remissions achieved by checkpoint inhibitors and chimeric antigen receptor T cells in a subset of patients have proven the potency and the memory of antitumor immunity [46, 47]. Similarly, alloSCT remains a standard of care therapy for many forms of leukemia because it, too, can cure a subset of patients with chemorefractory and otherwise lethal disease. New therapies have largely replaced alloSCT for CLL because of its attendant toxicity, though alloSCT remains a viable treatment option, particularly for the most refractory cases [48].

A clinical trial at MDACC demonstrated CLL's susceptibility to the GVL effect. Patients with refractory CLL were treated with a nonmyeloablative, HLA-matched alloSCT and 20 of 43

patients (47%) who received immune manipulation after transplant (DLI and/or withdrawal of immunosuppression) for persistent/recurrent disease experienced a complete remission. Overall, the estimated 5-year survival rate was 51% [49]. Similarly, the CLL3X trial studied reduced-intensity alloSCT in patients with poor-risk CLL. Among 90 patients, the relapse incidence with a median follow-up of 6 years was 46% and 6-year OS was 58% [50]. Single-center studies at the Fred Hutchinson Cancer Center and Dana Farber Cancer Institutes reported similar results for alloSCT in CLL [51, 52]. We chose to study CLL in the context of nonmyeloablative, HLA-matched alloSCT as a model of subclonal evolution since approximately half of cases prove sensitive to the GVL effect while the remaining half are resistant.

Hypothesis and Specific Aims: Treatment-driven selection of subclones within a tumor population can cause a patient's malignancy to evolve, often leading to more aggressive disease. To improve the durability of cancer therapy and to design rational combinations, it is important to understand how various treatments impact malignant evolution. Longitudinal genomic analyses from pre- and post-treatment tumor samples can provide insight into the genetic factors that confer treatment resistance and responsiveness. The evolution of leukemia after chemo- and targeted therapies has been investigated [17, 20, 53-56], however, the impact of alloSCT on subclonal evolution is unknown.

We investigated the therapeutic selective pressure imposed by alloSCT and DLI and hypothesized that allogeneic T cells reshape the subclonal architecture of CLL by eliminating immunogenic subclones and permitting the expansion of immune evasive subclones. In turn, we expected that leukemia exerts selective pressure on anti-tumor donor T cells, characterized by their increased clonal frequency within the T-cell repertoire at the time of disease response (see proposed model below).



The primary objective of this proposal was to analyze longitudinal samples from a cohort of CLL patients that were treated with alloSCT and DLI. We did so via the following specific aims:

AIM 1. *To characterize the hierarchical architecture of CLL subclones and clinical evolution after alloSCT and DLI*

AIM 2. *To identify T-cell clones that mediate anti-CLL activity and characterize their phenotype and persistence in longitudinal patient samples*

Here, we tested this hypothesis utilizing whole exome sequencing (WES) of purified leukemia from a cohort of 24 CLL patients treated with alloSCT/DLI. In the same cohort, we assessed for reciprocal changes within the allogeneic T cell compartment. Our cohort included 11 patients who relapsed after alloSCT, and in these patients, we sampled longitudinal time points to examine leukemic and T cell coevolution. We provide evidence of ongoing CLL subclonal evolution after alloSCT/DLI and provide a strategy to track candidate GVL T cell clones through their expansion within the T cell repertoire during periods of leukemia response.

Chapter 2: Genomic evolution of CLL after alloSCT

Background: We hypothesized that allogeneic T-cell immunotherapies, including alloSCT and DLI, would cause subclonal leukemic evolution through the application of selective immunologic pressure. To test this prediction, we initially focused on 11 CLL patients who relapsed after HLA-matched, nonmyeloablative alloSCT and for whom longitudinal post-transplant leukemia samples were available. Using whole exome sequencing (WES), we compared the allelic fraction (AF) of somatic mutations and CNAs over time at pre-alloSCT and serial post-alloSCT time points to assess for molecular changes in the disease.

Methods: *Patient samples* Patient and healthy donor samples were obtained after appropriate informed consent through institutional review board approved protocols at The University of Texas MD Anderson Cancer Center (MDACC). Patient peripheral blood (PB) and bone marrow (BM) mononuclear cells were separated using histopaque 1077 prior to initial cryopreservation (FBS with 10% DMSO) and were stored in liquid nitrogen until the time of analysis. The clinical charts of CLL patients with stored longitudinal samples were reviewed and patients with eventual post-transplant relapse were selected for the initial analysis. For 20/24 patients, HLA testing was conducted at the MDACC HLA typing laboratory. For 4 patients with only serologic HLA typing available, refined HLA typing was inferred from exome data using the winners output from Polysolver [57]. *Cell purification* PB mononuclear cells and BM aspirate cells were thawed and stained with the following antibodies prior to electrostatic droplet-based cell sorting: anti-CD19 FITC (clone SJ2SC1), anti-CD5 PE (clone UCHT2), anti-Ig λ light chain Pacific Blue (clone MHL-38), anti-Ig κ light chain Pacific Blue (clone MHL-49), anti-CD3 PE/Cy7 (clone SK7), anti-CD8 PE/Cy7 (clone SK1) all from BioLegend and Sytox Red live/dead stain. CD19+CD5+ CLL cells and CD3+ T cells were sorted on the FACS Aria Fusion using a 70 μ M nozzle at 70 psi with a purity mask (Y32-P32-Ph0) after excluding debris, doublets, and dead cells. Cells were thawed and stained in phosphate-buffered saline containing 2% FBS. Sorted cells were

collected in RPMI before DNA isolation. DNA extraction and WES Genomic DNA was extracted from CLL cells and pre-alloSCT T cells (for germline DNA) using the QIAamp DNA Blood Mini Kit (Qiagen). Tumor and germline DNA concentration and quality were measured using fluorometric quantification (Qubit, ThermoFisher, and Fragment Analyzer, Advanced Analytical). Libraries were constructed from genomic DNA using the KAPA Library Preparation Kit (Roche). Exome capture was performed using the NimbleGen SeqCap EZ Exome Enrichment Kit v3.0 (Roche). Multiplex sequencing of samples was conducted on the Illumina HiSeq 2000 using 76 base pair paired-end reads at the MDACC Sequencing and Microarray facility. The mean target coverage was 120X per tumor sample (range: 39 – 389; SD +/- 45) and 112X per germline sample (range: 50 – 162; SD +/- 31). Paired-end sequencing reads in FASTQ format were generated from BCL raw data using Illumina CASAA software and aligned to the human reference genome (UCSC Genome Browser, hg19) using Burrows-Wheeler Aligner on default settings with the following exceptions: seed length of 40, maximum edit distance of 3, and maximum edit distance in the seed of 2 [58]. Aligned reads were processed using the GATK Best Practices of duplicate removal, indel realignment, and base recalibration. Sequencing was targeted to an overall coverage of 120X for target samples and 100X for matched germline samples. Variant calling Somatic single nucleotide variants (sSNVs) located within the exome were identified using MuTect [59]. These data were further filtered using the following criteria: (i) minimum total read count in the tumor ≥ 20 (ii) minimum total read count in the germline ≥ 10 (iii) minimum alternate allele frequency in the germline ≤ 0.01 and (iv) variants in positions listed in ESP6500 [60] and 1000G [61] with minor allele frequencies > 0.01 were removed. For the CLL evolution analyses, only somatic non-silent mutations detected recurrently across longitudinal samples with an allelic fraction (AF) > 0.05 were considered. To categorize clonal versus subclonal sSNVs, we determined the mean AF at which recurrent CLL driver genes (initial AF filter at 0.25) appeared in the sort-purified pre-transplant leukemia, which was

calculated to be 0.477. sSNVs were considered clonal if the AF was greater than two standard deviations (SD, 0.066) below the mean, which set the threshold at an AF > 0.34. For the AF scatter plots, synonymous mutations and variants located within non-coding RNA and 5'/3' untranslated regions at an AF > 0.15 were included to aid visualization of mutation clusters. Small somatic insertions and deletions (indels) were identified using Pindel [62]. The following stringent filtering criteria were added to increase specificity of the indel calls: (i) present within exons of reported CLL driver genes [56, 63] at an AF > 0.05 and (ii) events manually viewed and confirmed using the integrated genome viewer [64]. Substitutions and indels were annotated using ANNOVAR [65] based on known genes within UCSC. To assess donor cell contamination of post-transplant leukemia, donor samples for 6 patients were sequenced to a depth of 112X at the exome (3 samples from donor pheresis products and 3 samples from host PB drawn when the host had full donor chimerism at a time of molecular remission). Platypus [66] was used to detect variants, primarily single nucleotide polymorphisms (SNPs), in the donor sample. The Platypus calls from germline donor exomes were then compared to the MuTect exonic calls from host sort-purified leukemia to determine the AF of any overlapping variants.

Detection of copy number alterations: Genome-wide copy number calls for the leukemia samples and their patient matched germline DNA were derived from exome data using Circular Binary Segmentation [67], followed by an in-house tool, exome CN, for generating log₂ ratio scores. A segment of gain had a log₂ score of >0.5 while a segment of loss had a log₂ score of <-0.5. These calls were all verified by manual inspection. For integration of CNA and mutation data in the scatter plots, we converted the log₂ value to an absolute copy number and assumed a diploid genome in the germline. The copy number data were intersected with the somatic non-silent mutation data for each patient to confirm that evolution of somatic SNVs was due to variation in AF rather than change in copy number. CLL immunoglobulin heavy chain (IGH)

CDR3 analysis To investigate the clonality and relatedness of early and late disease, the CDR3 region of the IGH was directly analyzed for CLL patients 3 and 9 using longitudinal sort-purified

CLL samples. Targeted massively parallel sequencing was conducted by Adaptive biotechnologies and the data were analyzed using the Immunoseq analyzer software.

Results: *Patients* To test our hypothesis that allogeneic T cells impose selective pressure on heterogeneous leukemic subclones, we focused on 11 CLL patients who relapsed after alloSCT and for whom a pre-transplant and longitudinal post-transplant CLL samples were available (CLL patients 1-11, Tables 1 and 2). All patients were heavily pretreated with chemotherapy, including with fludarabine, cyclophosphamide, and rituximab (FCR). This is a critical detail because 10 of the 11 patients also received FCR as their conditioning chemotherapy prior to alloSCT. In this context, we would expect that any observed post-alloSCT disease evolution would be attributable, at least in part, to the graft rather than the conditioning chemotherapy because of the recent and significant exposure to FCR. After alloSCT, the median time to retreatment for the 11 patients was 396 days (range, 124-1072 days). We analyzed serial samples for each patient (median, 4; range, 2-12) and, for some patients, our investigation of longitudinal post-alloSCT leukemia spanned several years (median, 4.4; range: 1.7-8.6).

WES of post-alloSCT CLL The post-alloSCT setting poses unique challenges for genomic analyses as there are potentially 3 types of exomes present: (i) host germline DNA (varies with hematopoietic chimerism) (ii) host tumor DNA and (iii) donor germline DNA. To maximize tumor purity and to minimize contamination by host and donor germline DNA, we sort-purified the autologous pre-alloSCT T cells (CD3+, for germline) and CLL cells (CD19+CD5+) prior to DNA extraction and WES (Figure 1). We then examined the level of donor DNA contamination in the post-alloSCT purified CLL samples from 6 patients to confirm the detected variants were host tumor variants and not SNPs in donor DNA that were called against the host germline DNA.

We directly assayed for donor SNPs in 23 post-transplant CLL samples using the strategy described in the 'Variant Calling' methods section. For 2 patients (CLL patients 1 and 4), there was no donor contamination observed. For 3 patients (CLL patients 2, 3, and 5), genetic contamination from the alloSCT donor was detected at a low level in 6 of 8 samples.

The donor variants, which are disparate SNPs between host and donor, were detected in the sorted CD19+CD5+ fraction at a mean allelic fraction (AF) of 0.047 (SD \pm 0.014) in CLL patient 2, 0.048 (SD \pm 0.018) in CLL patient 3, and 0.052 (SD \pm 0.014) in CLL patient 5 (Figure 2A). As a positive control, we analyzed the post-transplant FFPE BM sample from CLL patient 3, which was taken at a time point when clinical hematopathology reported CLL in more than 50% of examined cells and full donor chimerism in T and myeloid cells. Since the leukemia was not

Figure 1

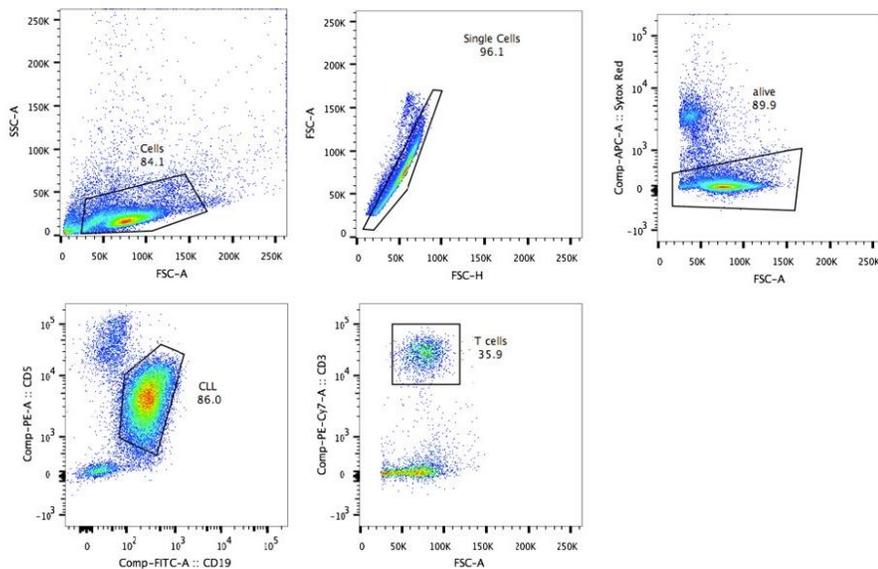


Figure 1. Sort purification of CLL and T cells

Gating strategy for the FACS of CD19+CD5+ CLL cells for leukemic DNA and CD3+ T cells for germline DNA. Cells were also evaluated for Ig k or Ig λ light chains to verify monoclonal expression.

Table 1. CLL prognostic features and prior treatments for the alloSCT cohort

CLL case	Age at diag. / Sex	IGHV status	Therapies prior to AlloSCT	Cytogenetics prior to AlloSCT (by G banding - GTG)	Cytogenetics prior to AlloSCT (FISH) (for 11q, 13q, and 17p deletions and trisomy 12)	CD19/CD38+ pre-AlloSCT (as %age of lymphs)	Disease status prior to AlloSCT (% CD19/5+ in BM)
1	53M	unmutated	FC, Chl, R, VCP	del11q14, del13q13q21	not tested	51.2%	active disease (99%)
2	58F	unmutated	F, FCR, RA	diploid	negative	6.2%	active disease (36.9%)
3	47F	unmutated	F, Chl, R, RA FCR	diploid	not tested	37.1%	active disease (88.2%)
4	48M	unmutated	F, FCR, CFAR	diploid	not tested	34.2%	CR
5	47M	unmutated	F, FCR	diploid	not tested	82.0%	active disease (49%)
6	64F	unmutated	R, FCR, RA	diploid	negative	54.3%	active disease (13.8%)
7	54M	unmutated	Chl, FCR	diploid	not tested	11.5%	active disease (10.2%)
8	56M	unmutated	PChl, F, R, FCR	diploid	no trisomy 12	3.1%	active disease (59.3%)
9	54M	unmutated	F, C, RA, Hyper-CVAD/MTX-Ara-C	diploid	not tested	3.5%	active disease (44.4%)
10	50M	unmutated	PR, FCR	del13q12q22	del 13q	63.1%	active disease (32.8%)
11	50F	unmutated	P, FCR, CFAR	del11q13q23	del 11q and del 17p	60.2%	active disease (92.5%)
12	44M	unmutated	F, FCR, R-CHVP, CVP-R, OFAR, Hyper-CVAD/MTX-Ara-C	del11q13q22.2	negative	7.3%	active disease (48.5%)
13	58M	unmutated	Chl+P, R, FC, R-CHVP, FCR + CD23 Mab, OFAR	del13q14q21 [2]; del11q21 [1]; diploid [17]	del 11q and del 13q	96.2%	active disease (79.4%)
14	54M	unmutated	Chl, FCR, CFAR, OFAR	complex: 45,XY,-1,del(3)(q26.2),-5,inv(6)(p21.3q25),add(9)(p13),-17,-18,+3mar	del 17p	96.9%	active disease (68%)
15	41M	not tested	Chl, Chl+P, FC, FCR, RA	del11q13	not tested	63.4%	active disease (83.7%)
16	53M	not tested	FCR, Hyper-CVAD/MTX-Ara-C	diploid	not tested	80.6%	active disease (19.4%)
17	55M	unmutated	FCR, RA	del11q14	not tested	58.6%	active disease (72.4%)
18	43F	unmutated	FCR, RA, pentostatin+CR	trisomy 12	trisomy 12	37.5%	active disease (58%)
19	53M	unmutated	CP, FCR, OFAR, hyper-CVAD-R	diploid	trisomy 12	4.6%	active disease (79%)
20	49F	unmutated	R-CHVP, A, FCR	diploid	del 13q14.3, del 13q34, del 17p	29.2%	active disease (17.9%)
21	44M	unmutated	FCR, OFAR, Hyper-CVAD-R	diploid	del 17p	97.9%	active disease (82.7%)
22	47M	unmutated	CHVP, R, F, R-ESHAP, Hyper-CVD	diploid	del 11q	14.8%	active disease (92.9%)
23	57M	unmutated	FC, FCR, A	add(11)q22	del 11q, del 13q	92.2%	active disease (92%)
24	52M	not tested	FCR, AR	45, X,-Y [13], diploid [7]	del 11q	96.5%	active disease (17.1%)

Abbreviations: A, alemtuzumab; C, cyclophosphamide; Chl, chlorambucil; ESHAP, etoposide, methylprednisone, cytarabine, cisplatin; F, fludarabine; Hyper-CVAD/MTX-Ara-C, fractionated cyclophosphamide, doxorubicin, vincristine, and dexamethasone alternating with high-dose methotrexate and cytarabine; O, oxaliplatin; P, prednisone; R, rituximab; V, vincristine.

Table 2. Patient and donor alloSCT characteristics

CLL case	AlloSCT source	HLA matching	AlloSCT conditioning (MA, NMA, or RIC)	AlloSCT conditioning (detailed)	Time from diagnosis to AlloSCT (yrs)	Response at 2 months post-AlloSCT*	Failure-free survival post-AlloSCT (days)	# of DLIs	Acute GVHD (grade III or IV)**	Chronic GVHD (moderate or severe)	COD
1	brother	identical by serology and PCR-SSP	NMA	FC	9.9	PD	124	6	no	no	progressive disease
2	MUD	10 of 10	NMA	FCRA	3.1	PR	396	2	yes, stage 3 skin	no	N/A
3	sister	10 of 10	NMA	FCR	5.5	PD	450	1	yes, stage 3 skin	moderate	progressive disease
4	MUD	10 of 10	NMA	FCRA	10.4	CRi w/ MRD	271	3	no	severe	GVHD and infection
5	sister	9 of 10 (1 mismatch at DQB1)	NMA	FCR	6.2	PR	165	5	no	no	progressive disease and infection
6	sister	10 of 10	NMA	FCR	2.4	SD	1072	0	yes, stage 3 skin	no	progressive disease, Richter's syndrome
7	sister	10 of 10	NMA	FCR	5.5	PR	396	1	yes, stage 3 GI	no	progressive disease and infection
8	sister	10 of 10	NMA	FCR	3.6	PR	177	2	no	no	progressive disease and infection
9	brother	identical by serology and PCR-SSP	RIC	FA + M [#]	6.3	CRi w/ MRD	406	1	no	moderate	FTT and infection
10	brother	10 of 10	NMA	FCRI	1.1	PR	619	3	no	no	N/A
11	cord	10 of 10	NMA	FCRT	7.4	PR	690	0	no	severe	progressive disease and infection
12	MUD	10 of 10	NMA	FCRA	7.9	CR	2320	0	no	severe	unknown
13	MUD	10 of 10	RIC	FR + M	6.0	SD	3124	0	no	moderate	progressive disease
14	MUD	10 of 10	RIC	FR + M	11.3	PR	3955	0	yes, stage 3 skin	moderate	N/A
15	sister	10 of 10	NMA	FCR	18.1	PR	4687	0	no	no	N/A
16	sister	10 of 10	NMA	FCR	6.4	CR	4500	0	yes, stage 3 skin	no	N/A
17	MUD	9 of 10 (1 mismatch at C locus)	NMA	FCR	3.1	PR	4808	0	yes, stage 3 skin	moderate	N/A
18	MUD	10 of 10	MA	BEAM+AR	8.1	PR	3311	0	no	no	N/A
19	MUD	10 of 10	NMA	FCRI	13.9	PR	2521	0	no	no	N/A
20	brother	10 of 10	NMA	FCR	1.6	CR	3821	0	no	no	N/A
21	MUD	10 of 10	NMA	FCR	5.9	PR	2662	0	no	moderate	N/A
22	sister	10 of 10	RIC	FR + M	7.8	SD	426	0	no	severe	steroid refractory GVHD and infection
23	sister	10 of 10	RIC	F + M	9.4	PR	463	0	yes, stage 3 skin	severe	steroid refractory GVHD and infection
24	brother	10 of 10	NMA	R + I	7.5	PR	3591	0	yes, stage 3 skin	no	N/A

Abbreviations: A, alemtuzumab; BEAM, carmustine, etoposide, cytarabine, melphalan; C, cyclophosphamide; COD, cause of death; cord, umbilical cord graft; CR, complete remission; CRi w/ MRD, complete remission with incomplete marrow recovery and evidence of minimal residual disease; F, fludarabine; FTT, failure to thrive; I, ibrutinomab tiuxetan; M, melphalan; MA, myeloablative; MUD, matched unrelated donor; NMA, nonmyeloablative; PD, progressive disease; PR, partial remission; R, rituximab; RIC, reduced-intensity conditioning; SD, stable disease; SSP-PCR, sequence-specific amplification (PCR) for HLA typing; T, low dose TBI (2 Gy).

* Responses graded according to the guidelines set by the International Workshop on CLL

** GVHD graded according to the NIH consensus criteria

Patient had 2 transplants from his brother; first with FCR conditioning (failed engraftment) and next with FA + M

sort-purified, a large number of donor SNPs ($n = 3184$) were detected among the CLL somatic calls at a mean AF of 0.13 ($SD \pm 0.047$) (Figure 2A, inset). Of note, the CLL driver mutations that were detected in this sample (*ATM*, *XPO1*, and *SF3B1* mutations) appeared well above the level of contamination at AF's of 0.36, 0.33, and 0.28, respectively.

Our sequencing pipeline flags all variants present in the dbSNP database [68] and correctly labeled > 97% of the contaminating donor variants as SNPs. Importantly, our approach resulted in concordant detection of mutations in serial samples as is shown over 7 years for CLL patient 10 (Figure 2B). For this patient, the WES data demonstrated the absence of post-transplant disease evolution despite our detection of donor cell contamination in the initial post-alloSCT CLL sample (represented with an asterisk). Taken together, these data indicate that our

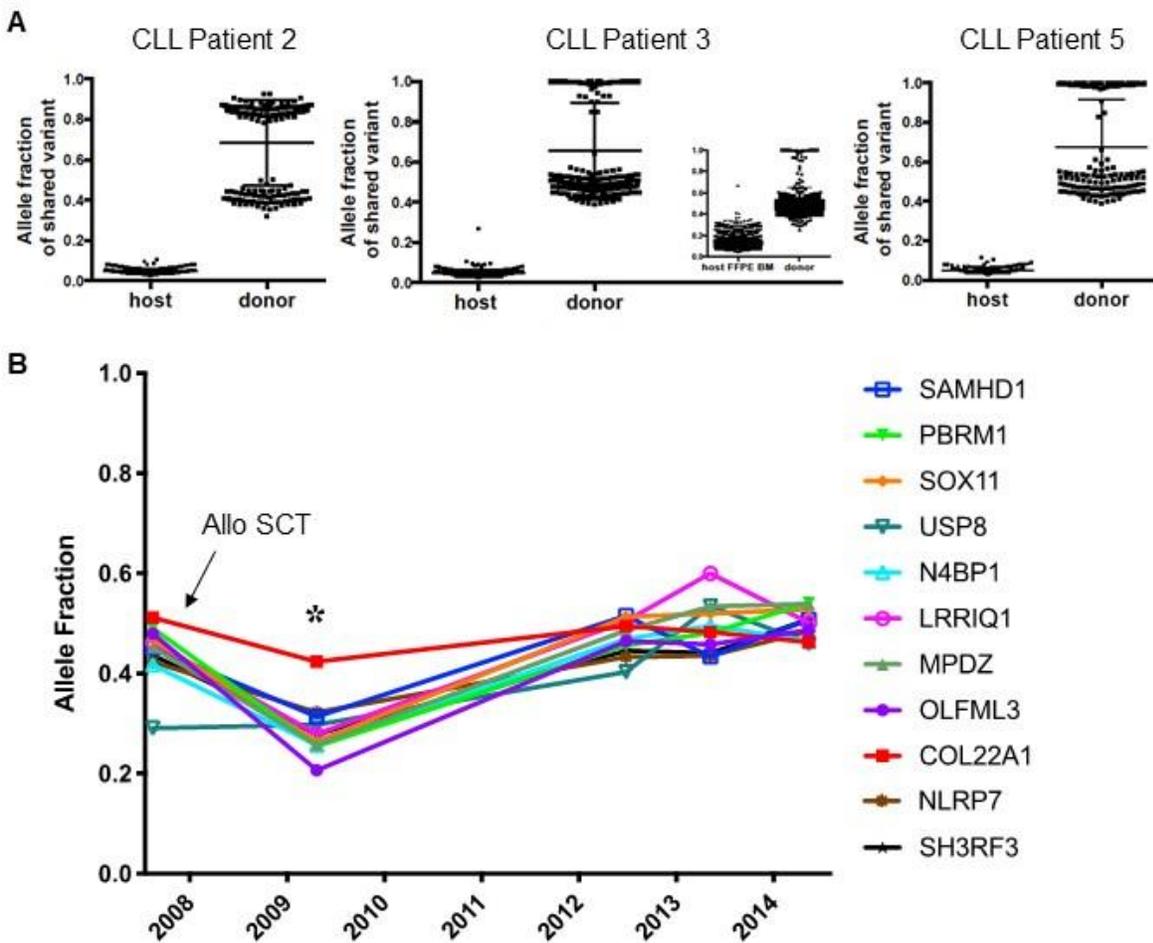


Figure 2. WES of post-alloSCT CLL is feasible despite donor chimerism

Overlapping SNVs found in both the normal donor cells and host post-transplant, sort purified samples are plotted at the AF detected in the host and donor samples for CLL 2, CLL 3, and CLL 5 (A). The inset for CLL 3 shows overlapping SNVs found in normal donor cells and host unsorted post-alloSCT BM as a comparison. (B) Recurrent, somatic exonic mutations and their corresponding AFs are plotted to demonstrate the concordance in WES across ~7 years of PB/BM samples from CLL 10 whose disease showed no evolution after transplant. The timing of alloSCT is indicated (the first sample, on the Y-axis, is a pre-alloSCT sample) and the asterisk denotes the post-SCT CLL sample where the greatest degree of donor SNP contamination was detected.

strategy resulted in reliable detection of host-specific somatic CLL mutations in post-transplant leukemia samples.

CLL subclonal evolution occurs after DLI

To examine leukemic evolution after alloSCT, we compared the AF of somatic mutations detected in longitudinal CLL patient samples. Clear patterns of evolution emerged, and in 2 patients (CLL 5 and CLL 8), evolution coincided with DLI administration for relapsed disease. The clinical course and WES windows are shown for CLL 5 (Fig. 3A) and CLL 8 (Fig. 3B). Sort purified CLL was sequenced at 5 time points (blue asterisks) in both patients (Figs. 3C and 3D). Branched leukemic evolution was observed with a branch point between time points 3 and 4, when CLL 5 received 3 DLIs (Fig. 3C). Outgrowth of a leukemic subclone containing non-silent mutations in *EGR2*, *NOTCH1*, *XPO1*, and a new mutation in *ASXL1* (p.M1345V) was seen post-DLI in concert with the elimination of a related subclone containing the *EGR2* mutation as well as a distinct *NOTCH1* nonsense mutation (p.S2492X) among several other variants.

CLL 8 experienced a similar relapsing and remitting clinical course (Fig. 3B). Branched leukemic subclonal evolution again coincided with DLI treatment (Fig. 3D). The pre-alloSCT

leukemic clone, harboring missense mutations in *RAG1*, *HERC2*, *CHD11*, and *LIFR* as well as a chromosome 11q deletion, contracted after the lymphocyte infusion, while a subclone expanded to compose the relapsed disease. This refractory subclone shared the *RAG1*, *HERC2*, and *CDH11* mutations, implying they were acquired prior to the branch; however, the late subclone carried additional mutations in *TP53* and *ASXL1* as well as an amplification of chromosome region 2p (Fig. 3E).

We performed this analysis in all 11 patients with available longitudinal post-alloSCT samples and observed branched CLL evolution in 5 patients, linear evolution in 3 patients, and no evolution in 3 patients (Figure 4). Notably, studies of CLL evolution after chemotherapy have also demonstrated mixed patterns of branched and linear evolution (6, 7). In summary, the data support our hypothesis that allogeneic T cells shape leukemic subclonal architecture after transplant.

Figure 3

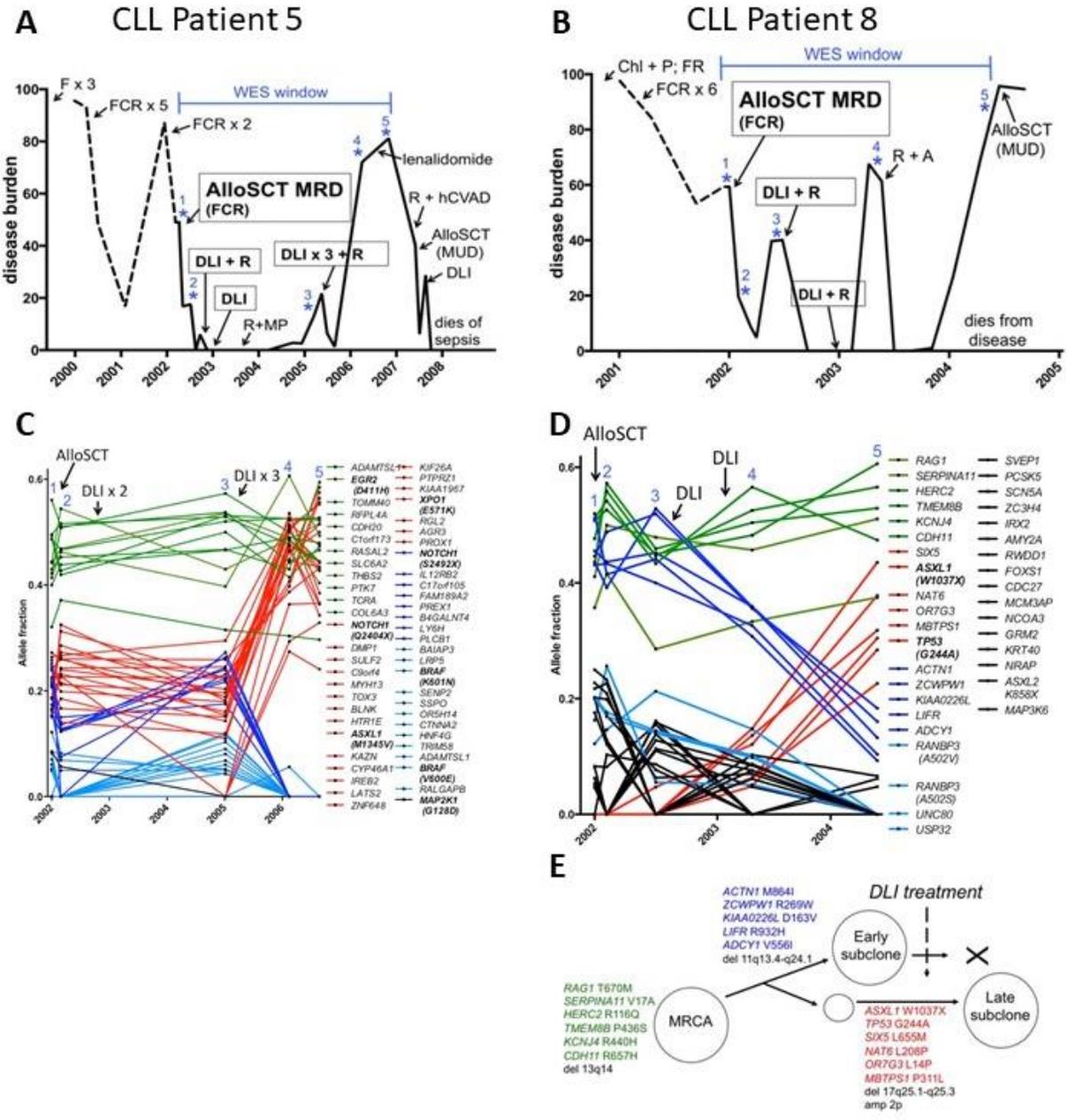


Figure 3. Branched evolution of CLL subclones occurs after DLI in patients CLL5 and CLL8

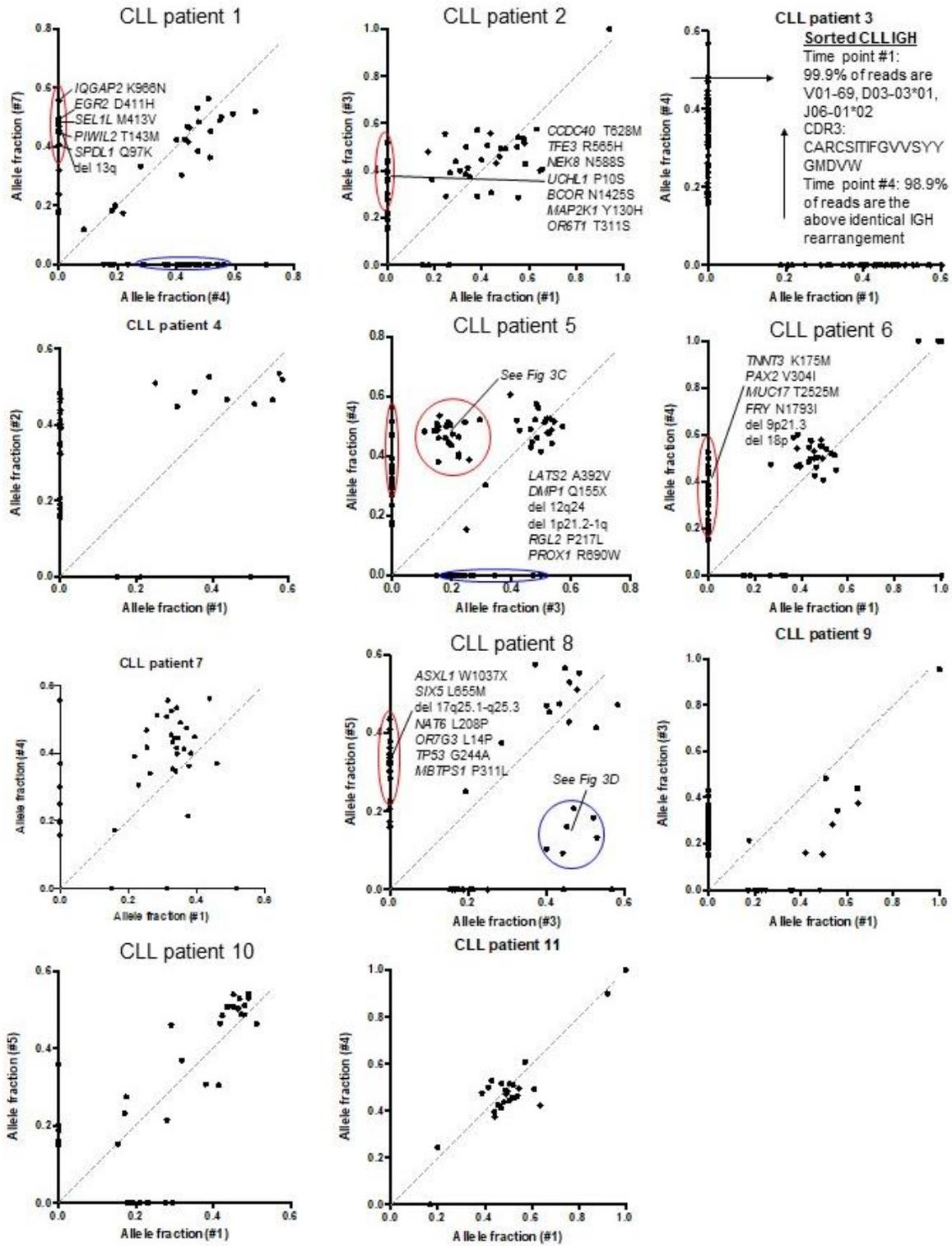
Clinical course for CLL 5 (A) and CLL 8 (B). The y-axis indicates disease burden as the percentage of lymphocytes within the bone marrow aspirate sample that are CD19+CD5+ CLL cells. Pre-transplant disease is represented as a dotted line and post-transplant disease as a

solid line. Administered treatments are indicated with arrows and number of cycles. Numbered asterisks (blue) delineate the time points when WES of sort purified CLL cells was performed. Longitudinal AF plots for CLL 5 (C) and CLL 8 (D). Recurrent, somatic exonic nonsynonymous mutations are shown over time for the 5 time points indicated by the asterisks in A and B. Previously reported CLL candidate driver genes are bolded, and the protein substitutions are listed. Clonal mutations (green), mutations enriched after DLI (red), and those that diminish after DLI (dark blue for AF changes by > 0.15 and light blue for AF changes by < 0.15) are shown. Mutations in black represent additional mutations seen in the disease course for CLL 8 (E) Model of subclonal evolution for CLL 8. MRCA denotes the most recent common leukemic ancestor. Abbreviations: A, alemtuzumab; F, fludarabine; FCR, fludarabine, cyclophosphamide, rituximab; hCVAD, hyperfractionated cyclophosphamide, vincristine, doxorubicin, dexamethasone; MP, methylprednisolone; MRD, HLA-matched related donor; MUD, HLA-matched unrelated donor; P, prednisone; R, rituximab; Chl, chlorambucil.

Figure 4. AF scatter plots for CLL patients 1-11 showing somatic variants across 2 time points for each patient's course

The earlier time point is on the x-axis and the later time point is on the y-axis. The dotted diagonal line shows $y=x$, indicating no change in AF over time. Red and blue circles highlight clusters of mutations that include nonsynonymous mutations with a >0.2 AF change over time. Linear evolution is defined by clusters of nonsynonymous mutations with an AF > 0.2 arising after alloSCT and branched evolution is defined as both an increase and decrease in clusters of nonsynonymous mutations with an AF > 0.2 post-alloSCT. According to these criteria, patients were categorized as follows: CLL patients 1, 3, 5, 8 and 9 – branched, CLL patients 2, 4, and 6 – linear, CLL patients 7, 10 and 11 – no evolution.

Figure 4



CLL drivers evolve after alloSCT

We anticipated there would be immunoediting of the leukemic subclones by allogeneic T cells; however, given that CLL subclonal selection had likely already occurred after prior treatments [17], it was unknown whether the post-alloSCT/DLI relapsed disease would differ substantially from the pre-alloSCT leukemia, particularly with regard to driver lesions. To answer this question, we looked for changes in CLL drivers over time in the 11 patients with sequential samples available post-transplant.

We observed marked evolution of variants in established CLL driver genes and chromosome regions after alloSCT (Figure 5). For example, CLL patient 2 demonstrated linear evolution of CLL post-alloSCT, and the relapsed leukemia contained mutations in the drivers *BCOR* and *MAP2K1* [69] as well as a novel mutation in *EWSR1* (p.T100S). CLL patient 6 also demonstrated linear evolution post-alloSCT and acquired an 18p deletion late in her course.

After alloSCT, 44% of CLL drivers remained unchanged in their AFs, 36% expanded, and 20% contracted. Six of 11 patients experienced the emergence of at least one previously undetectable driver post-transplant, including mutations in *EGR2*, *XPO1*, *SF3B1*, and *TP53* as well as novel mutations in *DDX3X* (p.L320F) and *ATM* (p.N1094S). We also detected 2 new missense mutations in *SAMHD1* [70] (p.P227L in CLL patient 2 and p.M240V in CLL patient 10), both found clonally throughout their courses.

The canonical drivers evolved in the context of subclonal expansions/contractions along with additional sSNVs. AF scatter plots that highlight the evolving subclones are shown in Figure 4. These data suggest that selective pressure on the CLL population continues after alloSCT.

Figure 5

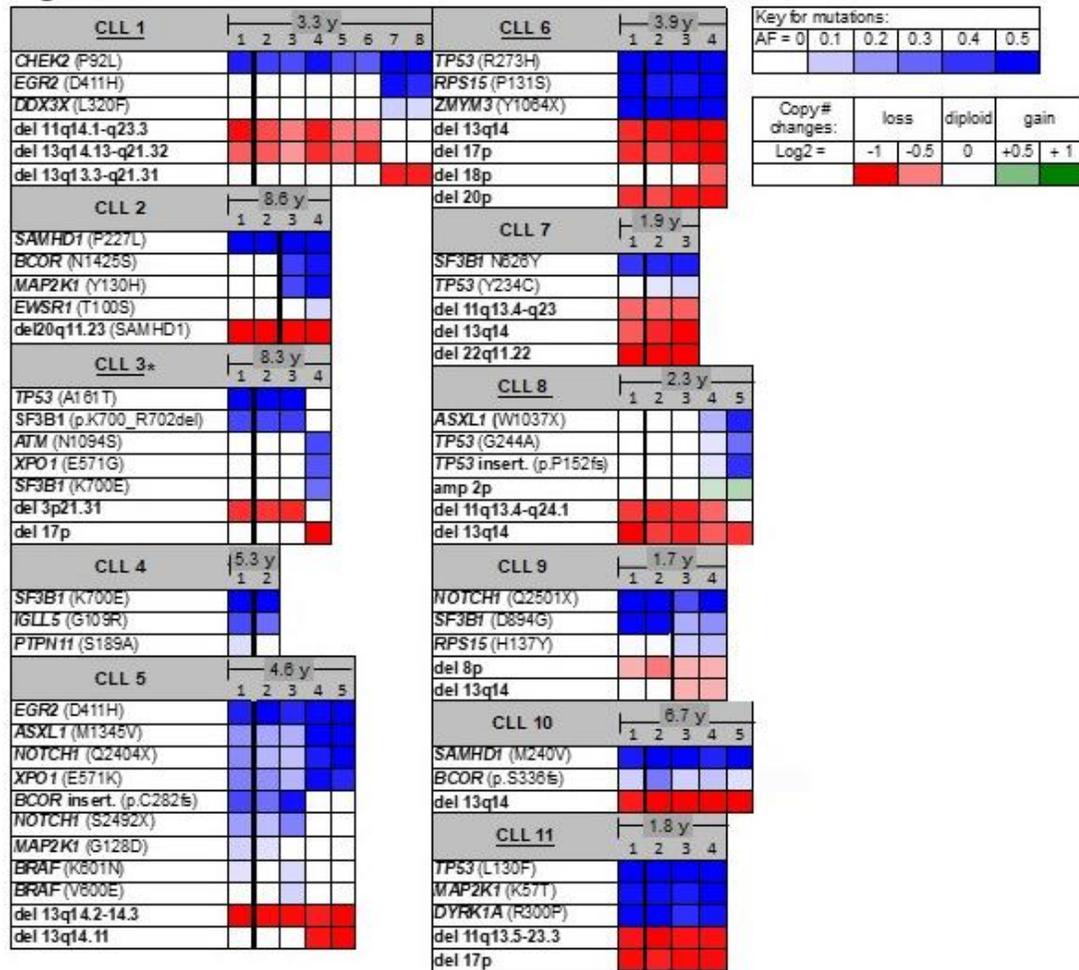


Figure 5. Evolution of CLL drivers and enrichment for novel mutations occurs after alloSCT and DLI

Heatmap indicating shifts in reported CLL driver lesions over time. The 11 CLL patients with longitudinal samples are shown with driver gene mutations shaded by AF (blue) and copy number alterations by log2 value (losses in red, gains in green). Each column represents a longitudinal time point in a patient's course and the bar above indicates the number of years between the first and last WES time point. Solid bars in bold indicate interval alloSCT. *The IGH CDR3 region of the early and late disease from CLL patient 3 was directly sequenced and was identical.

Chapter 3: Identification of a new candidate CLL driver through study of leukemic evolution

Background: When we began our study of peri-alloSCT CLL, the landscape of somatic mutations in untreated CLL was newly reported [71-74]. The landscape of somatic mutations and CNAs in heavily pretreated CLL remained unknown. In the first patient (CLL1), we uncovered a unique pattern of molecular disease evolution that led us to investigate a previously unreported alteration in *CHEK2*. While performing these experiments, a new report was published that listed *CHEK2* as a potential CLL driver, further supporting our hypothesis that the alteration we were investigating could have biologic relevance [17].

Methods: Microarray For CLL patient 1, genomic DNA from sort-purified populations of T cells (pre-alloSCT germline) and CLL cells was analyzed at 750,000 SNPs and 1.9 million non-polymorphic probes after hybridization to Affymetrix Cytoscan HD arrays (Affymetrix, Santa Clara, CA). Arrays were scanned using the GeneChip Scanner 3000 7G System at the MDACC SMF. Data were analyzed using the Affymetrix Chromosome Analysis Suite. Reverse transcription and CHEK2 cDNA sequencing Total RNA was isolated using the RNeasy Mini Kit (Qiagen) from a BM aspirate, which was collected at a time point when > 90% of BM cells were CLL, confirmed by pathology and flow cytometry for CLL patient 1. cDNA was prepared using iScript Reverse Transcription Supermix (Bio-Rad, Hercules, CA) according to the manufacturer's protocol. Primers were designed to amplify the cDNA from the region of interest within *CHEK2*: AAAGTCCAGCCAGTCCTCTC (forward) and TCTTTTCAGCAGTGGTTCATCA (reverse). Purified amplicons were sequenced from both strands using Big Dye terminator chemistry and an ABI 3730XL sequencer at the MDACC SMF Core facility. Chromatograms were viewed using FinchTV (Perkin Elmer, Waltham, MA) and sequences were compared to the germline reference sequence using Lasergene 12 software (DNASTAR, Madison, WI). CHK2 Co-Immunoprecipitation and Analytical Ultracentrifugation The plasmids 3XFlag-Chk2(P92L)

and Myc-Chk2(P92L), were generated by site-directed mutagenesis of 3XFlag-Chk2 and Myc-Chk2, respectively, using the QuickChange II XL kit (Agilent, Santa Clara, CA) and the following primers: 5'GAGCCTACCCCTGCCCTCTGGGCTCGATTATGGGCC-3' and 5'-GGCCATAATCGAGCCCAGAGGGCAGGGGTAGGCTC-3'. For oligomerization assays, HCT116 Chk2 *-/-* cells were co-transfected with 1 µg 3XFlag-Chk2 [wild-type (WT), I157T, or P92L] and 1 µg Myc-Chk2 (WT, I157T, or P92L) using Lipofectamine 2000. Cells were harvested 32 hours later and lysed in Mammalian Cell Lysis Buffer (MCLB); 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.5% NP40, 5mM EDTA, 2 mM DTT, 1 mM PMSF, 1 µM microcystin-LR, 1X Protease Inhibitor Cocktail (Sigma, St. Louis, MO), and 1X Phosphatase Inhibitor Cocktail Set II (Millipore, Billerica, MA). Protein from lysates was pre-cleared with Protein A Agarose (ThermoFisher) and then immunoprecipitated with anti-Flag M2 affinity gel (Sigma) or anti-c-Myc agarose (Santa Cruz Biotechnology, Dallas, TX) for 3 h at 4°C, and the resin was washed 4x with MCLB. Pre-cleared lysates and resin containing the immunoprecipitated proteins were boiled in Laemmli buffer for 5 min, and proteins were resolved by SDS-PAGE followed by Western blotting with the indicated antibodies. Antibody incubations were performed at the following dilutions in 1X PBST: 1:5000 mouse anti-Flag M5 (Sigma), 1:100 rabbit anti-c-Myc (Santa Cruz Biotechnology), 1:5000 HRP goat anti-mouse IgG (BD Pharmingen), and 1:5000 HRP goat anti-rabbit IgG (BD Pharmingen). Blots were visualized using SuperSignal Dura Substrate (ThermoFisher) and a Syngene G:box. Analytical ultracentrifugation Chk2 proteins (WT, P92L, I157T, and D368N) were purified from *E.coli* by the Center for Biomolecular Structure and Function at MDACC. Sedimentation equilibrium analytical ultracentrifugation (SEQ) experiments were performed at 20°C using a Beckman XL-I instrument with an AnTi 60 rotor. All samples were prepared in 20 mM HEPES, 200 mM NaCl, 2 mM TCEP at pH 7.5 and loaded into sample chambers with Epon double sector centerpieces and sapphire windows. SEQ scans were recorded using absorbance at 280 nm after 48 h incubation at 8000 rpm. The protein partial specific volume and solvent density were calculated using Sednterp 1.09 [75, 76].

Data analysis was performed using Sedphat 10.58d [77]. For each protein, the SEQ profiles were fitted to a monomer-dimer equilibrium model. A F-statistics error mapping approach [78] was used to determine the 99% confidence intervals for the dissociation constant.

Results:

Candidate driver *CHEK2* c.C275T (p.P92L) affects Chk2 dimerization

An unusual pattern of post-transplant subclonal evolution in CLL patient 1 facilitated the identification of a novel pathologic mutation in *CHEK2*. The AF plot and corresponding interval therapies for CLL patient 1 are shown in Figure 6A. Branched subclonal evolution was observed and the late disease contained a recurrent mutation in the CLL driver *EGR2* (p.D411H) as well as the characteristic CLL chromosome loss – deletion 13q. There was an 11q deletion in the pre-transplant and early post-transplant leukemia that was not detected in the late disease, implying that a rare subclone with diploid 11q may have expanded late in the patient's course.

To verify this result, we re-analyzed the samples using high-resolution microarray and confirmed post-alloSCT evolution of the chromosome 11q deletion as well as convergent evolution of distinct 13q deletions (Figure 6B). The branched leukemic evolution in CLL patient 1 suggested the 11q and 13q deletions emerged later in the patient's disease rather than early in leukemogenesis – the latter scenario being far more common (Figure 6C) [17]. Among the genetic lesions in the inferred most recent common leukemic ancestor, the *CHEK2* missense mutation (c.C275T; p.P92L) emerged as a potential pathogenic lesion and was confirmed to be expressed by Sanger sequencing of the mutant transcript (Figure 6D). Chk2, a serine/threonine protein kinase integral to the DNA damage response, is implicated in breast cancer and other solid malignancies [79, 80]. Importantly, *CHEK2* was added to list of most recurrently mutated CLL genes by Landau et al. who detected *CHEK2* mutations in 5 of 538 patients. Two of these mutations affected the same forkhead-associated (FHA) domain of Chk2 as in our patient but neither altered proline 92 [17].

Figure 6

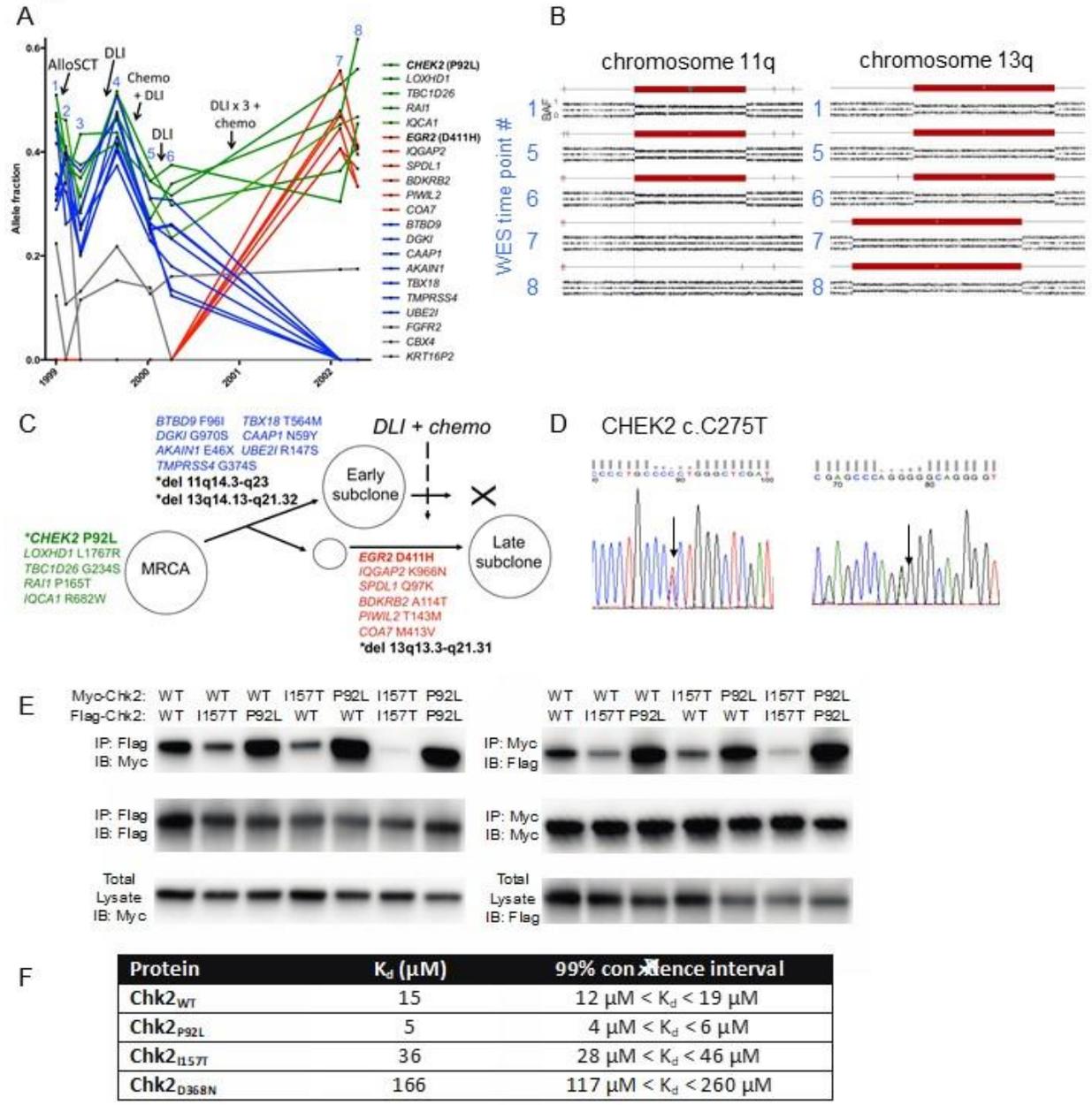


Figure 6. Candidate driver CHEK2 c.C275T (p.92L) affects Chk2 dimerization.

(A) AF plot for CLL patient 1. Recurrent, somatic exonic nonsynonymous mutations are plotted over time across 8 time points. Samples from time points 2, 3, 5, and 6 were unsorted, which accounts for the lower AF of mutations in these samples. Clonal mutations (green), mutations enriched after therapy (red), and those that diminish after therapy (dark blue) are shown and

CLL drivers are bolded. Interval treatments are listed. (B) Microarray B allele frequency plots for purified CLL samples from Patient 1 along chr11q (left) and chr13q (right) at the 5 longitudinal WES time points indicated in blue. Each data point represents a single SNP and the red bars span the breakpoints of chromosomal deletions. (C) Model of subclonal evolution for CLL patient 1. (D) Sanger sequencing chromatograms of *CHEK2* using patient-derived, reverse-transcribed leukemic mRNA. The heterozygous *CHEK2* c.C275T (p.P92L) mutation is validated within the CLL transcript in both the forward (left) and reverse (right) directions. (E) Blots show increased co-immunoprecipitation of WT and P92L Chk2 with the P92L mutant. HCT116 *CHEK2* ^{-/-} cells were co-transfected with plasmids expressing the indicated Flag-tagged and Myc-tagged forms of Chk2 (WT, I157T, and P92L). Lysates were incubated with resin conjugated to antibodies recognizing either the Flag (IP: Flag) or Myc (IP: Myc) tag. Protein from total lysates and immunoprecipitations was resolved by SDS-PAGE and immunoblotted with Flag (IB: Flag) and Myc (IB: Myc) antibodies. (F) The P92L mutant has a lower K_d than WT Chk2. The K_d 's of recombinant WT, I157T, D368N (kinase dead), and P92L Chk2 purified from *E.coli* were measured by analytical ultracentrifugation.

We sought to determine the consequence of the Chk2 P92L substitution. The FHA domain affected by the Chk2 P92L substitution mediates dimerization and activation of the kinase [81, 82]. A neighboring FHA domain substitution (I157T) is implicated in the Li-Fraumeni syndrome, breast cancer, and other tumors [80, 83], and was found to impair Chk2 dimerization, auto-phosphorylation, and activation of the protein [81, 82]. We hypothesized that the P92L mutation would be similarly disruptive.

To interrogate the P92L substitution, we co-expressed Flag- and Myc-tagged Chk2 proteins in HCT116 cells lacking Chk2 and evaluated the homo- and heterodimerization of mutant (I157T and P92L) and wild-type (WT) Chk2 by co-immunoprecipitation. In agreement with prior studies [81], the I157T mutation impaired both homo- and heterodimerization when

compared to WT Chk2. In contrast, the P92L substitution from CLL patient 1 significantly strengthened both homo- and heterodimerization of the kinase when compared to WT Chk2 (Figure 6E). To confirm this unexpected observation, we measured the dissociation constants (K_d) of purified recombinant WT and mutant Chk2 proteins by analytical ultracentrifugation. This experiment validated our co-immunoprecipitation results as the K_d for Chk2 P92L was 5 μ M compared to 15 μ M for WT Chk2 and 36 μ M for Chk2 I157T (Figure 6F and Figure 7). Taken together, these data suggest that the *CHEK2* candidate driver mutation (c.C275T, p.P92L) may prevent Chk2 dimers from dissociating and phosphorylating downstream substrates. Moreover, through an in-depth longitudinal study of late-stage, post-transplant leukemia in a single patient, we identified Chk2 as a candidate CLL driver, a finding that required hundreds of patients using a statistical approach [17].

Figure 7

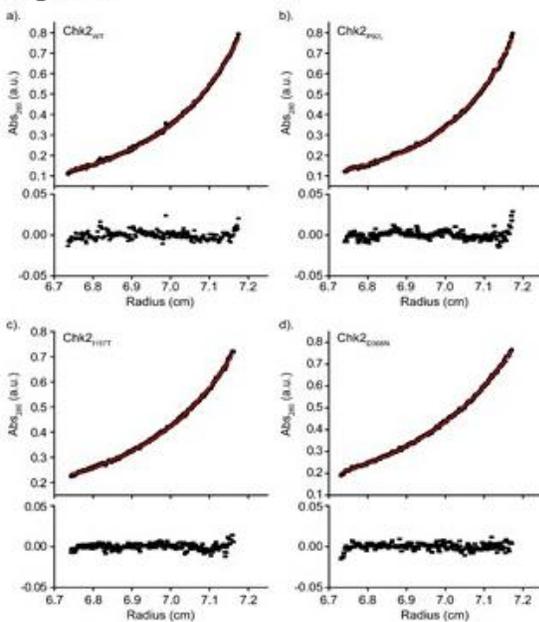


Figure 7. Chk2 P92L mutants form tighter homodimers

The K_d 's of recombinant WT, P92L, I157T, and D368N (kinase dead) Chk2 were measured by analytical ultracentrifugation. The non-linear regression fitting is shown for the Chk2 proteins in A-D. SEQ profiles were fitted to a monomer-dimer equilibrium model. An F-statistics error mapping approach was used to determine the 99% confidence intervals for the dissociation constant.

Chapter 4: Mutation load and neoantigen burden do not predict CLL response to alloSCT

Background: Our study of CLL disease evolution required patients who were nonresponders to alloSCT and who had sufficient post-alloSCT disease for longitudinal sequencing analyses.

Next, we included CLL patients who had a complete response (CR) to alloSCT to assess whether molecular disease features could predict transplant response. Pre-alloSCT CLL from responders was sort-purified and subjected to WES and the pre-transplant samples from the disparate response cohorts were compared.

Methods: *Immunoglobulin heavy chain variable gene (IGHV) mutation status* The somatic mutation status at the *IGHV* locus was determined by the MDACC clinical molecular diagnostics lab. For patients without documentation of the CLL mutation status and with adequate sample available, testing was performed according to an established protocol [13, 84]. Briefly, RNA was extracted from sort-purified CLL cells using the RNeasy Mini Kit (Qiagen), reverse transcribed, and multiplex PCR amplification of IGH transcripts was performed using consensus variable and segment primers. The presence of $\geq 2\%$ variation in the sequenced V-segment of clonal IGH sequences from wild type sequences was considered positive for somatic hypermutation.

Neoantigen prediction To predict potential neoantigens for each patient, a peptide list was generated from each missense mutation that was identified from the exome sequencing data at each sample time point. Peptides included all possible 9- and 10-mer peptides containing the alternate amino acid that resulted from the missense mutations [85]. The binding affinities for each wild type and mutant peptide to the patient-specific HLA molecules (HLA-A and HLA-B) were then tabulated using NetMHCpan (v.2.8) [86]. Peptide-HLA complexes with IC50 values less than 150 nM were considered strong binding neoepitopes and those with IC50 between 150 nM and 500 nM were considered intermediate binding neoepitopes. *Subclonal analysis* The Sequenza package v3.0 [87] was used on the paired tumor-normal BAM (Binary Alignment

Map) files to estimate the global parameters of cellularity and ploidy as well as allele-specific CNA profiles. The reads with low-quality mapping were excluded with the parameter *-q30*. Randomly selected samples from previously published FCR refractory/remission CLL WES data sets [17] were downloaded from dbGaP following institutional approval and converted to BAM format using the SRA Toolkit from NCBI. Statistical analysis Statistical analysis was performed using GraphPad Prism software. FFS post-alloSCT was defined as the number of days from transplant to re-treatment. Data were censored at the last date of MDACC follow-up. The Kaplan-Meier method was used to generate OS and failure-free survival (FFS) curves. Differences between groups were assessed using the log-rank test. The Mann-Whitney test was used to compare differences in mutation and copy number data between groups. Categorical variables were compared using Fisher's exact test. The Wilcoxon rank sum test was used to compare pre-treatment CLL samples from the alloSCT and chemotherapy responder/nonresponder cohorts analyzed using Sequenza v3.0. For the pre/post FCR samples from the Landau *et al.* data sets [17], paired Wilcoxon signed rank test was used. Two-tailed *P* values were calculated and *P* values of less than .05 were considered statistically significant.

Results:

CLL alloSCT patients

To identify molecular predictors of transplant response, we added 13 CLL patients to the cohort who experienced a CR to alloSCT (patients 12-24, Tables 1 and 2). Of these 13 patients, 11 had a durable CR (> 2 years) with a median post-transplant OS of 9.8 years. CLL patients 22 and 23 had a CR to alloSCT and neither patient experienced disease relapse; however, both patients died within 1.5 years of alloSCT from steroid refractory GVHD and so they were not considered to have had a durable CR.

Overall, the cohort was relatively young at the time of CLL diagnosis (median age 53), likely because younger patients tend to have more aggressive disease [88] and are more fit for transplant. All patients demonstrated FCR-refractory disease prior to alloSCT, and all but 1 of

the 24 patients received nonmyeloablative or reduced-intensity conditioning (Table 2). Notably, all 21 patients that were tested had CLL with an unmutated immunoglobulin heavy-chain variable region (*IGHV*) gene, which is among the strongest adverse risk factors in CLL (3 patients had insufficient sample) [89]. We utilized WES to an average depth of 120X to detect exonic sSNVs in the leukemia. We detected a mean of 41.2 (SD \pm 15) sSNVs (exonic silent, non-silent, and sIndels) per case (Figure 8A).

The patients in our cohort had leukemia enriched with variants in recognized CLL drivers (Figure 8B) [17, 63]. A median of 3 drivers was observed per case (range, 2-9; 88% clonal, 12% subclonal), and 10 patients harbored 5 or more drivers in their pre-alloSCT leukemia. Moreover, the prevalence of CLL drivers that adversely impact OS was increased in our cohort [89-92] compared to the 501 untreated patients from Landau et al. [17], including 10 patients with *TP53* mutations and 6 patients with loss of 17p (both $P < .01$ by Fisher's exact test).

We compared several metrics between the 11 patients with a durable CR and the 11 patients with an early relapse (within 2 years) post-transplant. We did not detect any significant differences between groups with respect to the number/composition of exonic mutations, copy number changes, or the leukemic neoantigen burden (Figure 9). Importantly, the inclusion/exclusion of data from CLL patients 22 and 23, who had a CR but died within 1.5 years of alloSCT due to severe GVHD, did not affect the results. In summary, we observed aggressive molecular features in our cohort of 24 patients with chemorefractory CLL; however, they were not predictive of transplant outcome.

Figure 8

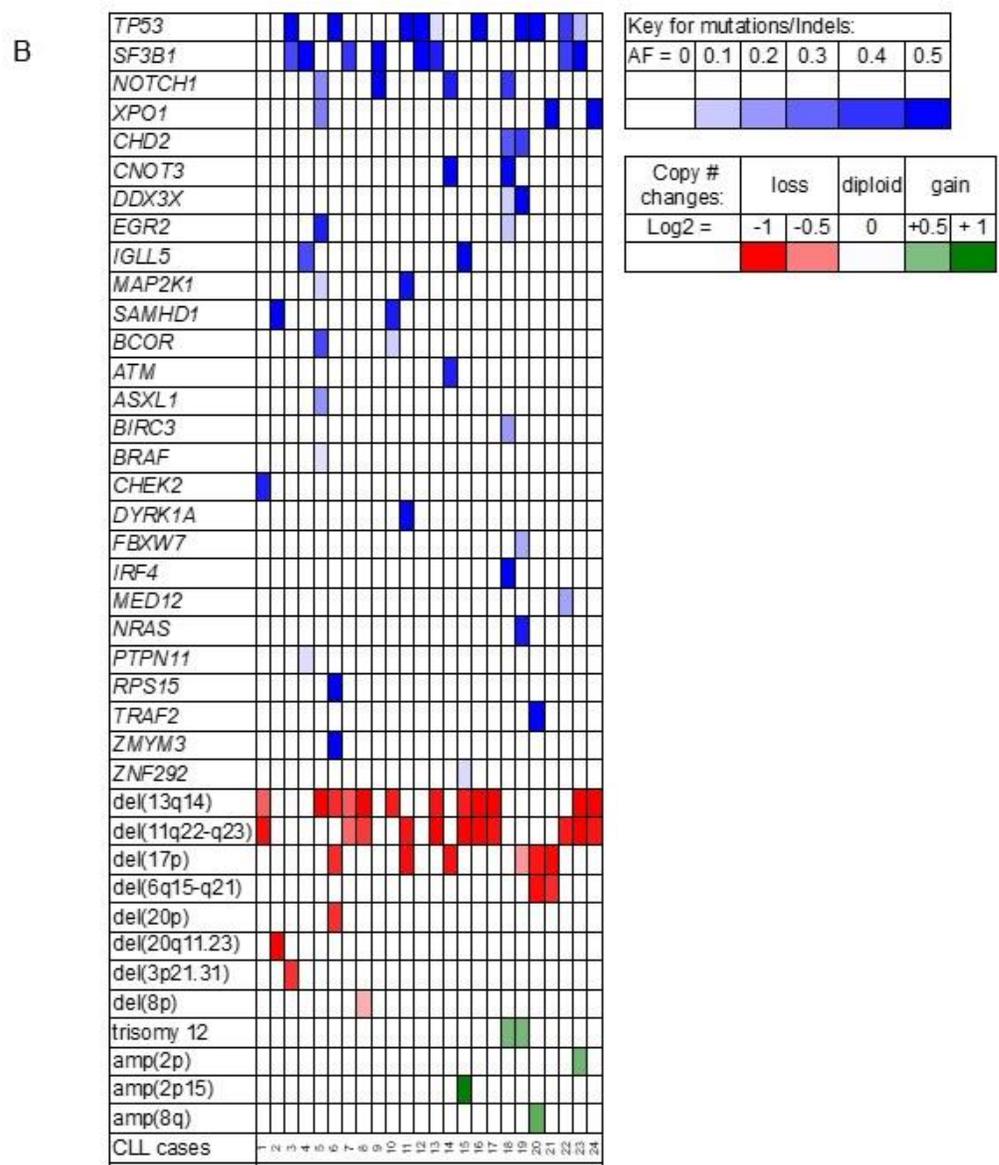
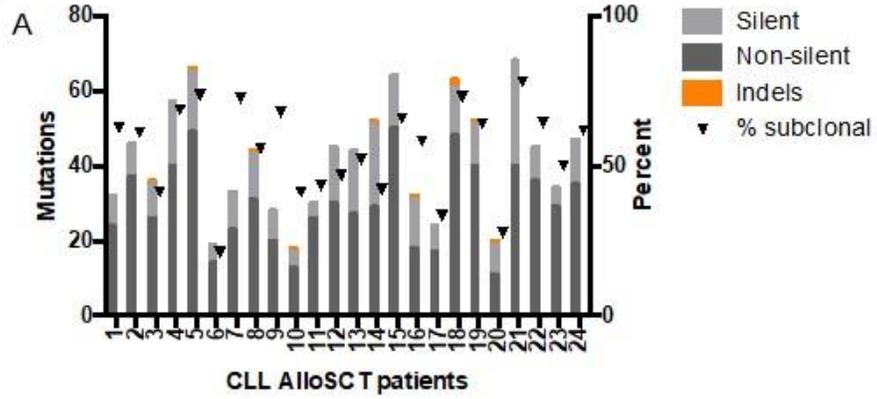


Figure 8. Somatic variants detected in the alloSCT patient cohort (n=24)

(A) Total numbers of somatic exonic silent and non-silent SNVs and indels are listed for each patient's pre-alloSCT leukemia (allelic fraction (AF) > 0.05). The percentage of mutations that are subclonal is also indicated (triangle). (B) The landscape of somatic variants in genes and chromosome regions recognized to be CLL drivers is shown for the CLL alloSCT cohort.

Variants are shaded according to their corresponding variant AF or exome-derived copy number log₂ ratio.

Figure 9. Mutation load and neoantigen burden do not predict response to alloSCT

(A) Post-transplant OS for patients in our cohort with a durable CR (>2 years) or early relapse.

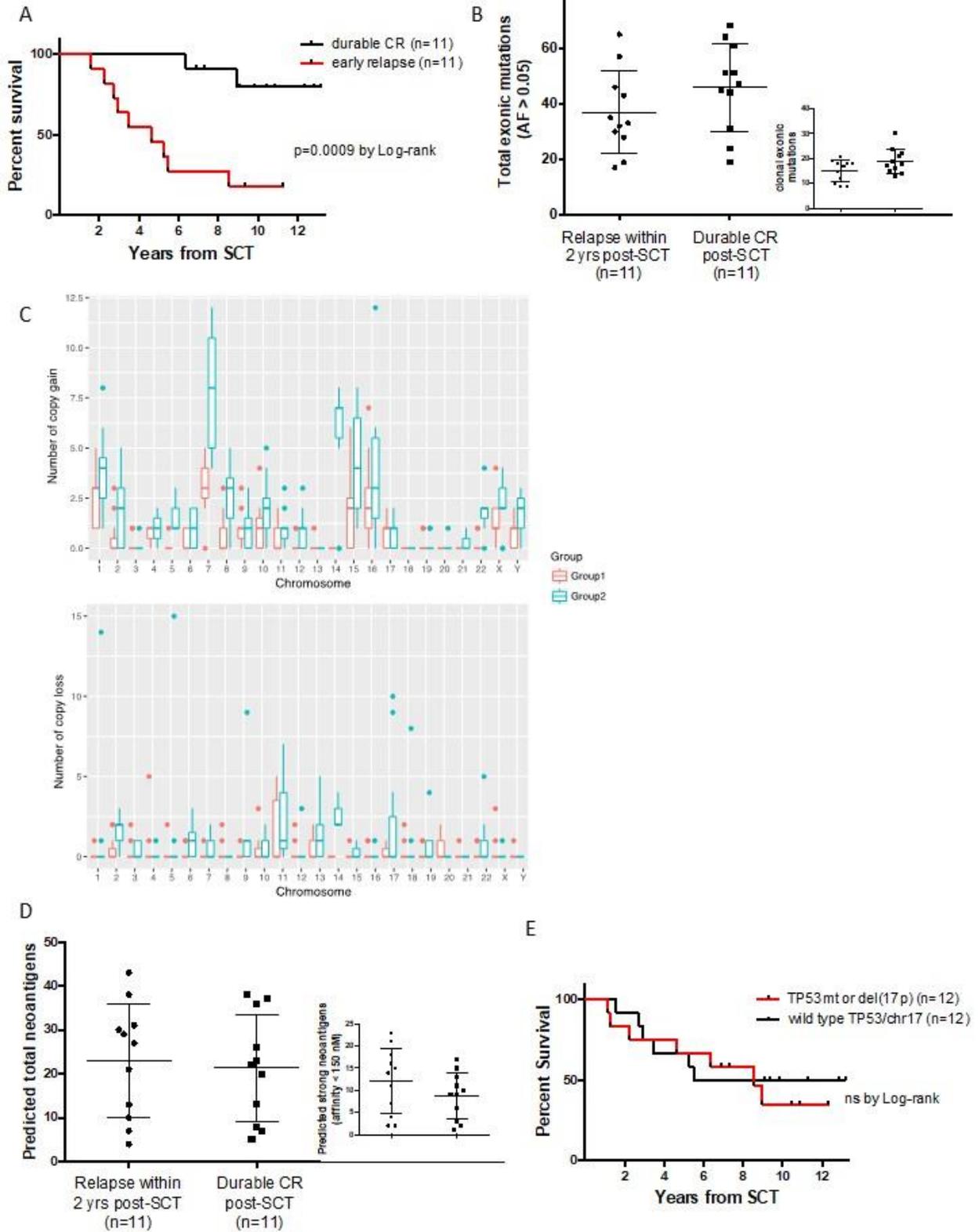
(B) Comparison of total exonic mutations and total clonal exonic mutations (inset) between the transplant early relapse and durable CR groups. Bars indicate mean +/- standard deviation. (C)

Comparison of copy losses and gains between the durable CR and early relapse groups. The Mann-Whitney test was used to compare differences. Copy number changes were similar between groups except at the T-cell receptor and immunoglobulin loci where copy number anomalies were probable artifacts from using sorted T cells as the normal reference. (D)

Comparison of the total and strong binding (inset) neoantigen burden between CLL transplant early relapse and durable responder groups. (E) Post-transplant OS for patients with either wild

type *TP53*/chr17p or those with mutant *TP53*/17p loss.

Figure 9



Structural heterogeneity differs between CLL alloSCT responders and nonresponders

Next, we used the WES data to evaluate tumor heterogeneity contributed by larger chromosomal aberrations such as CNAs and loss of heterozygosity (LOH) using the Sequenza algorithm [87]. Cellularity, as defined in Sequenza, is the fraction of tumor cells in a patient sample. Sequenza derives the cellularity parameter from WES data using two sources: normal cells (for germline analysis) and the tumor sample, which is assumed to be a mixture of normal and tumor cells. The program determines copy number variation and allele fraction at SNP sites across the exome and compares tumor and normal samples segment by segment to find the most likely cellular fraction (with values ranging between 0 and 1, where a value of 1 indicates a purely clonal tumor sample). Somatic mutated alleles are not used in the calculation. Since pre-transplant sorted leukemia cells were used for the analysis, we postulated that when cellularity was close to one, the tumor sample was pure according to copy number and SNP heterogeneity. Accordingly, when the cellularity was lower than one, it implied the tumor sample was composed of at least two structural clones.

While there were not significant differences in total CNAs between the CR and NR cohorts (Figure 9C), we did observe a significant difference in structural heterogeneity between the response groups. Specifically, the CR cohort had CLL that was more structurally clonal and AF change in the responder cohort was due to homogenous copy gain/loss within the CLL population ($P=0.003917$, Wilcoxon rank sum test). In the NR cohort, AF change was contributed by subclonal CNAs rather than clonal CNAs, resulting in increased subclonal structural heterogeneity (Figure 10A and B). These data suggest that structurally heterogeneous CLL may be more resistant to alloSCT.

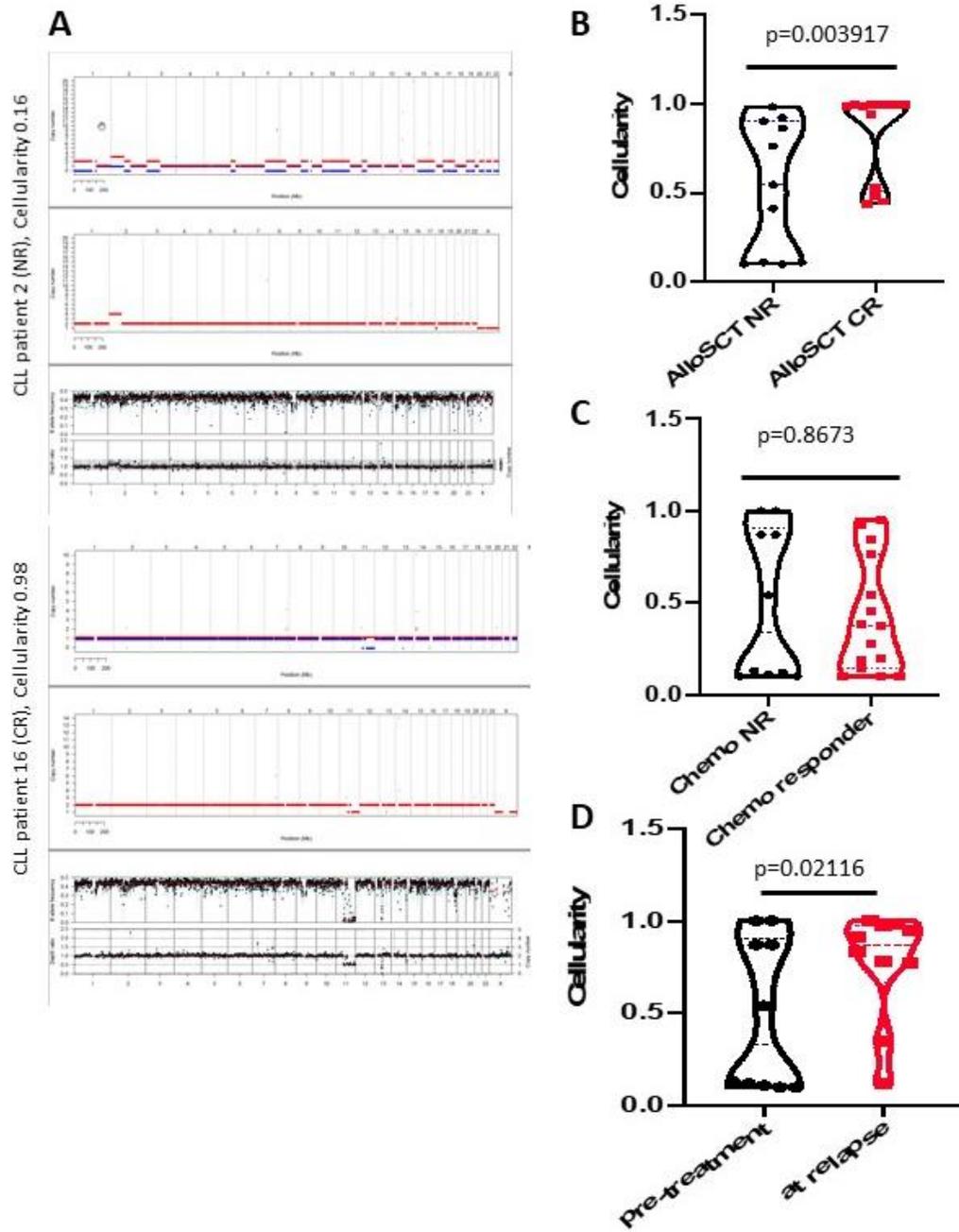
To understand whether the differences in structural heterogeneity were limited to the alloSCT setting, we performed similar analyses on CLL samples collected before and after chemotherapy using previously published data sets [17]. We evaluated tumor 'cellularity' of pretreatment samples from 10 randomly selected relapsed patients and compared them to

pretreatment samples from 15 patients that went into remission following FCR chemotherapy. There were no statistically significant differences between these two response groups in the Sequenza analysis ($P=.8673$, Wilcoxon rank sum test) (Figure 10C). Within the FCR relapse cohort, the cellularity of purified CLL was higher for post-FCR samples compared to pre-FCR paired samples ($P=.02116$, Wilcoxon's signed rank test, paired) (Figure 10D). These data suggest that chemotherapy can successfully eliminate some, but not all leukemia subclones, which may drive the tumor cells to a more clonal, chemorefractory population.

Figure 10. Structural heterogeneity of pre-transplant CLL differs between alloSCT response groups

(A) B allele frequencies of non-synonymous SNPs generated from WES data by Sequenza v3.0. CD19+CD5+ leukemia cells from pre-transplant samples from CLL patients 2 and 16 were analyzed. CLL patient 2 sample has high structural heterogeneity while the pre-transplant sample from CLL patient 16 is almost entirely clonal. (B) Comparison in structural heterogeneity between the nonresponder and responder cohorts, $P=.003917$ by Wilcoxon rank sum test. (C) Cellularity of sort-purified CLL cells from patients that went into remission ($n=15$; denoted as Chemo responder) following FCR treatment versus those who eventually relapsed ($n=10$; denoted as Chemo NR). Sequenza v3.0 analysis of WES data from pre-treatment samples was performed using data sets that were previously generated [17], $P=.8673$, using Wilcoxon rank sum test. (D). Comparison of tumor cellularity in paired samples from FCR relapsed patients before and after treatment. $P=.02116$ by paired Wilcoxon signed rank test.

Figure 10



Chapter 5: Post-alloSCT T cell repertoires

Background: Having observed significant and frequent post-alloSCT evolution of the leukemia population, we then assessed for reciprocal changes in the allogeneic T cell compartment. Delayed and impaired T cell reconstitution post-alloSCT is a well-known complication of the procedure that can lead to severe infection in the transplant recipient. T cell reconstitution post-alloSCT is influenced by many factors including conditioning type, graft type (related vs. unrelated vs. umbilical cord), infection, CMV/EBV reactivation, GVHD, recipient/donor age, and the post-alloSCT immunosuppression utilized to prevent/treat GVHD [93-95]. In some transplant recipients, defects in the T cell repertoire can persist for years. We assessed several features of the bulk allogeneic T cell repertoires in the responders and nonresponders from the CLL cohort, including T-cell receptor (TCR) clonality. We also utilized the post-alloSCT, longitudinal T cell repertoires of the CLL transplant recipients to track candidate anti-CLL T-cell clones identified through the techniques discussed in Chapter 6.

Methods: *Bulk TCR Sequencing* CD3+, CD4+, and CD8+ cells from post alloSCT samples were purified using the methodology described in the 'Cell purification' section (Chapter 2). Fractionation of CD3+ T cells was performed when sample quantity was sufficient. Genomic DNA was extracted from sorted cells using established protocols (QIAamp Blood mini kit, Qiagen). TCR sequencing was performed using the immunoSEQ hs TCRB kit from Adaptive Biotechnologies and sequencing was performed on a MiSeq System (Illumina) at the Sequencing and Microarray Facility at MDACC. A subset of the samples was shipped to Adaptive Biotechnologies for deep bulk TCR sequencing (hs TCRb $\nu\beta 3$ assay). The Adaptive Biotechnologies assay utilizes multiplex PCR with primers that anneal to the V and J segments, resulting in amplification of rearranged VDJ segments from each cell. T cell repertoires were analyzed using the immunoSeq analyzer software available through Adaptive Biotechnologies and using the tcR and ComplexHeatmap R packages [96]. Sample overlap was assessed using

two methods (i) scatter plot of clonotype abundance in each sample analyzed using the Pearson coefficient (ii) heatmap of the Morisita-Horn similarity index [97]. The Morisita-Horn index is a population overlap measure that compares the presence and frequency of T cells in 2 repertoires; values approaching 1 indicate highly correlated repertoires while values near 0 indicate that the 2 samples share very few T cells with the same TCR. Shannon Clonality is 1-normalized entropy and is also known as 1-Pielou's Evenness Index (see equation below) [98]. Simpson Clonality is the square root of Simpson's D. Simpson's D is the sum over all observed productive rearrangements of the square fractional abundances of each rearrangement (see equation below) [99].

$$\text{SIMPSON CLONALITY} = \sqrt{\sum_{i=1}^R P_i^2} \quad \text{SHANNON CLONALITY} = 1 - \frac{-\sum_{i=1}^R P_i \log_2 [P_i]}{\log_2 R}$$

Where R = the total number of productive T cell rearrangements, i = each rearrangement, and Pi = the productive frequency of rearrangement i.

Both the Shannon Clonality and the Simpson Clonality range from 0 to 1, where 0 represents a completely even sample and 1 represents a monoclonal sample.

Results:

We studied the post-alloSCT T cell repertoires in 19 of the 24 patients from the CLL WES cohort (Table 3). The remaining 5 patients had insufficient material for analysis. For many patients, serial samples were available that spanned several years. The multiplex PCR and NGS approach (Adaptive Biotechnologies) can replicate the historical method of assessing T cell repertoire diversity by CDR3 length (CDR3 spectratyping) but also permits deeper analyses since each T cell clone is resolved to its exact TCR Vbeta sequence. Importantly, the assay dedicates the highest sequencing fidelity to the hypervariable CDR3 segment (Figure 11).

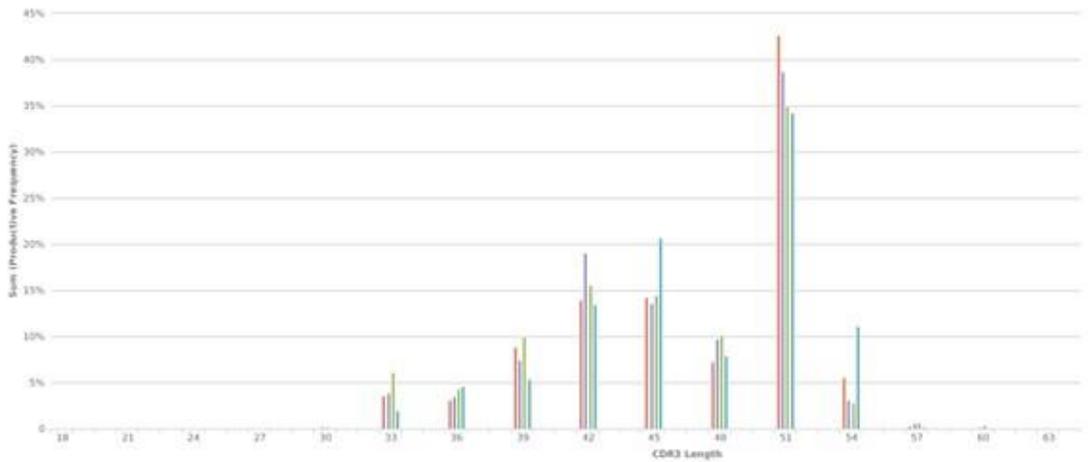
Table 3. Post-alloSCT T cell samples and TCR V β repertoire diversity

Sample ID	Time point (mths from alloSCT)	Sample Type	T cell subset	Total templates	Fraction productive	Productive rearrangements	Simpson clonality	Shannon clonality	Max productive frequency
01 CLL1	36.8	PB	CD3	110994	0.80	14020	0.1683	0.4223	0.10
02 CLL1	39.1	PB	CD3	58797	0.81	8737	0.1704	0.3966	0.10
03 CLL1	11.9	BM	unsorted	31431	0.84	11039	0.0779	0.1844	0.04
04 CLL1	14.7	PB	unsorted	78261	0.84	12492	0.1144	0.3237	0.05
05 CLL3	1.2	PB	CD3	21541	0.79	6150	0.1214	0.2183	0.10
06 CLL3	11.0	BM	CD3	31518	0.88	12776	0.1817	0.2387	0.16
07 CLL3	14.3	BM	CD3	30056	0.84	9985	0.0755	0.1564	0.06
08 CLL3	16.4	BM	CD3	12162	0.85	5241	0.1046	0.145	0.09
09 CLL4	8.5	BM	CD8	28586	0.61	682	0.44	0.6383	0.30
10 CLL4	8.5	BM	CD4	7307	0.84	3063	0.104	0.1416	0.07
11 CLL4	8.5	PB	CD8	67274	0.52	1325	0.5049	0.7034	0.37
12 CLL4	8.5	PB	CD4	1820	0.82	1101	0.058	0.0487	0.03
13 CLL4	11.2	BM	CD8	5493	0.52	242	0.4896	0.5906	0.40
14 CLL4	11.2	BM	CD4	4562	0.82	2063	0.1098	0.1431	0.07
15 CLL4	11.2	PB	CD8	47615	0.55	999	0.5305	0.7168	0.39
16 CLL4	11.2	PB	CD4	8963	0.82	4841	0.0745	0.0999	0.05
17 CLL4	16.1	BM	CD8	17811	0.78	1902	0.2303	0.4297	0.14
18 CLL4	16.1	BM	CD4	13691	0.81	7461	0.0636	0.09	0.04
19 CLL4	16.0	PB	CD8	2012	0.77	302	0.2442	0.329	0.14
20 CLL4	16.1	PB	CD4	51311	0.82	22745	0.0734	0.1447	0.04
21 CLL5	2.3	PB	CD3	73512	0.80	30172	0.0219	0.0693	0.01
22 CLL5	7.3	BM	CD3	79008	0.80	19922	0.0673	0.1745	0.05
23 CLL5	16.1	BM	CD3	36534	0.78	14709	0.043	0.1017	0.02
24 CLL5	35.2	BM	CD3	46527	0.78	20923	0.0409	0.0791	0.03
25 CLL5	35.2	PB	CD3	63888	0.79	36745	0.0182	0.0372	0.01
26 CLL5	35.2	PB	CD8	66445	0.80	31445	0.042	0.0921	0.03
27 CLL5	37.3	PB	CD8	65816	0.80	34060	0.0352	0.0727	0.02
28 CLL5	40.7	PB	CD8	59008	0.79	29366	0.0317	0.0733	0.02
29 CLL5	42.7	PB	CD8	39835	0.79	23423	0.0232	0.0409	0.01
30 CLL5	49.5	PB	CD8	58381	0.79	32270	0.0235	0.0492	0.01
31 CLL5	54.8	BM	CD3	24428	0.79	13555	0.0229	0.038	0.01
32 CLL6	3.0	BM	CD3	2789	0.80	1600	0.0738	0.0744	0.05
33 CLL6	34.7	PB	CD3	807	0.72	352	0.0932	0.0764	0.03
34 CLL6	43.4	BM	CD3	11357	0.82	4388	0.0718	0.1412	0.04
35 CLL7	17.1	BM	CD3	11866	0.78	870	0.1926	0.3488	0.11
36 CLL7	22.5	PB	CD3	68487	0.79	15090	0.0905	0.2399	0.05
37 CLL7	22.5	PB	CD3	75041	0.79	15857	0.0936	0.2498	0.05
38 CLL7	22.5	PB	CD8	43009	0.78	2330	0.1773	0.3991	0.10
39 CLL7	22.5	PB	CD8	41388	0.78	2444	0.1711	0.3932	0.09
40 CLL7	22.5	PB	CD4	132964	0.79	38533	0.0315	0.1216	0.01
41 CLL8	1.0	BM	CD3	115610	0.76	18681	0.119	0.2954	0.08
42 CLL8	1.0	BM	CD8	54610	0.69	2884	0.1713	0.3751	0.10
43 CLL8	5.8	BM	CD3	92314	0.78	25662	0.0931	0.2151	0.05
44 CLL8	15.2	BM	CD3	33422	0.79	10892	0.0949	0.2171	0.04
45 CLL8	27.4	BM	CD3	63231	0.84	7608	0.152	0.3582	0.08
46 CLL8	27.4	PB	CD8	60243	0.88	6359	0.2288	0.4312	0.16

47 CLL9	12.9	BM	CD8	83603	0.94	1394	0.2795	0.5286	0.16
48 CLL9	14.3	BM	CD8	3069	0.81	544	0.1174	0.1877	0.05
49 CLL10	48.8	BM	unsorted	124	0.24	27	0.216	0.0241	0.13
50 CLL10	58.2	PB	CD3	48444	0.81	15397	0.0509	0.1446	0.03
51 CLL10	58.3	BM	CD3	7689	0.83	2514	0.0698	0.1331	0.04
52 CLL10	68.9	PB	CD3	1482	0.78	625	0.0745	0.0762	0.03
53 CLL10	80.7	PB	CD3	244	0.50	110	0.0999	0.0083	0.02
54 CLL12	1.6	BM	CD8	2081	0.62	302	0.3159	0.324	0.29
55 CLL12	2.5	PB	CD8	4279	0.73	515	0.2992	0.3763	0.23
56 CLL12	2.5	PB	unsorted	48974	0.82	2935	0.2361	0.4266	0.14
57 CLL12	6.1	PB	CD8	7761	0.72	852	0.1728	0.2975	0.10
58 CLL12	9.6	BM	CD8	6079	0.93	343	0.2326	0.3995	0.11
59 CLL13	1.1	BM	CD3	857	0.77	533	0.0708	0.0347	0.05
60 CLL13	28.5	PB	unsorted	17734	0.82	7227	0.0791	0.1386	0.05
61 CLL13	17.6	PB	CD4	9211	0.84	6192	0.0211	0.0203	0.01
62 CLL13	17.6	PB	CD8	3143	0.84	1747	0.1336	0.1177	0.13
63 CLL14	5.7	PB	CD3	100028	0.86	13261	0.2848	0.4875	0.26
64 CLL14	5.7	PB	CD3	97003	0.84	13469	0.2695	0.4758	0.24
65 CLL14	8.4	PB	CD4	23551	0.81	14005	0.031	0.042	0.03
66 CLL14	8.4	PB	CD8	12788	0.75	2843	0.2346	0.3513	0.18
67 CLL14	14.7	PB	CD4	63663	0.82	35619	0.0155	0.0379	0.01
68 CLL14	14.7	PB	CD8	14013	0.78	4609	0.1764	0.2767	0.11
69 CLL14	50.8	PB	unsorted	11654	0.79	1365	0.2735	0.4731	0.16
70 CLL15	19.1	PB	CD4	67362	0.81	35633	0.0203	0.0594	0.01
71 CLL15	19.1	PB	CD8	71662	0.92	5185	0.18	0.4465	0.09
72 CLL15	35.9	PB	CD4	23852	0.78	12794	0.0246	0.0548	0.01
73 CLL15	35.9	PB	CD8	39886	0.88	2352	0.2488	0.4978	0.13
74 CLL15	47.7	PB	CD4	31325	0.79	17040	0.0253	0.0581	0.01
75 CLL15	47.7	PB	CD8	40186	0.85	2465	0.2457	0.5023	0.12
76 CLL16	6.5	PB	CD8	63965	0.85	7341	0.1489	0.3761	0.07
77 CLL16	18.0	PB	CD8	54643	0.82	10516	0.1333	0.3429	0.07
78 CLL16	27.6	PB	CD8	25788	0.78	1309	0.274	0.4697	0.22
79 CLL17	5.7	PB	unsorted	33679	0.79	20749	0.0148	0.0254	0.01
80 CLL17	36.8	PB	CD4	22414	0.79	11593	0.0361	0.0656	0.02
81 CLL17	36.8	PB	CD8	2655	0.81	588	0.4271	0.4148	0.41
82 CLL17	62.8	PB	CD4	77195	0.81	39842	0.0141	0.0445	0.01
83 CLL17	62.8	PB	CD8	9389	0.80	1824	0.3483	0.4621	0.23
84 CLL19	1.0	PB	CD3	13283	0.76	1966	0.1194	0.2537	0.05
85 CLL19	25.3	PB	CD4	2346	0.77	1303	0.0639	0.0611	0.04
86 CLL19	25.3	PB	CD4	1901	0.74	1082	0.0593	0.0479	0.04
87 CLL19	25.3	PB	CD8	18820	0.80	2037	0.1713	0.3417	0.13
88 CLL19	57.3	PB	unsorted	22999	0.79	4263	0.1422	0.3	0.09
89 CLL20	n/a	LP	CD3	47300	0.84	14609	0.2108	0.2198	0.21
90 CLL20	23.7	PB	CD8	9826	0.82	5778	0.0219	0.0285	0.01
91 CLL20	55.7	PB	CD8	58367	0.80	25303	0.0279	0.0646	0.02
92 CLL21	7.5	PB	CD8	47305	0.78	9754	0.0704	0.2168	0.03
93 CLL21	22.3	PB	CD8	49712	0.73	4770	0.0941	0.2832	0.04
94 CLL21	37.9	PB	CD8	5921	0.74	1265	0.0905	0.1775	0.03
95 CLL22	1.3	PB	CD4	16601	0.81	9592	0.0256	0.0437	0.02
96 CLL22	1.3	PB	CD8	15172	0.83	4747	0.1616	0.2182	0.15
97 CLL22	2.8	PB	CD8	517	0.63	298	0.061	0.0067	0.01
98 CLL22	8.93M and P _F	unsorted		1070	0.76	601	0.0611	0.0383	0.03

Figure 11

A



B

productive frequency	templa tes	amino acid	rearrangement	v resolved	d resolved	j resolved
0.2554069	21859	CASSQGGKWMNTEAFF	CTGGAGCTTGGTACTCTGCTGTGTAATTTCTGTGCCAGCAGCCAAGGCGGAAAATGGGCCATGAACACTGAAGCTTCTTTGGACAA	TCRBV03	unknown	TCRBJ01-01*01
0.1835476	1756	CASSEVAVILGQPQHF	ACAAAAGCTGGAGGACTCAGCCATGTAATTTCTGTGCCAGCAGTGAAGTGGCAGGGGTGATCTTAGGTCAGCCCCAGCATTITGGTGAT	TCRBV02-01*01	TCRB001-01*01	TCRBJ01-05*01
0.1571754	1449	CASSEVAVILGQPQHF	ACAAAAGCTGGAGGACTCAGCCATGTAATTTCTGTGCCAGCAGTGAAGTGGCAGGGGTGATCTTAGGTCAGCCCCAGCATTITGGTGAT	TCRBV02-01*01	TCRB001-01*01	TCRBJ01-05*01
0.1418809	1308	CASIQGAGPTYEQYF	AGCTCTCTGGAGCTGGGGGACTCAGCTTTGATTTCTGTGCCAGCAGTGAAGTGGCAGGGGTGATCTTAGGTCAGCCCCAGCATTITGGTGAT	TCRBV09-01	TCRB001-01*01	TCRBJ02-07*01
0.1132283	1236	CASSEVAVILGQPQHF	ACAAAAGCTGGAGGACTCAGCCATGTAATTTCTGTGCCAGCAGTGAAGTGGCAGGGGTGATCTTAGGTCAGCCCCAGCATTITGGTGAT	TCRBV02-01*01	TCRB001-01*01	TCRBJ01-05*01
0.1040243	959	CASSQGGKWMNTEAFF	CTGGAGCTTGGTACTCTGCTGTGTAATTTCTGTGCCAGCAGCCAAGGCGGAAAATGGGCCATGAACACTGAAGCTTCTTTGGACAA	TCRBV03	unknown	TCRBJ01-01*01
0.1004031	1096	CASSQGGKWMNTEAFF	CTGGAGCTTGGTACTCTGCTGTGTAATTTCTGTGCCAGCAGCCAAGGCGGAAAATGGGCCATGAACACTGAAGCTTCTTTGGACAA	TCRBV03	unknown	TCRBJ01-01*01
0.1003363	925	CASYTRTGWDRSYGYTF	GCTCCCTCCAAACATCTGTGTAATTTCTGTGCCAGCAGTGAAGTGGCAGGGGTGATCTTAGGTCAGCCCCAGCATTITGGTGAT	TCRBV06	TCRB001-01*01	TCRBJ01-02*01
0.0913557	874	CASSQGGKWMNTEAFF	CTGGAGCTTGGTACTCTGCTGTGTAATTTCTGTGCCAGCAGCCAAGGCGGAAAATGGGCCATGAACACTGAAGCTTCTTTGGACAA	TCRBV03	unknown	TCRBJ01-01*01
0.0870701	833	CASSTMFGGAEAFF	CTGGAGCTTGGTACTCTGCTGTGTAATTTCTGTGCCAGCAGTGAAGTGGCAGGGGTGATCTTAGGTCAGCCCCAGCATTITGGTGAT	TCRBV25-01*01	unknown	TCRBJ01-01*01
0.0626278	5360	CASSTMFGGAEAFF	CTGGAGCTTGGTACTCTGCTGTGTAATTTCTGTGCCAGCAGTGAAGTGGCAGGGGTGATCTTAGGTCAGCCCCAGCATTITGGTGAT	TCRBV25-01*01	unknown	TCRBJ01-01*01

Figure 11. CDR3 length plot for patient CLL 14

(A) Four longitudinal post-alloSCT T cell samples from CLL patient 14 were subjected to TCR sequencing (time points are from 6 (red), 8 (purple), 15 (green), and 51 months (blue) post-transplant). A CDR3 length histogram of the productive TCR rearrangements is displayed using data from the 4 samples and shows a persistently oligoclonal repertoire. CLL patient 14 experienced a CR to alloSCT and has been in remission for > 14 years as of August 2020. His course has been complicated by chronic extensive sclerodermatous GVHD and he was treated with tacrolimus for over a decade. (B) Representative rows from a table displaying the unique clonotypes identified within the T cell repertoire of CLL patient 14. Clonotypes are ranked by the productive frequency within a given sample. The TCR variable beta chain, including the CDR3 segment, is depicted by the amino acid and DNA sequences as well as the V, D, and J gene families. Duplicate sequences indicate that the clone was found in multiple samples.

A clone's unique variable beta chain sequence can be used to follow it across longitudinal samples over time and its expansion/contraction can be quantified. We were particularly interested in the variation within the T cell repertoire between 3 and 24 months post-alloSCT when the GVL effect occurs in responding patients (though there is considerable deviation and it can happen earlier or later). Samples more distant from transplant were useful for monitoring ongoing T cell reconstitution. As mentioned in the 'Background' section, post-alloSCT T cell reconstitution is complex and is affected by multiple variables. At present, it is not possible to examine bulk T cells over time and link an observed pattern of variation in the repertoire to a specific entity such as GVHD, GVL, or infection (apart from CMV/EBV reactivation [100, 101]). Nonetheless, we did see significant expansion and contraction of very frequent T cell clones post-alloSCT in both responders and nonresponders (Figure 12). The portion of the T cell repertoire occupied over time by the top 20 most frequent clones was far higher in the transplant recipients compared to the healthy donors.

The diversity of a T cell repertoire is a measure of its richness and evenness. Richness relates to the total number of clones present in a sample, which can be used to estimate the richness of the larger repertoire being measured. Evenness describes the degree to which one or a few clones dominate a sample repertoire. Simpson Clonality and Shannon Clonality are metrics that can be used to describe the overall shape of a T cell repertoire. Shannon Clonality, which is derived from Shannon's Entropy, is more sensitive to differences in sample size than Simpson Clonality and we confirmed that with our data (Pearson correlation coefficient for sample size vs. Shannon Clonality $r=0.25$ ($P=.012$) compared to $r=-0.045$ ($P=.7$) for Simpson Clonality). Because our samples were not uniform, we used Simpson Clonality to compare the shape of a T cell repertoire over time and between patients. Samples with very small numbers of T cells (<5000 total T cells) can show a bias towards increased Simpson Clonality (Figure 13A). For this reason, we excluded the 19 samples with less than 5000 productive TCR templates in the clonality analyses.

The cryopreserved BM and PB samples varied in their cell number and viability. For larger samples, we fractionated CD3 T cells into CD4 and CD8 subsets. For samples with low cell number, unfractionated CD3 T cells were sequenced. We assessed the overlap between CD8 and CD3 T cells that were sorted from samples taken at identical time points in several patients and observed a high degree of overlap (Figure 13B). There was no drop out in CD8 clones more frequent than 0.1% of the population within the CD3 bulk samples. Similarly, at some time points, only BM or PB samples were available, and we measured the overlap between BM and PB samples taken from patients at identical time points to address cross-source comparison. Again, sample overlap was high and all frequent T cells were detected in both the BM and PB samples (Figure 13C). CD8 T cells were of particular interest because they are thought to be the primary effectors of the GVL effect [102]. In addition, CD8 T cells were the cells isolated through MHC class I-restricted, neoantigen tetramer-based, single T cell sorting (Chapter 6). Despite this, we did collect data on CD4 T cells and observed a significantly lower Simpson clonality (a more even distribution of T cell clones) within the CD4 subset compared to the CD8 subset, which has been described previously (Figure 13D) [93]. Though the CD4 T cell subset is more diverse post-alloSCT, it lags behind CD8 T cells in terms of absolute cell number and the ratio of CD4:CD8 T cells is reduced and often inverted post-transplant [95].

The clonality of a given T cell sample describes the repertoire diversity at one time point but the repertoire adapts in response to multiple insults (viral infection, vaccination, autoimmunity, aging, etc.). In alloSCT recipients, the T cell repertoire diversity is dramatically restricted initially, but, in general, diversity improves over time owing to thymopoiesis [93]. We measured the clonality of serial T cell samples for the CLL cohort (Figure 14A). The median range in clonality over time per patient was 0.08. As context, in a separate dataset that monitored the peripheral blood T cell repertoire of 3 healthy donors at 8 time points over the course of a year, the average range in Simpson clonality per patient was 0.006 (Figure 12, I-K).

CLL patients 4, 16, and 17 experienced especially large changes in their T cell repertoire diversity. The serial clonality data for CLL patient 4 are intriguing because the initial two T cell samples were drawn in the context of CLL relapse post-alloSCT while the third sample was drawn 3 months after a DLI, which effectively put the CLL into a durable remission. Unfortunately, CLL patient 4 did experience severe chronic GVHD that resulted in his death 5 years post-transplant. CLL patient 16 experienced a CR post-alloSCT and is alive and well without chronic GVHD. However, one possible confounder is that the last serial PBMC sample for CLL patient 16 was drawn 3 months after he suffered a serious motor vehicle accident and multiple orthopedic fractures, requiring surgery and prolonged rehabilitation. Finally, CLL patient 17 remains alive and disease-free, though he experienced extensive chronic skin GVHD requiring systemic immunosuppression for over a decade post-alloSCT.

The median Simpson clonality at time points between 6-24 months post-alloSCT, the window for the GVL effect, was similar between nonresponders and responders (median 0.1 vs. 0.1, Figure 14B). As a comparison, in a separate study the median clonality of the peripheral blood T cell repertoire in 786 healthy donors was 0.02 (IQR 0.01-0.04; $P < .0001$ vs. CLL cohort) [100]. Among patients with treatment responsive disease, we compared the Simpson clonality of the last available time point between patients without chronic GVHD and those with chronic GVHD and, even in this small cohort, a more restricted repertoire was observed in the patients with chronic GVHD (Figure 14C).

Homeostatic peripheral expansion (HPE), a thymic independent process that involves the expansion of mature donor T cells, is the predominant source of recipient T cells in the early post-alloSCT period. Within the first year post-alloSCT, depending on multiple factors (e.g. GVHD, recipient age, viral reactivation, etc.), thymopoiesis begins to generate a new, repertoire of naïve T cells from engrafted hematopoietic stem cells that are educated in the thymus [103]. If thymic output is robust, recipients can eventually build new T cell repertoires that have minimal resemblance to the donor repertoire infused in the original graft [94]. For CLL patient

20, we performed TCR sequencing on an aliquot of the leukapheresis product that composed the original stem cell graft donated by the patient's brother. We then compared the infused donor T cell repertoire to the recipient T cell repertoire at 2 and 3 years post-alloSCT (Figure 15A). Surprisingly, the 'healthy donor' had a very frequent clone that composed 21% of the leukapheresis product and a relatively oligoclonal repertoire (Simpson's Index 0.211). In the recipient at 3 years, however, the very dominant donor clone was present among CD8 T cells at a frequency of only 0.028%, and the recipient had generated a new CD8 T cell repertoire distinct from and more diverse than the original product (Figure 15B).

Consistent with these data, we compared the repertoire overlap of two peripheral blood samples, taken 12 months apart in 3 healthy donors versus 3 complete responder CLL alloSCT recipients (Figure 15C). In the healthy donors, the T cell repertoires at 0 and 12 months showed a high degree of overlap (average Morisita index = 0.88) whereas the overlap in the transplant recipients was intermediate (average Morisita index = 0.73). This supports the dynamism of the post-alloSCT repertoire, especially among the low frequency T cell clones where naïve T cells reside. This can be overlooked when focusing only on the most frequent T cell clones as in Figure 12. Importantly, though, the additional new T cells were not in sufficient quantities to improve the T cell population diversity in these 3 transplant responders and the repertoires remained oligoclonal.

Our results highlight the potential utility of tracking the adaptation of the donor repertoire to the recipient environment over an extended timeframe. One can manipulate and expand candidate GVL T cells from the graft ex vivo; however, there is no substitute for monitoring the expansion/contraction of allogeneic T cells in the host and mining those rich data to identify potential anti-leukemia clones and to uncover signatures of the GVL effect. We hypothesize that donor GVL T cells participate in immunoediting of host leukemia, thereby triggering subclonal evolution. However, our data confirm the GVL T cells are part of a very restricted repertoire that is dramatically different than the input healthy graft.

Figure 12

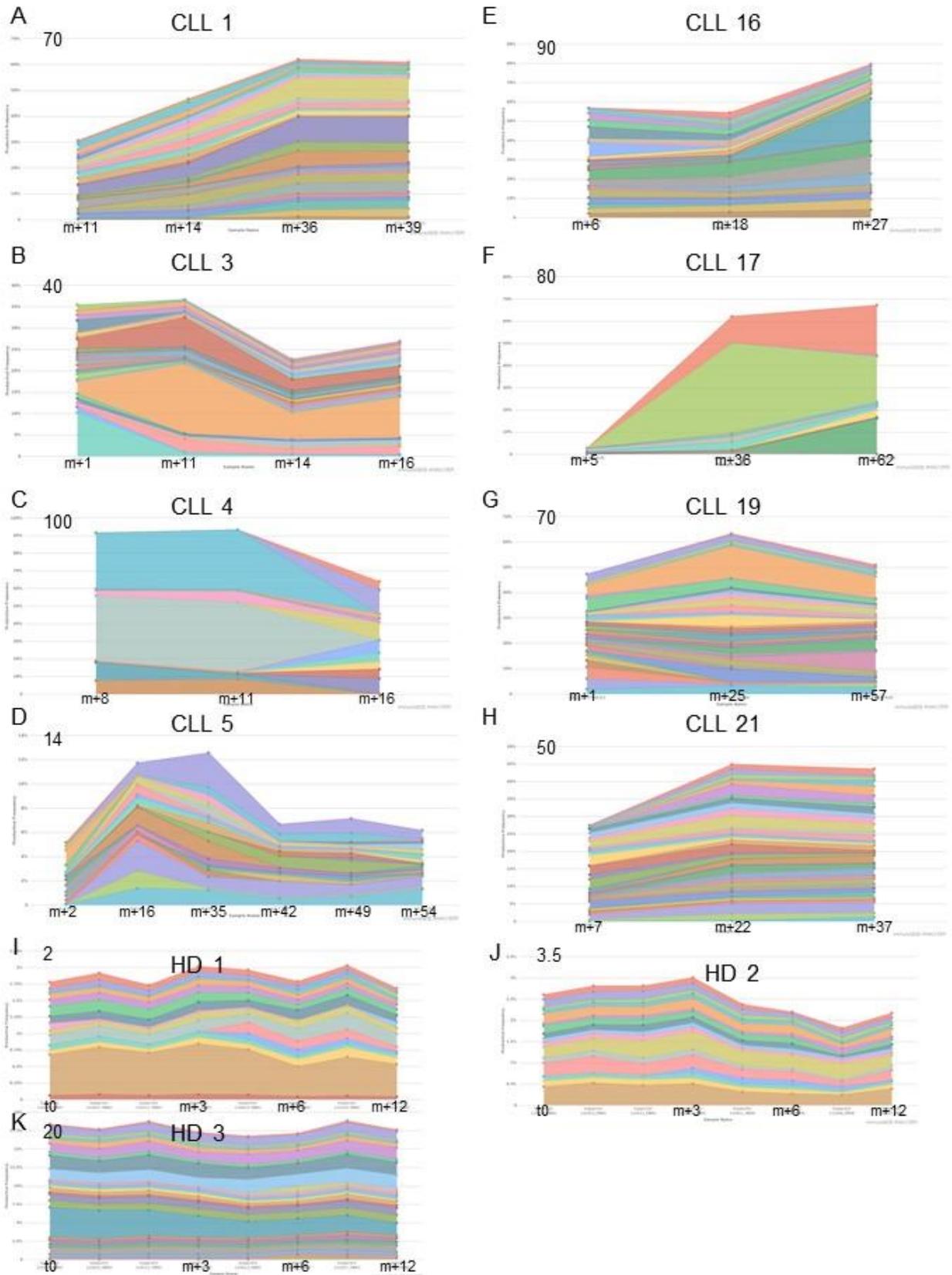


Figure 12. Top 20 T cell clonotypes

The top 20 most frequent T cell clonotypes at a given time point are shown for 8 patients [4 nonresponders (A-D) (though CLL patients 3 and 4 did respond to DLI) and 4 responders (E-H)] and 3 healthy donors (I-K). The Y axis shows the productive frequency of the clonotypes as a %age of the total T cell population and is not standardized across patients since the 20 clonotypes make up a different proportion of the total population in each patient (the top of the Y axis is labeled). The X-axis represents time though the intervals between samples are not to scale. Time points for each patient (in months since alloSCT) are as follows: CLL 1 (11, 14, 36, 39 months), CLL 3 (1, 11, 14, 16 months), CLL 4 (8, 11, 16 months), CLL 5 (2, 16, 35, 42, 49, 54 months), CLL 16 (6, 18, 27 months), CLL 17 (5, 36, 62 months), CLL 19 (1, 25, 57 months), and CLL 21 (7, 22, and 37 months). CLL patients 3, 4, 17, and 21 all had significant chronic GVHD. Longitudinal data for the 3 healthy donors was obtained from the Adaptive immunoSEQ immuneACCESS project 'TCRB Time Course 3 Subjects' dataset. HD1 age 18-24, time points 3/16/11, 4/15/11, 5/12/11, 6/9/11, 8/19/11, 9/15/11, 10/14/11, 3/20/12. HD2 age 18-24, time points 3/17/11, 4/15/11, 5/13/11, 6/9/11, 8/11/11, 9/8/11, 10/6/11, 3/27/12. HD 3 age 24-45, time points 3/16/11, 4/15/11, 5/13/11, 6/9/11, 8/12/11, 9/9/11, 10/7/11, 4/3/12.

Figure 13

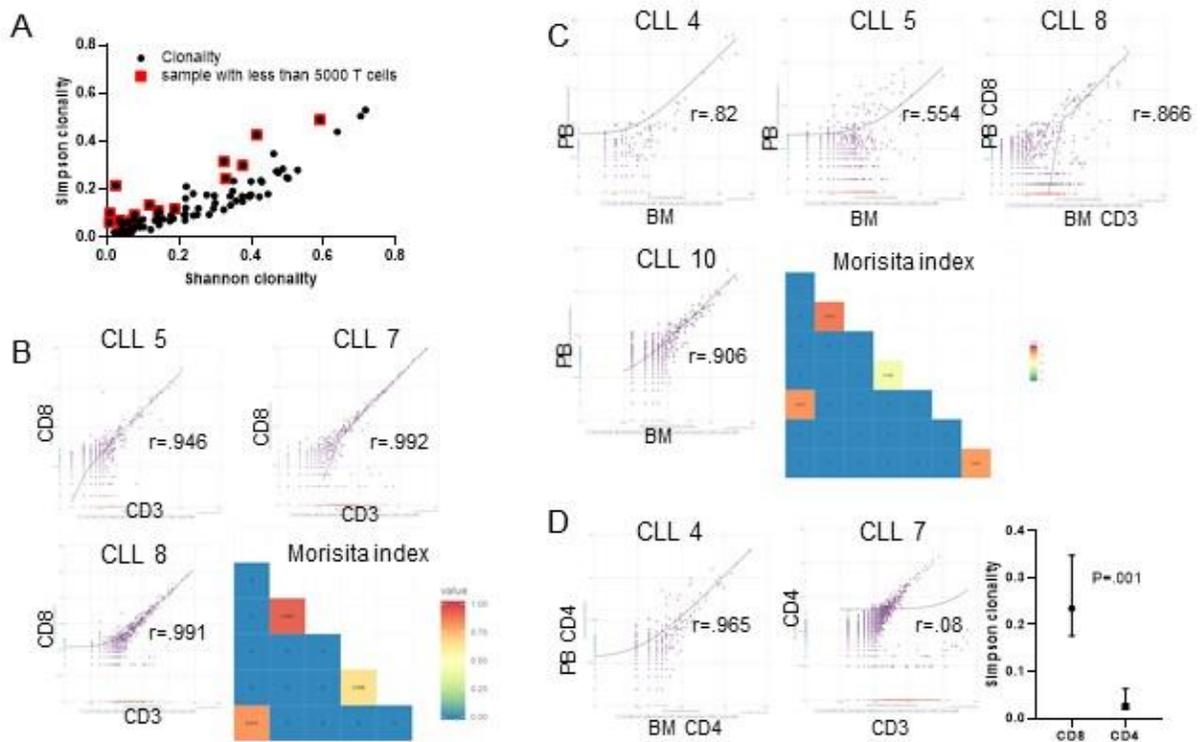


Figure 13. TCR repertoire analysis and reproducibility

(A) Shannon clonality versus Simpson clonality is graphed for all 98 T cell samples that underwent TCR analysis. Samples with < 5000 productive T cell templates are highlighted by red squares. (B) Scatter plots of the productive frequency of individual T cell clones in CD3 vs. CD8 sorted T cells from the same time point in 3 patients. The Pearson correlation coefficient, r , is labeled for each plot. The Morisita index showing the overlap between all 6 possible samples is also shown as a heatmap (red is index = 1 [complete overlap], blue is index = 0 [no overlap]). (C) Scatter plots showing the productive frequency of individual T cell clones in BM vs. PB sorted T cells from the same time point (sorted for the same type of T cell (CD3 or CD8 unless indicated)). (D) Scatter plots showing the productive frequency of individual T cell clones in BM vs. PB (and CD3 vs. CD4) sorted T cells from the same time point. There were 11 time points at which both CD8 and CD4 T cells were sorted from the same sample. The Simpson clonality for CD8 compared to CD4 T cells is shown, $P=0.001$ by Wilcoxon matched-pairs signed rank test.

Figure 14

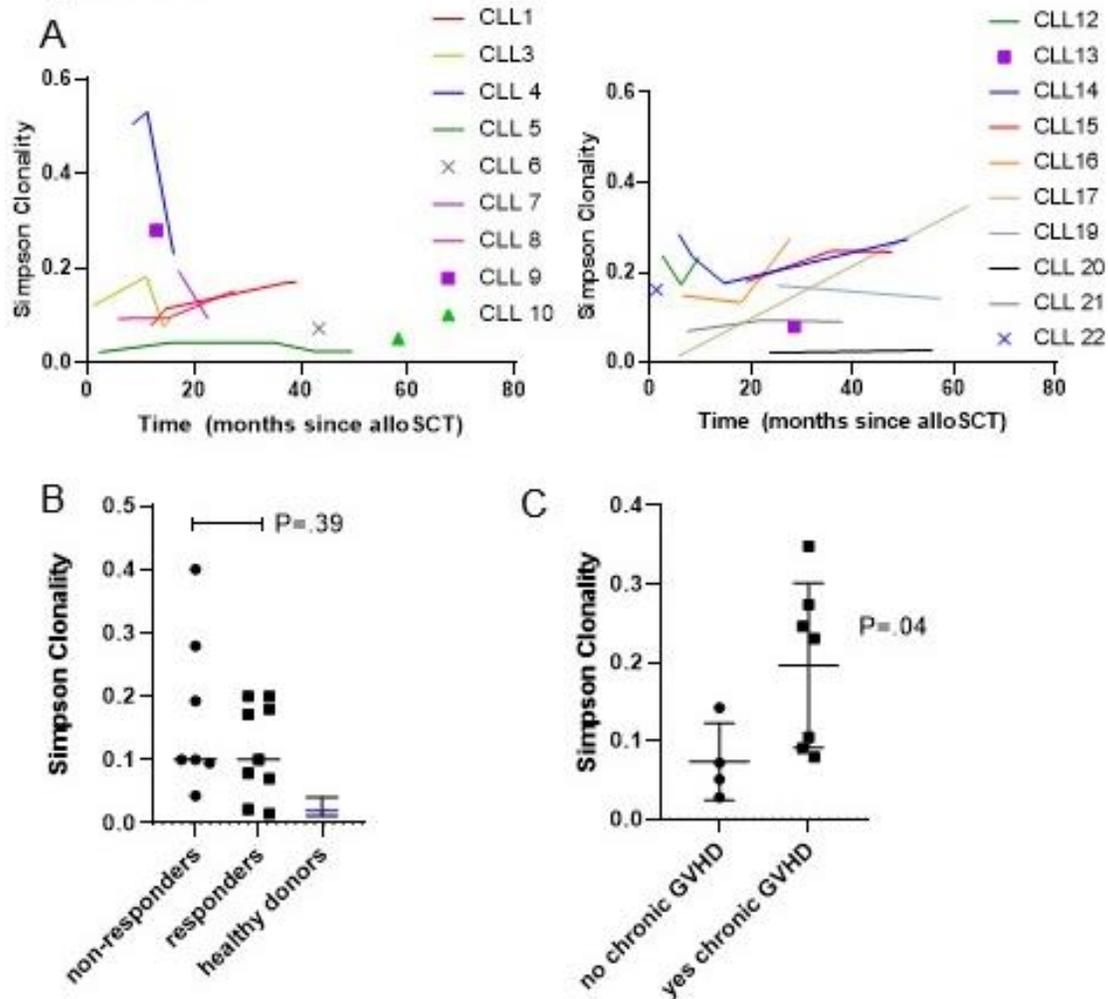


Figure 14. T cell repertoire diversity of post-alloSCT samples

(A) Simpson clonality over time for the alloSCT nonresponders (left) and responders (right). (B) The Simpson clonality at time points between 6-24 months was compared between nonresponders and responders using the Mann-Whitney test (for patients with multiple samples in this time window, the clonalities were averaged and included as one value per patient). The Simpson clonality for 786 healthy donors [100] is listed as median and interquartile range and differed significantly from the Simpson clonality of the grouped CLL cohort ($P < .0001$). (C) The Simpson clonality for patients with treatment responsive disease was compared at the last available time point (median 48 months, range 16-62) between patients without or with chronic GVHD. The Mann-Whitney test was used.

Figure 15

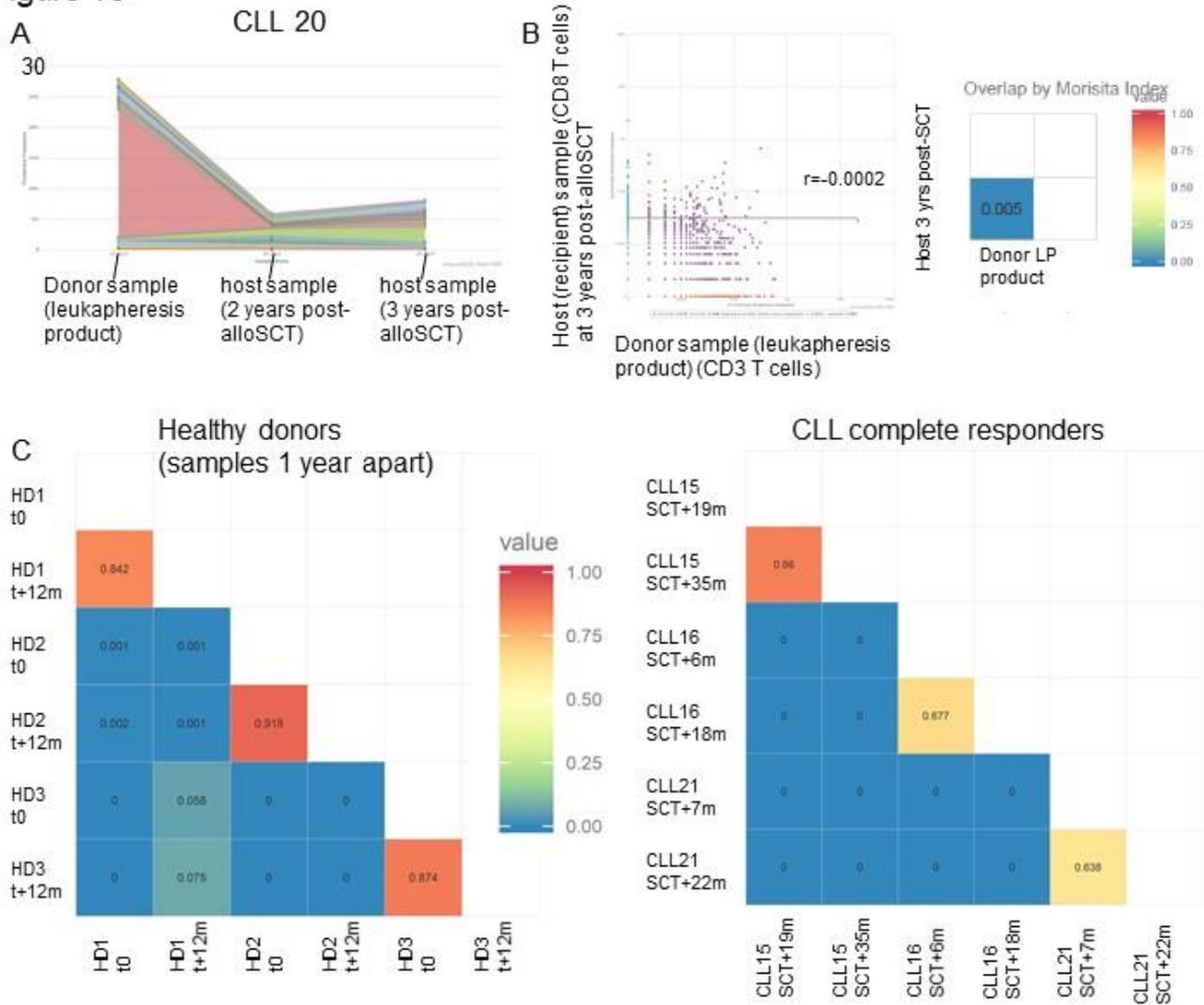


Figure 15. AlloSCT donor and host T cell repertoires

(A) TCR sequencing was conducted on an aliquot of the leukapheresis product that composed the stem cell graft donated by the brother (MRD) of CLL patient 20 and compared to the host repertoire at 2 and 3 years post-alloSCT. (B) Sample overlap between the stem cell graft and the recipient T cell repertoire at 3 years by scatter plot (Pearson correlation coefficient = -0.0002) and the Morisita overlap index. (C) Sample overlap between 2 sample time points, ~12 months apart for 3 healthy donors (left) and 3 CLL transplant complete responders (right).

Chapter 6: Single cell TCR sequencing of candidate GVL clones

Background: The longitudinal T cell repertoires of alloSCT recipients showed shifts in T cell diversity and repertoire composition over time, but, given the multiple variables, the sample size was too small to identify signatures predictive of GVL. Next, we focused on individual T cell clones with potential anti-leukemia activity. Donor T cells eliminate leukemia through the recognition of tumor antigens including leukemia-associated antigens (LAAs), minor histocompatibility antigens (mHAs), and leukemic neoantigens [38]. Our observation that disparate subclones, which harbor the same mHAs yet different molecular lesions, demonstrate differential sensitivity to allogeneic T cell therapy led us to hypothesize that LAAs and/or neoantigens were mediating, at least in part, the leukemic evolution that we observed.

We used the leukemia exome data to predict personalized CLL neoantigens for the CLL cohort patients. We then integrated the longitudinal TCR sequencing data with the single cell TCR sequencing data to investigate candidate GVL clones. For single T cell TCR sequencing, we adapted a technique that utilized tetramer-based single T cell sorting, multiplex nested PCR, and NGS [104]. TCR sequencing is distinct from other types of single cell sequencing because it requires uniquely high fidelity to account for the tremendous variability in the TCR CDR3 region.

Methods (abbreviated and detailed): *Single T cell TCR sequencing* Briefly (see detailed protocol below), CD3+CD8+ tetramer positive cells were sorted into 96 or 384 well plates. RT-PCR for TCR and gene expression of 17 transcripts associated with T cell ontogeny/activation was performed based on a previously described protocol [104]. cDNA was generated using Qiagen One Step RT-PCR kit with combined TCR/phenotyping primers, followed by two rounds of nested PCR to amplify TCR and phenotyping gene sets. Barcoding and paired end addition of Illumina compatible primers were performed to enable deep sequencing. Libraries were pooled, gel purified, and quantified using a Qubit fluorometer. DNA quality was analyzed on a TapeStation 4200. Normalized libraries were sequenced using an Illumina MiSeq (using MiSeq

Reagent kit v3, 600 cycles). Data were deconvoluted using a custom software pipeline that was generously shared by Dr. Jacob Glanville at Stanford University and adapted by Dr. Sahil Seth at MDACC. Briefly, the barcode tags were used to assign reads from a sequencing run to each plate and well. Paired ends were assembled by finding a consensus sequence of at least 100 bases in the middle of the read. The paired TCR V, D, and J segments were then assigned by VDJFasta [105].

T cell immunophenotyping Sorted T cells were stained with the following antibodies: Sytox blue live/dead stain (ThermoFisher), anti-CD3 PE/Cy7 (clone SK7), anti-CD39 PerCPCy5.5 (clone A1), anti-CD69 BV650 (clone FN50), anti-CD107 BV711 (clone H4A3), anti-CD197/CCR7 FITC (clone G043H7), anti-CD8a APC/Cy7 (clone HIT8a) (all from Biolegend), anti-CD103 PE (clone M290) (BD Biosciences) and anti-CD45RA ECD (clone 2H4LDH11LDB9) from Beckman Coulter. Tetramers against mutant PARPBP (for CLL 12) and mutant ACTN1 (for CLL 8) were generated at Baylor College of Medicine, Houston, TX. CD3 positive T cells were sorted on the FACS Aria Fusion (BD Biosciences) after excluding debris, doublets, and dead cells. Single cells were index sorted in 96 or 384 well plates followed by TCR sequencing and gene expression analyses.

Single leukemia cell gene expression analysis for neoantigens Single cells were generated using a Chromium controller and processed to generate gene expression libraries per 10X Genomics guidelines. Gene expression and ATAC-Seq libraries were generated from the pre-transplant sorted cells while only single cell gene expression libraries were generated from the leukemia cells of CLL 8. Standard 10X Genomics Chromium 5' libraries were generated after gel emulsification breakage and synthesis of first strand cDNA. cDNA amplification was followed by cleanup using SPRIselect followed by another round of amplification. The 10X libraries were pooled, normalized to 2nM, and sequenced on NovaSeq 6000 using a S2 kit (300 cycles) with 1% Phi-X spiked in the library. Single nuclei were generated per vendor guidelines for ATAC-Seq. Single cell ATAC-Seq libraries were run on a NovaSeq using an S1 kit (100 cycles). Single cell sequencing data were processed using Cell Ranger (v3.1.0) software from 10X Genomics Inc. with default parameters. Filtered gene counts

from Cell Ranger were analyzed and visualized through the Seurat R package with recommended parameters [106, 107].

Single T cell TCR sequencing (detailed protocol) The complete, adapted protocol for the single T cell TCR sequencing method is included here because it represented an important contribution from the author to the Mollidrem lab and the Jin Im lab. The protocol has since been refined but remains central to the work of the ECLIPSE platform at MDACC.

Single T cell TCR sequencing protocol

Reagents:

- primers - these were ordered from Eurofins Genomics in quantities of ~55 nmol as salt-free primers (list is in Table 4)
 - o Reaction 1:
 - 38 forward Valpha primers
 - 36 forward Vbeta primers
 - TRAC reverse and TRBC reverse primers
 - o Reaction 2: (nested primers and adds common sequence)
 - 36 forward Valpha primers
 - 36 forward Vbeta primers
 - TRAC2 and TRBC2 reverse primers
 - o Reaction 3:
 - Includes forward and reverse barcoding primers as well as Illumina paired end adaptor primers (Pepimer 1 and 2 aka Illumina PE1 and Illumina PE2)
 - o Optional: phenotyping primers
 - o Optional: The sanger sequencing primers used were:
 - Forward: ACACTCTTTCCCTACACGACGCTCTTCCGATCT
 - Reverse: CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
- PCR cyclers: can use either the old Bio-Rad PCR machine or the new Bio-Rad RT-PCR machine
 - o Different plates are used for these machines
 - New RT-PCR cycler: Bio-Rad hard-shell 96-well PCR plates – white shell (HSP9601)
 - When using this cycler, the ramp rate for the step-down to annealing temp **must be reduced** to 0.5 degrees per second or the PCR won't work
 - Old cycler: Bio-Rad hard-shell 96-well high profile semi-skirted PCR plates, clear shell/clear wall (HSS9601)
- Qiagen 1 step RT-PCR mix (210212)
- Qiagen Hotstart Taq DNA (203205)
- Qiagen dNTP Mix (201900)
- Plate covers: Bio-Rad Microseal 'B' seals (MSB1001)
- DNA benchtop 100 bp ladder G8291 promega
- Multichannel pipettes:

- Eppendorf 0.5 – 10 ul multichannel
- Eppendorf 5 – 50 ul multichannel
- Filter pipette tips (Art tips)
- Qiaquick Gel Extraction Kit (28704)
-

Protocol:

Sorting

1. Sort single T cells of interest (tetramer binding or otherwise), 1 cell per well, into the well of a 96-well plate (use the plate appropriate for TCRseq with reaction #1 master mix [20 ul per well] already aliquoted); the volume from a single sort drop is negligible
 - a. Dr. Karen Clise-Dwyer recommends using the 'single cell' sorting mode to increase the chance of getting exactly 1 T cell per well and further advises sorting at a slow speed (500 events/sec)
 - b. The plate with TCRseq master mix already aliquoted should be brought to the sorter (I always used the Aria 1) with the plate seal in place; remove plate seal only immediately prior to sorting
2. For reaction #1, you must decide if you are only assaying for TCR paired alpha/beta chains for each cell or if you are also assaying for the phenotypic parameters; if phenotyping is desired in addition to the paired TCR alpha/beta sequences, then the phenotypic forward and reverse primers must be included in the 20 ul master mix for reaction #1 (see below for example PCR master mixes)

Primer stocks

1. **all** primers were initially reconstituted with water at 100 uM
 - a. 10 uM working concentrations were also prepared for each primer

Alpha and Beta forward primer mixes for Reactions #1 and #2:

1. The 38 Valpha forward primers were combined at the volumes listed below and this mix was stored at -20 in aliquots to be used in the PCR master mixes (same for the 36 Vbeta forward primers for reaction #1 and the 36 Valpha forward/36 Vbeta forward primers for reaction #2)
2. Reaction #1:
 - a. Valpha forward mix:
 - i. 38 primers: JM_P1_001 -> JM_P1_038
 1. Final concentration in PCR reaction: 0.06 uM
 2. 10X = 0.6 uM
 3. For 1000 ul of 10X Valpha forward mix, combine 6 ul of each primer (at 100 uM stock) (6 ul x 38 primers = 228 ul) + 772 ul of water and store at -20 (I often made 4000 ul at a time and stored in 1000 ul aliquots)
 - b. Vbeta forward mix:
 - i. 36 primers: JM_P1_039 -> JM_P1_074
 1. final concentration in PCR reaction: 0.06 uM
 2. 10X = 0.6 uM
 3. For 1000 ul of 10X Vbeta forward mix, combine 6 ul of each primer (at 100 uM stock) (6 ul x 36 primers = 216 ul) + 784 ul of water and store
3. Reaction #2
 - a. Valpha forward mix Rxn #2:

- i. 36 primers: JM_P2_001 -> JM_P2_036
 1. final concentration in PCR reaction: 0.6 uM
 2. 4X = 2.4 uM
 3. For 2000 ul of 4X Valpha forward mix Rxn #2, combine 48 ul of each primer (at 100 uM stock) (48 ul x 36 primers = 1728 ul) and 272 ul of water and store at -20
- b. Vbeta forward mix Rxn #2:
 - i. 36 primers: JM_P2_037 -> JM_P2_072
 1. final concentration in PCR reaction: 0.6 uM
 2. 4X = 2.4 uM
 3. For 2000 ul of 4X Vbeta forward mix Rxn #2, combine 48 ul of each primer (at 100 uM stock) (48 ul x 36 primers = 1728 ul) and 272 ul of water and store at -20

Phenotyping primer mixes for Reactions #1 and #2:

- this is to incorporate the original 17 phenotypic parameters used in the original Han protocol (GATA3, TBET, FOXP3, etc.)
- 1. Reaction #1 forward phenotype mix:
 - a. Final concentration in PCR reaction: 0.1 uM
 - b. 10X = 1 uM
 - c. for 1000 ul of 10X forward phenotype mix, combine 10 ul of each primer (at 100 uM stock) (10 ul x 17 primers = 170 ul) and 830 ul water and store
 - d. use the same strategy for the 10X reverse phenotype mix
- 2. Reaction #2 forward phenotype mix: (JM_R2A_001 -> JM_R2A_017)
 - a. Final concentration in PCR reaction: 0.1 uM
 - b. 10X = 1 uM
 - c. for 1000 ul of 10X forward phenotype mix, combine 10 ul of each primer (at 100 uM stock) (10 ul x 17 primers = 170 ul) and 830 ul water and store
 - d. use the same strategy for the 10X reverse phenotype mix for reaction #2 (JM_R2B_001 -> JM_R2B_017)
 - e. ok to scale up batches to make 4000 ul at a time and store in 1000 ul aliquots

Master Mixes:

1. Reaction #1: no phenotyping

Component	Volume for 1 well of 96-well plate (made 112x for 1 96-well plate)
1. 5X RT-PCR buffer	4 ul
2. dNTPs	0.8 ul
3. 10X Valpha forward mix rxn #1 (see above)	2 ul
4. 10X Vbeta forward mix rxn #1 (see above)	2 ul
5. 10X TRAC reverse (a 1:3.33 dilution from 10 uM stock = 3 uM)	2 ul
6. 10X TRBC reverse (a 1:3.33 dilution from 10 uM stock = 3 uM)	2 ul
7. RT-PCR enzyme	0.8 ul
8. water	6.4 ul

Total volume =	20 ul
----------------	-------

2. Reaction #1: yes phenotyping

Component	Volume for 1 well of 96-well plate (made 112x for 1 96-well plate)
1. 5X RT-PCR buffer	4 ul
2. dNTPs	0.8 ul
3. 10X Valpha forward mix rxn #1 (see above)	2 ul
4. 10X Vbeta forward mix rxn #1 (see above)	2 ul
5. 10X TRAC reverse (a 1:3.33 dilution from 10 uM stock = 3 uM)	2 ul
6. 10X TRBC reverse (a 1:3.33 dilution from 10 uM stock = 3 uM)	2 ul
7. RT-PCR enzyme	0.8 ul
8. 10X Phenotyping forward	2 ul
9. 10X Phenotyping reverse	2 ul
10. water	2.4 ul
Total volume =	20 ul

3. Reaction #2: no phenotyping

Component	Volume for 1 well of 96-well plate (made 112x for 1 96-well plate)
1. 10X PCR buffer	2 ul
2. dNTPs	0.4 ul
3. 4X Valpha forward mix rxn # 2 (see above)	5 ul
4. 4X Vbeta forward mix rxn # 2 (see above)	5 ul
5. 10X TRAC2 reverse (a 1:3.33 dilution from 10 uM stock = 3 uM = 10X)	2 ul
6. 10X TRBC2 reverse (a 1:3.33 dilution from 10 uM stock = 3 uM = 10X)	2 ul
7. HotStar Taq	0.1 ul
8. water	2.5 ul
9. (template from reaction #1)	1 ul per well from plate #1
Total volume =	20 ul

4. Reaction #2: yes phenotyping → this is done **separately** from the TCR alpha/beta reaction 2 in a different plate

Component	Volume for 1 well of 96-well plate (made 112x for 1 96-well plate)
1. 10X PCR buffer	2 ul
2. dNTPs	0.4 ul

3. 10X Phenotype forward rxn #2	2 ul
4. 10X Phenotype reverse rxn #2	2 ul
7. HotStar Taq	0.1 ul
8. water	12.5 ul
9. (template from reaction #1)	1 ul per well from plate #1
Total volume =	20 ul

5. For reaction #3, you can choose to obtain the paired TCR alpha-beta sequences via Sanger sequencing or through next-generation sequencing (NGS) via the Miseq. Phenotypic markers can only be assessed using NGS. NGS is preferred for TCR alpha-beta sequences since there are often 2 alpha chains expressed in a cell (see Han et al.). Sanger sequencing is useful to pilot the reactions and for troubleshooting.
 - a. For Sanger sequencing (choose any forward and reverse alpha **or** beta barcode primers and add in the Illumina paired-end primers 1 & 2). The alpha and beta reactions **must be performed separately** with a 3' reverse alpha barcode primer and a 3' reverse beta barcode primer, respectively.

Component	Volume for 1 well of 96-well plate (made 112x for 1 96-well plate)
1. 10X PCR buffer	2 ul
2. dNTPs	0.4 ul
3. 10X 5' Forward barcode (e.g. JM_31_002) (1:20 dilution from 10 uM stock = 0.5 uM = 10X)	2 ul
4. 10X 3' Reverse barcode for beta (e.g. JM_32_014) (1:20 dilution from 10 uM stock = 0.5 uM = 10X)	2 ul
5. 10X Illumina PE1 (1:2 dilution from 10 uM stock = 5 uM = 10X)	2 ul
6. 10X Illumina PE2 (1:2 dilution from 10 uM stock = 5 uM = 10X)	2 ul
7. HotStar Taq	0.1 ul
8. water	8.5 ul
9. template	1 ul per well from plate #2
Total volume =	20 ul

6. For NGS reaction #3, a similar barcoding strategy described in the Han et al. publication was used except that the plate and row barcodes were reversed (both within the forward primer) as requested by our bioinformatics collaborator Sahil Seth
 - a. To accomplish this, dedicated 96-well plates were filled with the desired plate/row barcodes as well as the desired alpha column and desired beta column barcodes so that they could easily be added to the barcoding plate with a multichannel pipette
 - b. Master mix was added first (see below) followed by the plate/row barcode and then the column barcodes; the 1 ul template from reaction #2 was added last (see attached TCR barcode example sheet). Barcoding for the phenotype amplicons is done in a similar fashion in a separate plate.

Component	Volume for 1 well of 96-well plate (made 112x for 1 96-well plate)
1. 10X PCR buffer	2 ul
2. dNTPs	0.4 ul
3. 10X Illumina PE1 (1:2 dilution from 10 uM stock = 5 uM = 10X)	2 ul
4. 10X Illumina PE2 (1:2 dilution from 10 uM stock = 5 uM = 10X)	2 ul
5. HotStar Taq	0.1 ul
6. water	6.5 ul
volume =	13 ul

+ 2 ul alpha column barcodes (at 10X = 1.5 uM = a 1:6.666 dilution from 10 uM working stock) + 2 ul beta column barcodes (at 10X = 0.5 uM = a 1:20 dilution from 10 uM working stock) + 2 ul row/plate barcodes (at 10X = 0.5 uM = 1 1:20 dilution from 10 uM working stock)
+ 1 ul template from reaction #2 plate = total volume = 20 ul for reaction #3

PCR programs:

Reaction #1

1. 50°C 30 min; 95°C 15 min; 94°C 30 s; 62°C 1 min, 72°C 1 min x 30 cycles; 72°C 5 min; 4°

Reaction #2

2. 95°C 15 min; 94°C 30 s; 64°C 1 min, 72°C 1 min x 25 cycles; 72°C 5 min; 4°; 35 cycles for phenotyping only plate

Reaction #3

3. 95°C 15 min; 94°C 30 s; 64°C 1 min, 72°C 1 min x 36 cycles; 72°C 5 min; 4°

* I increased the # of PCR cycles in reaction #1 from 25 to 30 and this seemed to increase efficiency though I did not rigorously test this effect

* When using the old PCR cycler, the ramp rate for the step-down to annealing temp (62 or 64°C must be **reduced** to 0.5 degrees per second)

For Sanger sequencing: (you must have used only alpha **or** only beta column reverse barcodes since you have to sequence the alpha and beta separately)

- run each alpha and/or beta amplicon on a gel and extract the bright band at 350 – 380 bp using the Qiaquick gel extraction kit (total elution volume of 35 ul in 2 steps) (or if you're very confident or have QC'd other amplicons from the same run, skip the gel) and send for Sanger sequencing using the forward and reverse Sanger sequencing primers (at 1 uM) included in the above reagents list (sequences obtained using the reverse primer were generally the better sequences)
- see attached for sample data

Pooling amplicons for NGS on the Miseq: (preferred)

- pool equal volumes from each well after reaction #3 (I pooled 2 ul from each well of the 96-well reaction #3 plate x 16 wells at a time = 32 ul of pooled amplicons)
- run pooled amplicons on a gel (I ran in an 8-lane gel to facilitate clean extraction)
- extract the bright bands at 350 – 380 bp using the Qiaquick gel extraction kit (I eluted in 35 ul total in 2 elution steps for each band)
- pool all extracted DNA together into 1 eppendorf in equal volumes (for 7 plates, I pooled together 20 ul/plate x 7 plates = 140 ul)
- quantify library using TapeStation and then via qPCR (Rebecca Thornton on South campus helped me with this step) (e.g. the concentration of my first library by qPCR was 47.8 nmol; the library was loaded onto the Miseq at a 1:10 dilution)
- sequencing is performed using 2 x 300 bp paired end reads on the Miseq (v3 Illumina kit) with 10% PhiX

Figure 16

TCR barcode example

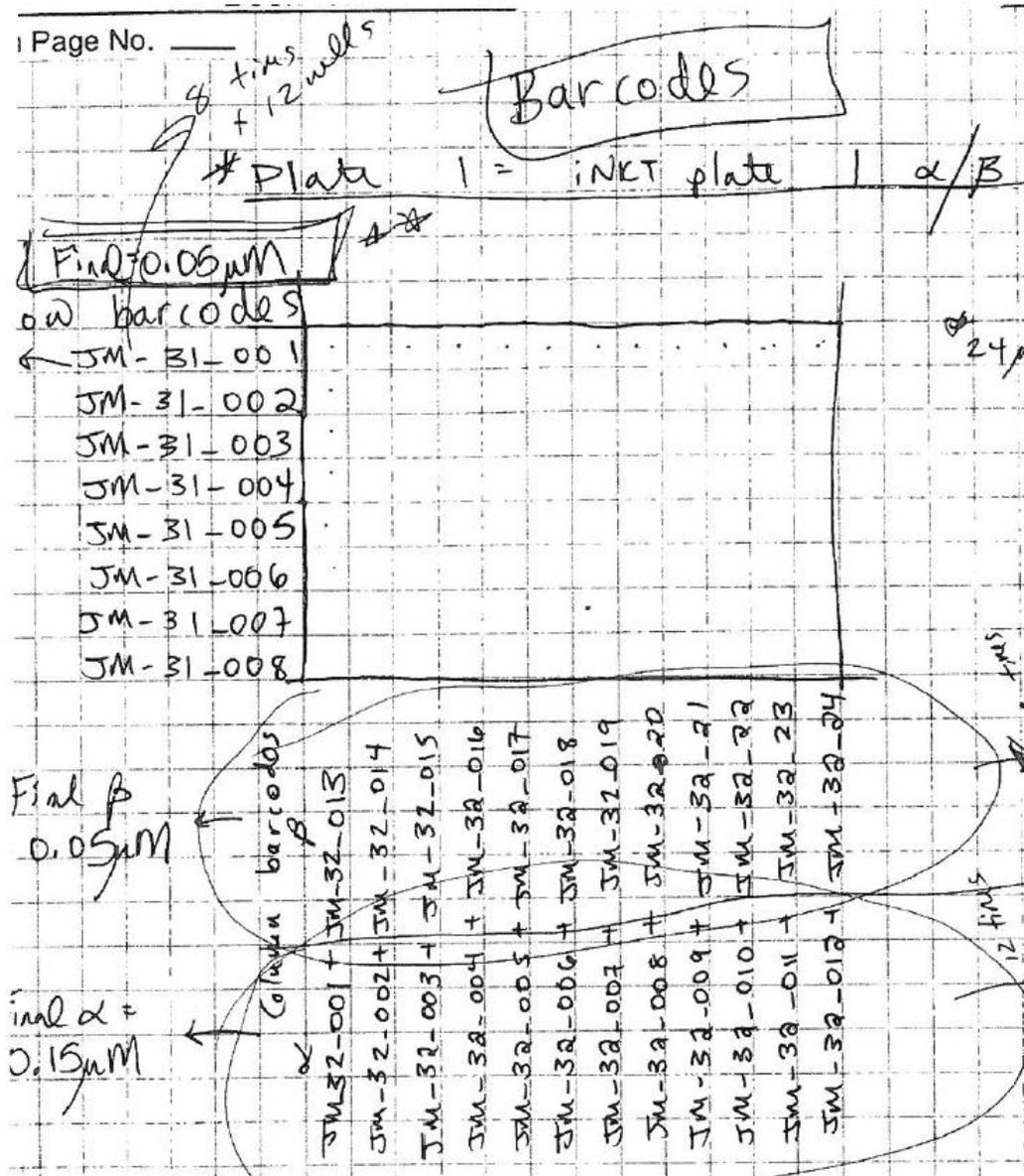


Figure 16. Example layout of TCR α and β barcoding primers in preparation for NGS of single cell TCR transcripts

Figure 17.

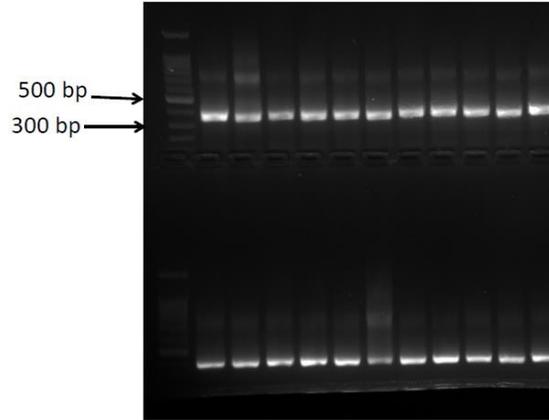


Figure 17. Example gel of the final TCR amplicon at 350 - 380 bp after PCR reaction #3

Figure 18

row	column	reads	Vbeta	Jbeta	CDR3 beta	% reads match	# reads match	dominan t alpha	CDR3 alpha	% match	# reads match	secondary alpha	
A	1	16068	TRBV25-1	TRBJ2-3	CASSDLKEAGDKPSTDTQYF	0.912	965	TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.914	13728	
A	2	27488	TRBV25-1	TRBJ2-5	CASSEASGSRKTQYF	0.933	3631	TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.947	22359	
A	3	99					TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.662	55		
A	4	32124	TRBV25-1	TRBJ2-1	CASSEPGGGEQFF	0.944	15179	TRAV10	TRAJ18	CVVKDRGSTLGRLYF	0.946	15163	
A	5	29071	TRBV25-1	TRBJ1-1	CASSGGHLNTEAFF	0.923	8715	TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.947	18580	
A	6	128					TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.576	64	TRAV10	
A	7	66					TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.551	27		
A	8	32607	TRBV25-1	TRBJ2-7	CASSESRRDRGSYEQYF	0.93	3053	TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.95	27858	
A	9	93					TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.79	64		
A	10	67					TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.81	47		
A	11	30385	TRBV25-1	TRBJ2-1	CASSERNGRYNEQFF	0.93	9102	TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.943	19422	
A	12	27941	TRBV25-1	TRBJ1-1	CASSDWSSNTEAFF	0.937	10322	TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.783	13253	TRAV8-2
B	1	18468	TRBV25-1	TRBJ1-5	CASSGGIGNQPQHF	0.932	8210	TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.847	8167	
B	2	33609	TRBV25-1	TRBJ2-3	CASSARRGPATDTQYF	0.936	19294	TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.946	12234	
B	3	28345	TRBV25-1	TRBJ1-5	CASSEGGRGQPQHF	0.94	11618	TRAV10	TRAJ18	CVVSDRGSTLGRLYF	0.933	14901	

Figure 18. MiSeq output

Example MiSeq output from a pilot experiment using singly sorted invariant natural killer T (iNKT) cells. The row and column numbers refer to each well of a 96 well plate where a single iNKT cell was sorted per well.

Table 4. Primers used for single T cell TCR sequencing

JM_P1_001	CTGCACGTACCAGACATCTGGGTT	JM_P2_025	CCAGGGTTTTCCAGTCACGACAGAAAGTCCAGCACCT	PCR#2 Nested alpha and beta forward and 2 reverse primers		
JM_P1_002	GGCTCAAAGCCCTCTCAGCAGG	JM_P2_026	CCAGGGTTTTCCAGTCACGACATCGCTGAAGACAGAAAGTCCAGT			
JM_P1_003	GGATAACCTGGTTAAAGGCAGCTA	JM_P2_027	CCAGGGTTTTCCAGTCACGACATAACCTTTTCAGTTTGTGATGC			
JM_P1_004	GGATACAAAGCAAAAGTACAAACGA	JM_P2_028	CCAGGGTTTTCCAGTCACGACCTTAACAAAGGTGCCAAGCACCTC			
JM_P1_005	GCTGACGTATATTTTTTCAAAATGGA	JM_P2_029	CCAGGGTTTTCCAGTCACGACAATATCTGCTTCATTTAATGAAAAAGG			
JM_P1_006	GGAAAGAGCCCTGTTTTCTTCTGT	JM_P2_030	CCAGGGTTTTCCAGTCACGACCAAGTGGATGAGAAAAAGCAGCA			
JM_P1_007	GCTGGATATGAGAAGCAGAAAGGA	JM_P2_031	CCAGGGTTTTCCAGTCACGACCTCAGTTTGGTATAACCCAGAAAGGA			
JM_P1_008	AGGACTCCAGCTTCTCTGGAAGTA	JM_P2_032	CCAGGGTTTTCCAGTCACGACGGAAGACTAAGTACGATATTAGATAAG			
JM_P1_009	GTATGTCCAATATCTCTGGAAGGT	JM_P2_033	CCAGGGTTTTCCAGTCACGACCTGTGAACCTTCCAGAAAGCAGCA			
JM_P1_010	CAGTGAAGACACAAGTGTCAACGG	JM_P2_034	CCAGGGTTTTCCAGTCACGACCTCAGTTGATACCAAGGCCGT			
JM_P1_011	CCTAAGTTTGTGTGTCCTGATAC	JM_P2_035	CCAGGGTTTTCCAGTCACGACGCGGAAATTAAGACAAAACCTC			
JM_P1_012	GGSAAGGCCCTGAGTTGATAATGT	JM_P2_036	CCAGGGTTTTCCAGTCACGACATAAATGGCCAAATAAACATACAGG			
JM_P1_013	GCTGATGTACACATACTCCAGTGG	JM_P2_037	CCAGGGTTTTCCAGTCACGACGCTGATGATCAAAATTCACCTCG			
JM_P1_014	CCCTTGGTATAAGCAAGACTGG	JM_P2_038	CCAGGGTTTTCCAGTCACGACTCTCACCTAAATCTCCAGCAAAAGCT			
JM_P1_015	CCTCAATTCATATAGACATTCGTTT	JM_P2_039	CCAGGGTTTTCCAGTCACGACCTGAATGCCCAACAGCTCTC			
JM_P1_016	GCAAAATGCACAACAGTCCGCTA	JM_P2_040	CCAGGGTTTTCCAGTCACGACCTCTGAGCTGAATGAAAGCCCT			
JM_P1_017	TAGAGAGAGCATCAAAAGGCTTCAC	JM_P2_041	CCAGGGTTTTCCAGTCACGACCGATTCCTCAGGGCCAGTCTCT			
JM_P1_018	CGTTCAAATGAAGAGGAAACACAG	JM_P2_042	CCAGGGTTTTCCAGTCACGACTGGCTACAATGCTCCAGATTAACAA			
JM_P1_019	CCTGAAAAGTTCAGAAAACCCAGGAG	JM_P2_043	CCAGGGTTTTCCAGTCACGACCCCTGATGGCTCAACATGCTCCAGA			
JM_P1_020	GCTGGGTATCTTGGAACTCCAG	JM_P2_044	CCAGGGTTTTCCAGTCACGACGTGCTCCAGAGCAACACAGAGGAT			
JM_P1_021	GCTGGGGAAGAAAAGGAGAAAGAAA	JM_P2_045	CCAGGGTTTTCCAGTCACGACGTCTCCAGATCAACACAGAGGAT			
JM_P1_022	GTCAGAGAGAGCAAAACAGTGGAA	JM_P2_046	CCAGGGTTTTCCAGTCACGACGTCTAGATTAACACAGAGGATTTT			
JM_P1_023	GGACAAAACAGATGGAAGATTAAGC	JM_P2_047	CCAGGGTTTTCCAGTCACGADGGCTACATGATATCCAGATCAACAA			
JM_P1_024	CCAGATGTGATGTAAGAAAGTGGAA	JM_P2_048	CCAGGGTTTTCCAGTCACGACTCGCTTCTGTCAGAGAGAGACTGG			
JM_P1_025	GACTTAAATGGGGGATGAAAAGAGA	JM_P2_049	CCAGGGTTTTCCAGTCACGACCGGTTCTTGCAGTCAGGGCTGA			
JM_P1_026	GGAGAAGTGAAGAGAGAGAAAAGC	JM_P2_050	CCAGGGTTTTCCAGTCACGACCCAGTATCGCTTCTTTGAGAAA			
JM_P1_027	CAAATGAAATGGCCTCTCTGATCA	JM_P2_051	CCAGGGTTTTCCAGTCACGACTCTCCACTCTGAMGATCCAGCCCA			
JM_P1_028	GCAATGTGAACAACAGTGGAAAGT	JM_P2_052	CCAGGGTTTTCCAGTCACGACGAGAGAGGCTGAGGGATGAT			
JM_P1_029	GGTGGAGAGTGAAGAGAGCTGAAG	JM_P2_053	CCAGGGTTTTCCAGTCACGACTGAGAGAGGCTAAGGGATCT			
JM_P1_030	GGATAAAATGAAGATGGAAGATTCAC	JM_P2_054	CCAGGGTTTTCCAGTCACGACTCCGGCACAACAGTTCCTGACT			
JM_P1_031	CCTGATGATATTACTGAAGGGTGG	JM_P2_055	CCAGGGTTTTCCAGTCACGACAGATGGCTAYAGTGTCTCTAGATCAA			
JM_P1_032	GGTGGGGAAGAGAAAAGTCAAG	JM_P2_056	CCAGGGTTTTCCAGTCACGACTGTTGCTCCAGATCAAAGACAGAGAA			
JM_P1_033	GGTGAATGACCTCAAATGGAAGAC	JM_P2_057	CCAGGGTTTTCCAGTCACGACGAGAGAGGCTCAAAGGATAGACT			
JM_P1_034	GCTAACTTCAAGTGGAAATGAAAAGA	JM_P2_058	CCAGGGTTTTCCAGTCACGACGTAAAGTGCCTAAGGATCATCTC			
JM_P1_035	GAACTTATAAGCAACAGATGCAAC	JM_P2_059	CCAGGGTTTTCCAGTCACGACTCAGCAGAGATGCTCATGCAACT			
JM_P1_036	GGAGCAGTGAAGCAGAGGGAC	JM_P2_060	CCAGGGTTTTCCAGTCACGACTCTCACTCAACAGTTCAGTGACTA			
JM_P1_037	GAGAGACAATGGAAGACAGCAAAAAC	JM_P2_061	CCAGGGTTTTCCAGTCACGACGCTGAAAGGACTGGAGGGAGCTAT			
JM_P1_038	GCTGAGCTCAGGGAAGAAAGAGC	JM_P2_062	CCAGGGTTTTCCAGTCACGACGATAAAGTCAATCCAGAGGGCGC			
JM_P1_039	CTGAAATATCCGATGATCAATTCGAC	JM_P2_063	CCAGGGTTTTCCAGTCACGACGCTAAGTGGCTCCCAAAATCCACC			
JM_P1_040	TCATATAAATGAAGCAGTTCCAAATCG	JM_P2_064	CCAGGGTTTTCCAGTCACGACGGAACGATTTCTGCTGAATTTCCCA			
JM_P1_041	AGTGTGCCAAGTGGCTCTCAG	JM_P2_065	CCAGGGTTTTCCAGTCACGADGGTACAGGCTCTCTCGGGAAGAA			
JM_P1_042	CACAGGAACATYCCCTCTAGATT	JM_P2_066	CCAGGGTTTTCCAGTCACGACGAGCAAGTTCCTCATCAACATGCA			
JM_P1_043	CAGACACAGAGAAACAAAGCAACTC	JM_P2_067	CCAGGGTTTTCCAGTCACGACTGGATACAGTGTCTCTCCAGAGGC			
JM_P1_044	GGTACCCTGCAAAAGGAGAGCTCC	JM_P2_068	CCAGGGTTTTCCAGTCACGACCAACAGTCTCCAGATAAAGCAGGA			
JM_P1_045	GAGGTACAACCTGCCAAAGGAGAGT	JM_P2_069	CCAGGGTTTTCCAGTCACGACTACAAGTCTCTGCAAAAGCAGAGAGA			
JM_P1_046	GGCAAGAGAGAGTCCCTGATGTT	JM_P2_070	CCAGGGTTTTCCAGTCACGACGGGTACAGTGTCTCTAGAGACA			
JM_P1_047	AAGGAAAGTCCSAAATGGCTACA	JM_P2_071	CCAGGGTTTTCCAGTCACGACTTCCATCAGCCGCCAACCTA			
JM_P1_048	CTGACAAAAGTCCCAATCGCTCAC	JM_P2_072	CCAGGGTTTTCCAGTCACGACAGACCCAGCAGCCGCGAGTTCAT			
JM_P1_049	CAGTACAAAAGGAGAGTCCCGCAT	JM_P2_073	CAGACAGACTGTACTGATTTAG			
JM_P1_050	AGACAAATCAGGGCTGCCAGTGA	JM_P2_074	CTTTGGGTGGGAGATCTCTG			
JM_P1_051	GACTCAGGGCTGCCAAGT	JM_R2A_001	CCAGGGTTTTCCAGTCACGACGCGGAGGAGGTGATGTGCTT		Phenoty ping primers PCR#2	
JM_P1_052	CCAGAATGAAGCTCAACTGACAAA	JM_R2A_002	CCAGGGTTTTCCAGTCACGACCCCAACACAGGAGCCGACTGG			
JM_P1_053	GGTCTCTCCGACAGAGCTCCGAG	JM_R2A_003	CCAGGGTTTTCCAGTCACGACGCGACCCCAAGCCCTGTGCT			
JM_P1_054	GGCTGCCAGTATGCTGGTCTCT	JM_R2A_004	CCAGGGTTTTCCAGTCACGACAGAGGAATCCATGTGGGAGATGT			
JM_P1_055	GACTTACTTCCAGAATGAAGCTCAACT	JM_R2A_005	CCAGGGTTTTCCAGTCACGACGCGAGTGGTGGCAGCCACA			
JM_P1_056	GAGCAAAGGAAACATCTTCAACGATT	JM_R2A_006	CCAGGGTTTTCCAGTCACGACGGCAGCCAGGCGAGCTGTG			
JM_P1_057	GGCTRATCCTTACTCATATGAGTGT	JM_R2A_007	CCAGGGTTTTCCAGTCACGACCTCACGCGCCCACTGCTC			
JM_P1_058	GATAAAGGAGAGTCCCGCTAGTGT	JM_R2A_008	CCAGGGTTTTCCAGTCACGACCCACAGACTGAAACATCTTCAGT			
JM_P1_059	GATTCACAGTTGCTTAAGTGCAT	JM_R2A_009	CCAGGGTTTTCCAGTCACGACCCCAAGCTGGAACCAAGCCCA			
JM_P1_060	GATTCAGGATGCCCGAGGATGG	JM_R2A_010	CCAGGGTTTTCCAGTCACGACAGACCTTTTATGATGGCCCGCT			
JM_P1_061	GATTCGGGATGCCCAAGATCG	JM_R2A_011	CCAGGGTTTTCCAGTCACGACGGTATGGAGCATCAACCTGACAG			
JM_P1_062	GCAGAGCATAAAGCAAGCATCCCT	JM_R2A_012	CCAGGGTTTTCCAGTCACGACAACTGAAACATCATAAACCGGAA			
JM_P1_063	TCGGTATGCCCAACATCAATCT	JM_R2A_013	CCAGGGTTTTCCAGTCACGACGGTCTCTTGGCTGTACTCG			
JM_P1_064	GATTTAAACATGAAGCAGACCCCT	JM_R2A_014	CCAGGGTTTTCCAGTCACGACGGAGCGCTCCCAAGAAAGC			
JM_P1_065	GATGAACAGGTATGCCCAAGAAAG	JM_R2A_015	CCAGGGTTTTCCAGTCACGACCGAGCGGCTCCAGCCCAAGAA			
JM_P1_066	TATCATAGATGAGTCAAGAAATGCCAAAG	JM_R2A_016	CCAGGGTTTTCCAGTCACGACGCCAATTTGACAGCCCAAGAGC			
JM_P1_067	GACTTTCAGAAAGGAGATATAGCTGAA	JM_R2A_017	CCAGGGTTTTCCAGTCACGACCCCAATATAAAGAACAGGAGCC			
JM_P1_068	CAAGGCCACATACGAGCAAGGGTCC	JM_R2B_001	AGCGGATAACAATTTTACACAGGAGGGGAGGGCGGTGGTGCCT			
JM_P1_069	CAAAGATATAAACAAGGAGAGATCTCT	JM_R2B_002	AGCGGATAACAATTTTACACAGGACGCTGTTGGAAAGGTTGACAGT			
JM_P1_070	AGAGAAGGGAGATCTTCTCTGAGT	JM_R2B_003	AGCGGATAACAATTTTACACAGGACAGGATGATGAGCCCGGATGA			
JM_P1_071	GACTGATAAGGGAGATGTTCTCTGAAG	JM_R2B_004	AGCGGATAACAATTTTACACAGGATCAGCATTGTAGCCCGGACATC			
JM_P1_072	GGCTGATCTATTCTCATATGATTTAA	JM_R2B_005	AGCGGATAACAATTTTACACAGGAGGCTGCGGTAGCATTTCTCAGCT			
JM_P1_073	GCCACATATGAGAGTGGATTTGTCTATT	JM_R2B_006	AGCGGATAACAATTTTACACAGGACGGCCGAGGATTGGCCAGCT			
JM_P1_074	GGTGGCCAGAAATCTCTCAGCGT	JM_R2B_007	AGCGGATAACAATTTTACACAGGAGGGTGCATGTAGAGTGGTAGTG			
JM_P1_075	CGGTGAAATAGCCAGACAGACTGT	JM_R2B_008	AGCGGATAACAATTTTACACAGGATTCACAATGGTTGCTGCTCA			
JM_P1_076	ACCAGTGTGGCCTTTGGGGTGG	JM_R2B_009	AGCGGATAACAATTTTACACAGGAGTCAAACTCCTCATGGCTTTGTA			
JM_P2_001	CCAGGGTTTTCCAGTCACGACAGTCTGTTTTCTTCACTCTAGTCT	JM_R2B_010	AGCGGATAACAATTTTACACAGGAGGCACAGTCTCCTGTTGAAATGCA			
JM_P2_002	CCAGGGTTTTCCAGTCACGACAGCATACAACATGACCTATGAAGCG	JM_R2B_011	AGCGGATAACAATTTTACACAGGAGGCTCTTACAAACTGGGCCAC			
JM_P2_003	CCAGGGTTTTCCAGTCACGACCTTGAAGCTGAAATTAACAAGAGCC	JM_R2B_012	AGCGGATAACAATTTTACACAGGAGGGACAGAGTTCATGGGTAGT			
JM_P2_004	CCAGGGTTTTCCAGTCACGACACTCCCTGTTTTATCCCTGCCGAC	JM_R2B_013	AGCGGATAACAATTTTACACAGGAGTTTGAAGTAAAGGAGACAAATTTG			
JM_P2_005	CCAGGGTTTTCCAGTCACGACAAACAAGACCAAGACTCACTGTCT	JM_R2B_014	AGCGGATAACAATTTTACACAGGACGAGAAGTATGATGACTGCTGCT			
JM_P2_006	CCAGGGTTTTCCAGTCACGACAAAGACTGAAGTCACTTTGATACC	JM_R2B_015	AGCGGATAACAATTTTACACAGGACCCGACAACTCCGGTGAATCA			
JM_P2_007	CCAGGGTTTTCCAGTCACGACATAAATGCTACTACTGAAGAATGG	JM_R2B_016	AGCGGATAACAATTTTACACAGGAGGGTGGCTAGTGGAGATAAG			
JM_P2_008	CCAGGGTTTTCCAGTCACGACGATCAACGGTTTTGAGGCTGAATTTAA	JM_R2B_017	AGCGGATAACAATTTTACACAGGAGCCACACTGCATGTCTGCCCT			
JM_P2_009	CCAGGGTTTTCCAGTCACGACGAAACCACTTCTTCCACTTGGAGAA	Peprimer1	AATGACGCGGACCCAGGATCTACACTCTTCCCTACACAGGCTCTCCGATCT			Paired end adaptor primers and sequencing primers
JM_P2_010	CCAGGGTTTTCCAGTCACGACTACAGCAACTCTGGATGCAGACAC	Peprimer2	AAGCAGAAGCAGGCATACGAGATCGGTCTCGGCTCTGCTGTAACCCGCTCTCCGATCT			
JM_P2_011	CCAGGGTTTTCCAGTCACGACGAAAGTGAAGTTTTACAGCACA	sequencing primer	ACACTCTTCCCTACACAGCGCTCTCCGATCT			
JM_P2_012	CCAGGGTTTTCCAGTCACGACGACATTCGTTCAAATGTGGGGGAA	reverse	CGGTCTCGGCATTCTGCTGAACCCGCTCTCCGATCT			
JM_P2_013	CCAGGGTTTTCCAGTCACGACGGCAAGGCAAGAGTCAAGCT					
JM_P2_014	CCAGGGTTTTCCAGTCACGACTCCAGAAAGGCAAGAAATCCGGCA					
JM_P2_015	CCAGGGTTTTCCAGTCACGACGCTGACCTTAAACAAGGCGAGACA					
JM_P2_016	CCAGGGTTTTCCAGTCACGACTTAAAGTCAAGCTGTGACACTTCCA					
JM_P2_017	CCAGGGTTTTCCAGTCACGACGCAAGGTTTTCAAGCCAGTCTCA					
JM_P2_018	CCAGGGTTTTCCAGTCACGACTCCACAGTCTCTCAACTTCCACC					
JM_P2_019	CCAGGGTTTTCCAGTCACGACGCCACATTAACAAGAAAGGAAAGCT					
JM_P2_020	CCAGGGTTTTCCAGTCACGACGCTCGCTGGATAAATCATCAGGA					
JM_P2_021	CCAGGGTTTTCCAGTCACGACAGACTGTCGCTACGGAAGCCTA					
JM_P2_022	CCAGGGTTTTCCAGTCACGACCAACTTCTCTCAATAAAGGTGCCA					
JM_P2_023	CCAGGGTTTTCCAGTCACGACAGAAATAAGTCCACTCTTAATACCA					
JM_P2_024	CCAGGGTTTTCCAGTCACGACTTGGAGAAGCAAAAAGAACAGCT					

Table 4. continued.

CCTACACGACGCTCTCCGATCTTTAAGCGAGCAGACAGGGTTTCCCAGTCACGAC	IM_31_001	CCTACACGACGCTCTCCGATCTACCTCAGGAATGCCAGGGTTTCCCAGTCACGAC	IM_31_075
CCTACACGACGCTCTCCGATCTGCTGCAGGACAGACAGGGTTTCCCAGTCACGAC	IM_31_002	CCTACACGACGCTCTCCGATCTCGGGAAATGATGCCAGGGTTTCCCAGTCACGAC	IM_31_076
CCTACACGACGCTCTCCGATCTACTCAGGAGCAGACAGGGTTTCCCAGTCACGAC	IM_31_003	CCTACACGACGCTCTCCGATCTGACGAGGGAATGCCAGGGTTTCCCAGTCACGAC	IM_31_077
CCTACACGACGCTCTCCGATCTCGGGAATGAGCAGACAGGGTTTCCCAGTCACGAC	IM_31_004	CCTACACGACGCTCTCCGATCTAAAGGAGGGAATGCCAGGGTTTCCCAGTCACGAC	IM_31_078
CCTACACGACGCTCTCCGATCTGACGAGGAGCAGACAGGGTTTCCCAGTCACGAC	IM_31_005	CCTACACGACGCTCTCCGATCTGTTGGAATGCCAGGGTTTCCCAGTCACGAC	IM_31_079
CCTACACGACGCTCTCCGATCTAAAGGAGGAGCAGACAGGGTTTCCCAGTCACGAC	IM_31_006	CCTACACGACGCTCTCCGATCTCTCAACTGAATGCCAGGGTTTCCCAGTCACGAC	IM_31_080
CCTACACGACGCTCTCCGATCTGTTGTTGGAGCAGACAGGGTTTCCCAGTCACGAC	IM_31_007	CTGTGAAACCGCTCTCCGATCTCGGTTCACTGACTGGATTAGAGTCTCTCAG	IM_32_001
CCTACACGACGCTCTCCGATCTCTCAACTGAGCAGACAGGGTTTCCCAGTCACGAC	IM_31_008	CTGTGAAACCGCTCTCCGATCTCCAGGACTCACTGGATTAGAGTCTCTCAG	IM_32_002
CCTACACGACGCTCTCCGATCTTTAAGCGATCGAACAGGGTTTCCCAGTCACGAC	IM_31_009	CTGTGAAACCGCTCTCCGATCTCATTATAGTCACTGGATTAGAGTCTCTCAG	IM_32_003
CCTACACGACGCTCTCCGATCTGCTGCAGATCGAACAGGGTTTCCCAGTCACGAC	IM_31_010	CTGTGAAACCGCTCTCCGATCTGACTGTGCTCACTGGATTAGAGTCTCTCAG	IM_32_004
CCTACACGACGCTCTCCGATCTACTCAGGATCGAACAGGGTTTCCCAGTCACGAC	IM_31_011	CTGTGAAACCGCTCTCCGATCTTACCAGGCTCACTGGATTAGAGTCTCTCAG	IM_32_005
CCTACACGACGCTCTCCGATCTCGGGAATGATCGAACAGGGTTTCCCAGTCACGAC	IM_31_012	CTGTGAAACCGCTCTCCGATCTATACTAGTCACTGGATTAGAGTCTCTCAG	IM_32_006
CCTACACGACGCTCTCCGATCTGACGAGGAGTCAACAGGGTTTCCCAGTCACGAC	IM_31_013	CTGTGAAACCGCTCTCCGATCTCTGTAGGCTCACTGGATTAGAGTCTCTCAG	IM_32_007
CCTACACGACGCTCTCCGATCTAAAGGAGGATCGAACAGGGTTTCCCAGTCACGAC	IM_31_014	CTGTGAAACCGCTCTCCGATCTTGGAGTGTCACTGGATTAGAGTCTCTCAG	IM_32_008
CCTACACGACGCTCTCCGATCTGTTGTTGGAGCAGACAGGGTTTCCCAGTCACGAC	IM_31_015	CTGTGAAACCGCTCTCCGATCTTAGGCTAGTCACTGGATTAGAGTCTCTCAG	IM_32_009
CCTACACGACGCTCTCCGATCTCTCAACTGATCGAACAGGGTTTCCCAGTCACGAC	IM_31_016	CTGTGAAACCGCTCTCCGATCTCGGAAATGCTCACTGGATTAGAGTCTCTCAG	IM_32_010
CCTACACGACGCTCTCCGATCTTTAAGCGAAACCAACAGGGTTTCCCAGTCACGAC	IM_31_017	CTGTGAAACCGCTCTCCGATCTCACCAGCTCACTGGATTAGAGTCTCTCAG	IM_32_011
CCTACACGACGCTCTCCGATCTGCTGCAGGAAACCAACAGGGTTTCCCAGTCACGAC	IM_31_018	CTGTGAAACCGCTCTCCGATCTGTGAGAGCTCACTGGATTAGAGTCTCTCAG	IM_32_012
CCTACACGACGCTCTCCGATCTACTCAGGAAACCAACAGGGTTTCCCAGTCACGAC	IM_31_019	CTGTGAAACCGCTCTCCGATCTGCTCTCAGAGATCTCTGCTCTGATGGCTC	IM_32_013
CCTACACGACGCTCTCCGATCTCGGGAATGAGGAAACCAACAGGGTTTCCCAGTCACGAC	IM_31_020	CTGTGAAACCGCTCTCCGATCTCCAGGAGAGATCTCTGCTCTGATGGCTC	IM_32_014
CCTACACGACGCTCTCCGATCTGACGAGGGAACCAACAGGGTTTCCCAGTCACGAC	IM_31_021	CTGTGAAACCGCTCTCCGATCTCATTATAGAGATCTCTGCTCTGATGGCTC	
CCTACACGACGCTCTCCGATCTAAAGGAGGAAACCAACAGGGTTTCCCAGTCACGAC	IM_31_022	CTGTGAAACCGCTCTCCGATCTGACTGTGAGATCTCTGCTCTGATGGCTC	
CCTACACGACGCTCTCCGATCTGTTGTTGGAGCAGACAGGGTTTCCCAGTCACGAC	IM_31_023	CTGTGAAACCGCTCTCCGATCTTACCAGGAGATCTCTGCTCTGATGGCTC	
CCTACACGACGCTCTCCGATCTCTCAACTGAAACCAACAGGGTTTCCCAGTCACGAC	IM_31_024	CTGTGAAACCGCTCTCCGATCTATACTAGAGATCTCTGCTCTGATGGCTC	
CCTACACGACGCTCTCCGATCTTTAAGCGAGGTTGCCAGGGTTTCCCAGTCACGAC	IM_31_025	CTGTGAAACCGCTCTCCGATCTGCTAGGAGATCTCTGCTCTGATGGCTC	
CCTACACGACGCTCTCCGATCTGCTGCAGGAGGTTGCCAGGGTTTCCCAGTCACGAC	IM_31_026	CTGTGAAACCGCTCTCCGATCTTGGAGTGTGAGATCTCTGCTCTGATGGCTC	
CCTACACGACGCTCTCCGATCTACTCAGGAAACCAACAGGGTTTCCCAGTCACGAC	IM_31_027	CTGTGAAACCGCTCTCCGATCTTAGGCTAGGATCTCTGCTCTGATGGCTC	
CCTACACGACGCTCTCCGATCTCGGGAATGAGGTTGCCAGGGTTTCCCAGTCACGAC	IM_31_028	CTGTGAAACCGCTCTCCGATCTCGGAAATGAGATCTCTGCTCTGATGGCTC	
CCTACACGACGCTCTCCGATCTGACGAGGAGGTTGCCAGGGTTTCCCAGTCACGAC	IM_31_029	CTGTGAAACCGCTCTCCGATCTCACCAGGAGATCTCTGCTCTGATGGCTC	
CCTACACGACGCTCTCCGATCTAAAGGAGGAGGTTGCCAGGGTTTCCCAGTCACGAC	IM_31_030	CTGTGAAACCGCTCTCCGATCTGTGAGAGATCTCTGCTCTGATGGCTC	
CCTACACGACGCTCTCCGATCTGTTGTTGGAGGTTGCCAGGGTTTCCCAGTCACGAC	IM_31_031	CTGTGAAACCGCTCTCCGATCTCAAGCGGATAACAATTCACACAGGA	IM_32_025
CCTACACGACGCTCTCCGATCTCTCAACTGAGGTTGCCAGGGTTTCCCAGTCACGAC	IM_31_032	CTGTGAAACCGCTCTCCGATCTCCAGGAGCGGATAACAATTCACACAGGA	IM_32_026
CCTACACGACGCTCTCCGATCTTTAAGCGATGTTGCCAGGGTTTCCCAGTCACGAC	IM_31_033	CTGTGAAACCGCTCTCCGATCTCATTAAAGCGGATAACAATTCACACAGGA	IM_32_027
CCTACACGACGCTCTCCGATCTGCTGCAGATGTTGCCAGGGTTTCCCAGTCACGAC	IM_31_034	CTGTGAAACCGCTCTCCGATCTGACTGTAGCGGATAACAATTCACACAGGA	IM_32_028
CCTACACGACGCTCTCCGATCTACTCAGGATGTTGCCAGGGTTTCCCAGTCACGAC	IM_31_035	CTGTGAAACCGCTCTCCGATCTTACCAGGAGGATAACAATTCACACAGGA	IM_32_029
CCTACACGACGCTCTCCGATCTCGGGAATGATGTTGCCAGGGTTTCCCAGTCACGAC	IM_31_036	CTGTGAAACCGCTCTCCGATCTATACTAAGCGGATAACAATTCACACAGGA	IM_32_030
CCTACACGACGCTCTCCGATCTGACGAGGAGTGGTCCAGGGTTTCCCAGTCACGAC	IM_31_037	CTGTGAAACCGCTCTCCGATCTCTGAGAGCGGATAACAATTCACACAGGA	IM_32_031
CCTACACGACGCTCTCCGATCTAAAGGAGGATGTTGCCAGGGTTTCCCAGTCACGAC		CTGTGAAACCGCTCTCCGATCTTGGAGTGTAGCGGATAACAATTCACACAGGA	IM_32_032
CCTACACGACGCTCTCCGATCTGTTGTTGGAGCAGACAGGGTTTCCCAGTCACGAC		CTGTGAAACCGCTCTCCGATCTTAGGCTAAGCGGATAACAATTCACACAGGA	IM_32_033
CCTACACGACGCTCTCCGATCTCTCAACTGATGTTGCCAGGGTTTCCCAGTCACGAC		CTGTGAAACCGCTCTCCGATCTCACCAGCGGATAACAATTCACACAGGA	IM_32_034
CCTACACGACGCTCTCCGATCTTTAAGCGAGATGCCAGGGTTTCCCAGTCACGAC		CTGTGAAACCGCTCTCCGATCTGTGAGAGCGGATAACAATTCACACAGGA	IM_32_035
CCTACACGACGCTCTCCGATCTGCTGCAGGAAACCAACAGGGTTTCCCAGTCACGAC		GACGGGGCGAGTACCCCGT	GATA3
CCTACACGACGCTCTCCGATCTAAAGGAGGAAATGCCAGGGTTTCCCAGTCACGAC	IM_32_044	GCCCTGACGTCCACCCGGACT	TBBT
CCTACACGACGCTCTCCGATCTGACGAGGGAATGCCAGGGTTTCCCAGTCACGAC	IM_32_045	GGCTTCTGTCGATGATGCTGCT	FOXP3
CCTACACGACGCTCTCCGATCTAAAGGAGGAAATGCCAGGGTTTCCCAGTCACGAC	IM_32_046	CCCGGGAGGAAATGATCTGGCTA	RORC
CCTACACGACGCTCTCCGATCTGTTGTTGGAGCAGACAGGGTTTCCCAGTCACGAC	IM_32_047	CCGCAACATGGTGGAGGTGCT	RUNX1
CCTACACGACGCTCTCCGATCTCTCAACTGACATGCCAGGGTTTCCCAGTCACGAC	IM_32_048	CGCTCGATGGTGGAGTGGCT	RUNX3
CCTACACGACGCTCTCCGATCTTTAAGCGAATGCCAGGGTTTCCCAGTCACGAC	IM_31_049	GCCAAACAGAGGGGCGCTGAG	BCL6
CCTACACGACGCTCTCCGATCTGCTGCAGAAATGCCAGGGTTTCCCAGTCACGAC	IM_31_050	CTCACATTTAAGTTTACATGCCCAA	IL2
CCTACACGACGCTCTCCGATCTACTCAGGAAATGCCAGGGTTTCCCAGTCACGAC	IM_31_051	CCAGTTTACCTGGAGGAGGTGA	IL10
CCTACACGACGCTCTCCGATCTCGGGAATGAAATGCCAGGGTTTCCCAGTCACGAC	IM_31_052	GGGAGTTGCCCTGGCTCCAGAA	IL12A
CCTACACGACGCTCTCCGATCTGACGAGGAAATGCCAGGGTTTCCCAGTCACGAC	IM_31_053	CCCAGAAACAGAAAGCTCCGCT	IL13
CCTACACGACGCTCTCCGATCTAAAGGAGGAAATGCCAGGGTTTCCCAGTCACGAC	IM_31_054	GACAAAGACTCCCCCGGACTG	IL17A
CCTACACGACGCTCTCCGATCTGTTGTTGGAGCAGACAGGGTTTCCCAGTCACGAC	IM_31_055	GGCTTTACGCTCTGATCGTTTT	IFNG
CCTACACGACGCTCTCCGATCTCTCAACTGAAATGCCAGGGTTTCCCAGTCACGAC	IM_31_056	CATGATCCGGGACGTGGAGCT	TNFA
CCTACACGACGCTCTCCGATCTTTAAGCGAGGTTGCCAGGGTTTCCCAGTCACGAC	IM_31_057	GCATATATATGTTCTCAACACATCA	TGFB
CCTACACGACGCTCTCCGATCTGCTGCAGAGGTTCCAGGGTTTCCCAGTCACGAC	IM_31_058	GTGCTGTGGCCGGCTCACAC	PERFORIN
CCTACACGACGCTCTCCGATCTACTCAGGAGGTTCCAGGGTTTCCCAGTCACGAC	IM_31_059	GGGAAGCTCCATAAATGTCACCTT	GRANZYMB
CCTACACGACGCTCTCCGATCTCGGGAATGAGCAGGTTCCAGGGTTTCCCAGTCACGAC	IM_31_060	GGAGAAGGGCTGAGATCCAG	GATA3
CCTACACGACGCTCTCCGATCTGACGAGGAGGTTCCAGGGTTTCCCAGTCACGAC	IM_31_061	CTGGGTTCTTGGAAATGAAGATAT	TBBT
CCTACACGACGCTCTCCGATCTAAAGGAGGAGGTTCCAGGGTTTCCCAGTCACGAC	IM_31_062	GTCGGCTGCTCTCTGGAGCT	FOXP3
CCTACACGACGCTCTCCGATCTGTTGTTGGAGCAGACAGGGTTTCCCAGTCACGAC	IM_31_063	CCATGCCACGGTATTTGGCTTCAA	RORC
CCTACACGACGCTCTCCGATCTCTCAACTGACGGTTCCAGGGTTTCCCAGTCACGAC	IM_31_064	GGTCATTAATCTTGGCAACTGGTT	RUNX1
CCTACACGACGCTCTCCGATCTTTAAGCGAATGCCAGGGTTTCCCAGTCACGAC	IM_31_065	CGTTGACCTGGCCACTGGTT	RUNX3
CCTACACGACGCTCTCCGATCTGCTGCAGAAATGCCAGGGTTTCCCAGTCACGAC	IM_31_066	GAGAGCCGAGGACGTGCACTT	BCL6
CCTACACGACGCTCTCCGATCTACTCAGGAATGCCAGGGTTTCCCAGTCACGAC	IM_31_067	GACAAAGGTAACTCTGTTGAG	IL2
CCTACACGACGCTCTCCGATCTCGGGAATGAAATGCCAGGGTTTCCCAGTCACGAC	IM_31_068	GTAGGCTCTATGTAGTTGATGAAGA	IL10
CCTACACGACGCTCTCCGATCTGACGAGGAAATGCCAGGGTTTCCCAGTCACGAC	IM_31_069	CGGTTCTTCAAGGAGGATTTTTGT	IL12A
CCTACACGACGCTCTCCGATCTAAAGGAGGAAATGCCAGGGTTTCCCAGTCACGAC	IM_31_070	CCCTCCGAAAGTTCTTTAAAT	IL13
CCTACACGACGCTCTCCGATCTGTTGTTGGAGCAGACAGGGTTTCCCAGTCACGAC	IM_31_071	GGACCCAGATCTCTGCTGGAT	IL17A
CCTACACGACGCTCTCCGATCTCTCAACTGAAATGCCAGGGTTTCCCAGTCACGAC	IM_31_072	GGATGCTCTGGTCACTTTAAAGTT	IFNG
CCTACACGACGCTCTCCGATCTTTAAGCGAATGCCAGGGTTTCCCAGTCACGAC	IM_31_073	GGCTACAGGCTTGTCACTCG	TNFA
CCTACACGACGCTCTCCGATCTGCTGCAGAAATGCCAGGGTTTCCCAGTCACGAC	IM_31_074	CCCTCCAGGGCTCAACCACT	TGFB
		CCGATATGGGCCACCCAGCT	PERFORIN
		GTTTTCCAGGGGGGCGCTC	GRANZYMB

Column barcodes for alpha, beta, and phenotype

Plate and row barcodes

Phenotype forward and reverse primers PCR#1

Results:

Single cell analysis of candidate GVL T cells

We isolated CLL neoantigen-specific T cells using the tetramer-based, single cell method described. We then mapped the TCR beta chain sequence obtained from the single cell method onto the bulk Adaptive TCR beta chain data. This allowed us to visualize the expansion/contraction of the relevant clones during periods of leukemia response or significant subclonal leukemic evolution.

First, we predicted neoantigens for each patient at each time point across their specific HLA molecules using the leukemia WES data. Next, we used public CLL RNA sequencing (RNA-seq) data to evaluate leukemic expression of neoantigens of interest (particularly strong binders with a predicted IC50 < 150 nM) [17]. Single cell RNA-seq and scATAC-Seq of sorted pre-alloSCT CLL cells were then used to confirm expression of neoantigens using the *vartrix* algorithm (Figure 19).

We analyzed pre-transplant, sort-purified leukemia cells from patient CLL 13, a complete responder with a CLL driver *SF3B1 K700E* mutation, for whom there was sufficient material available for both scRNA-seq and scATAC-Seq assays (Figure 19A and B). Both assays confirmed expression of the *SF3B1 K700E* mutation. Similarly, we detected *CHEK2 P92L* (see Chapter 3) in both scRNA-seq and scATAC-Seq assays in sorted pre-transplant cells from CLL 1, but not CLL 13, who was wild type for *CHEK2* (Figure 19C and D). We were unable to confirm the expression of all relevant neoantigens due to sample availability; however, these data support the utility of single cell expression analyses for this purpose when sample quantity is limited.

CLL patient 12 experienced a CR of heavily pretreated CLL after an HLA-matched, nonmyeloablative alloSCT from a MUD. His course was complicated by fungal infection and chronic GVHD of the gut and skin. His CLL never relapsed and he passed away 6 years post-

alloSCT from unknown causes. We conducted WES of 2 pre-alloSCT samples from CLL patient 12, 106 and 17 days prior to his transplant (Figure 20A). In the interval between the samples, he received a fourth cycle of chemotherapy with hyperCVAD in combination with rituximab and alemtuzumab. WES of both samples confirmed mutations in *TP53* (2 separate mutations), *SF3B1* (K700E), and *PARPBP* with high allelic fractions (Figure 20A). Pyclone analysis estimated the presence of these mutations in nearly 100% of CLL cells in cluster 0 (Figure 20B). Several exonic mutations formed predicted strong neoantigens in the context of the patient's HLA molecules including SF3B1 K700E and PARPBP N355I, which were both predicted to bind to HLA A*02:05 (Figure 20C).

We generated PARPBP and SF3B1 neoantigen tetramers (GLVDEQQEV and GLSNFIIFI, both HLA A*02:05) and isolated CD8+, tetramer-binding T cells at 3 time points for CLL patient 12 (Figure 20D). Unfortunately, sample viability was intermediate in the first and last samples. Only 1 reliable TCR beta sequence was recovered for the SF3B1 tetramer in the single cell sorts – CASSYAISVPSYNEQFF – and this sequence was found in only 1 of the 4 longitudinal adaptive bulk repertoire samples (75 days post-alloSCT) at a frequency of 0.94%. For the PARPBP neoantigen tetramer, 5 TCR beta sequences were recovered (Figure 21A). Of these 5 TCR beta sequences from the single T cell neoantigen tetramer sorts, 3 of them were found at an appreciable frequency in all 4 time points from the longitudinal bulk TCR data (Figure 21B). In the day +50 post-transplant sample, the PARPBP-tetramer binding clones were present at expanded frequencies of 29%, 2.9%, and 7.3% of the T cell repertoire.

In an antigen-specific response, multiple T cell clonotypes are recruited [108]. In addition, different nucleotide (DNA) sequences can encode for the same TCR Vbeta amino acid sequence, known as 'convergent recombination', which can also indicate an antigen-specific T cell response [109]. Within CLL patient 12's longitudinal bulk TCR repertoire data, we did see evidence of convergent DNA recombination for the PARPB neoantigen-specific T cell clones. Specifically, there were 7 unique DNA sequences for CASSVTGGYNEQFF, 3 for

CASSLLPESADTQYF, and 2 for CASSISGRSVPGELFF. However, convergent DNA recombination was also seen for many other T cell clones in CLL patient 12's repertoire that did not bind the PARPBP neoantigen tetramer (Figure 21C). There are several potential explanations: these T cells could be responding to other antigens including other CLL neoantigens, LAAs, mHAs involved in GVHD/GVL, viral antigens, or it could be a general feature of the post-alloSCT repertoire, a question that has not been investigated.

Finally, in addition to sequencing the TCR, we utilized flow cytometry and the single cell gene expression data to investigate the immunophenotype of the singly sorted, mutant PARPBP tetramer-binding CD8⁺ T cells. Tetramer sorted T cells had either a central memory (CCR7⁺, CD45RA⁻) (clones 3A19 and 3B19) or an effector phenotype (CCR7⁻, CD45RA⁺) (clones 3E6 and 3E23) and an activated (CD69⁺CD107⁺) surface expression profile (Figure 21D). This is expected since post-transplant, reconstituting T cells are skewed towards an antigen-experienced, effector memory phenotype [110]. Together, we identified neoantigen-specific T cells whose expansion and contraction post-alloSCT coincided with clinical CR. However, tumor cytotoxicity assays are necessary to confirm anti-CLL activity but were beyond the scope of this work.

CLL patient 8 was discussed in Chapter 2 as an example of CLL subclonal genomic evolution. Briefly, the patient's leukemia persisted after a nonmyeloablative alloSCT from his sister. He then received 2 DLIs, also from his sister, and briefly responded but ultimately relapsed and passed away from refractory CLL. Significant subclonal evolution was seen after the 2 DLIs as shown in Figure 22A. Strong predicted neoantigens included those resulting from somatic nonsynonymous mutations in *ZCWPW1*, *ACTN1*, and *ADCY1* (Figure 22B). Single cell RNA-seq confirmed expression of the *ACTN1* alternate allele in the early post-alloSCT disease but not in the late post-alloSCT disease (Figure 22C and D). The *ADCY1* mutant allele was not detected using single cell RNA-seq (data not shown), though emergence of an expressed *TP53*

mutation was confirmed (Figure 22D, inset). Importantly, the single cell RNA-seq data aligned with the WES data and showed clear post-alloSCT leukemic evolution at the population level.

A persistently restricted T cell repertoire was observed in CLL patient 8 post-alloSCT (Figure 22E, left panel). A post-DLI bone marrow sample was subjected to neoantigen, tetramer-based (ACTN1 M460I in HLA*B 15:01 and ZCWPW1 in HLA*A 02:01) single T cell sorting and TCR sequencing. None of the ZCWPW1 neoantigen-binding TCR beta sequences were found in the longitudinal TCR repertoire data. In contrast, similar to CLL patient 14, four of the ACTN1 (YCIARIAPY/HLA*B 15:01) neoantigen-binding T cell clones, with high relative tetramer binding intensity, could be mapped onto the longitudinal repertoires, linked by the TCR beta sequence. Leukemic subclonal evolution occurred between d+176 and d+463 and the ACTN1 neoantigen-binding clones were found at expanded frequencies during that time window (Figure 22E, right panel and arrows). Again, cytotoxic assays are needed to differentiate true anti-CLL activity from TCR promiscuity or frequent clones stochastically landing in the tetramer gate during single T cell tetramer sorting. Nevertheless, this work is proof of concept of a robust method that links detailed information about a single T cell, including its antigen specificity and immunophenotype, to the broader context of where the clone resides within the shape of a T cell repertoire and its adaptation over time.

Figure 19

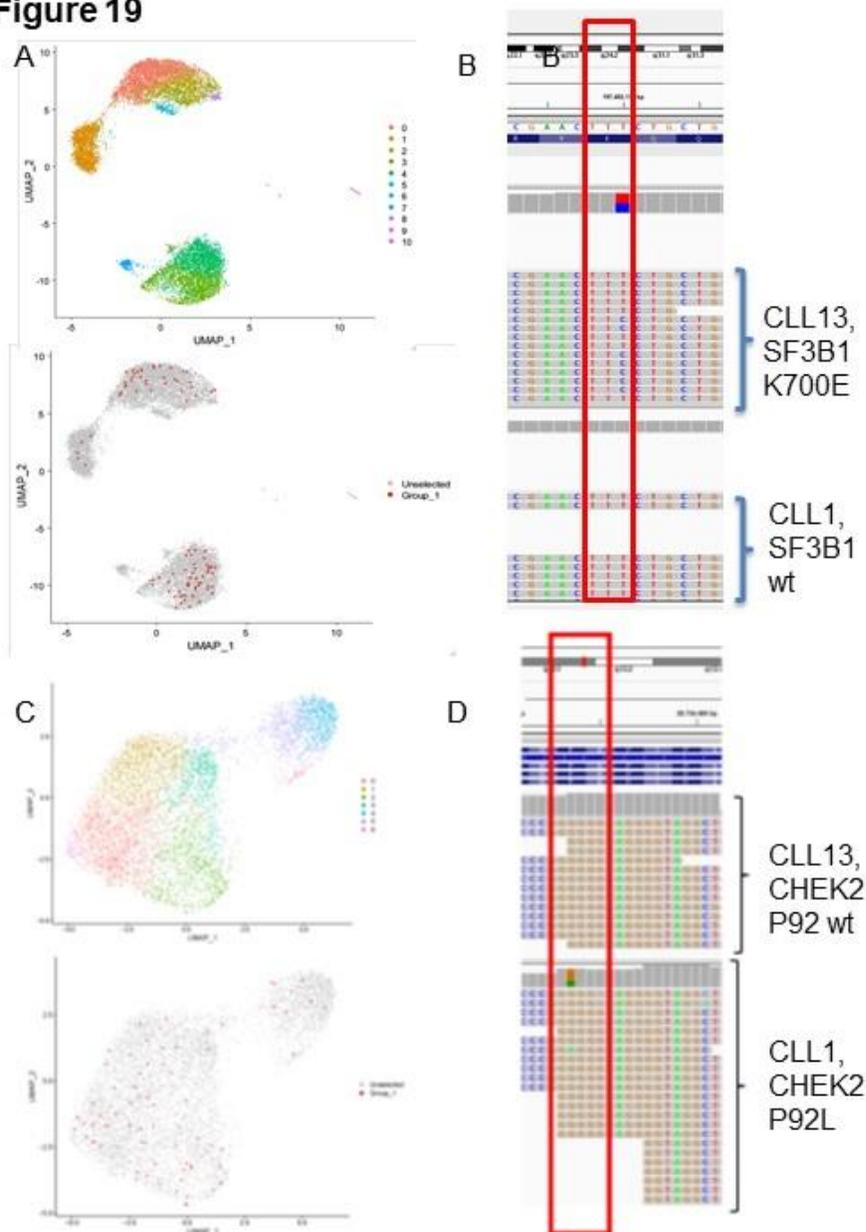


Figure 19. Single leukemia cell expression of neoantigens

(A) Clustering of gene expression from sorted leukemia cells from CLL patient 13 using uniform manifold approximation and projection (UMAP). Transcripts encoding SF3B1 K700E were detected using Vartrix, a software tool from 10X Genomics. The upper panel indicates that 11 different gene expression clusters were identified within the CLL population. The lower panel indicates that the altered SF3B1

allele was detected in all clusters (red dots). (B) View of IGV reader from single cell ATAC-Seq data from sort-purified leukemia from CLL patient 13. The SF3B1 K700E allele is detected in CLL patient 13, but not in CLL patient 1 who was SF3B1 wild type. (C) Similar to A, UMAP analysis of sort-purified leukemia from CLL patient 1 reveals 7 clusters with the CHEK2 P92L mutant transcript found in all clusters. (D) As in B, the IGV reader view confirms expression of the mutant CHEK2 allele in CLL patient 1 but not CLL patient 13.

Figure 20

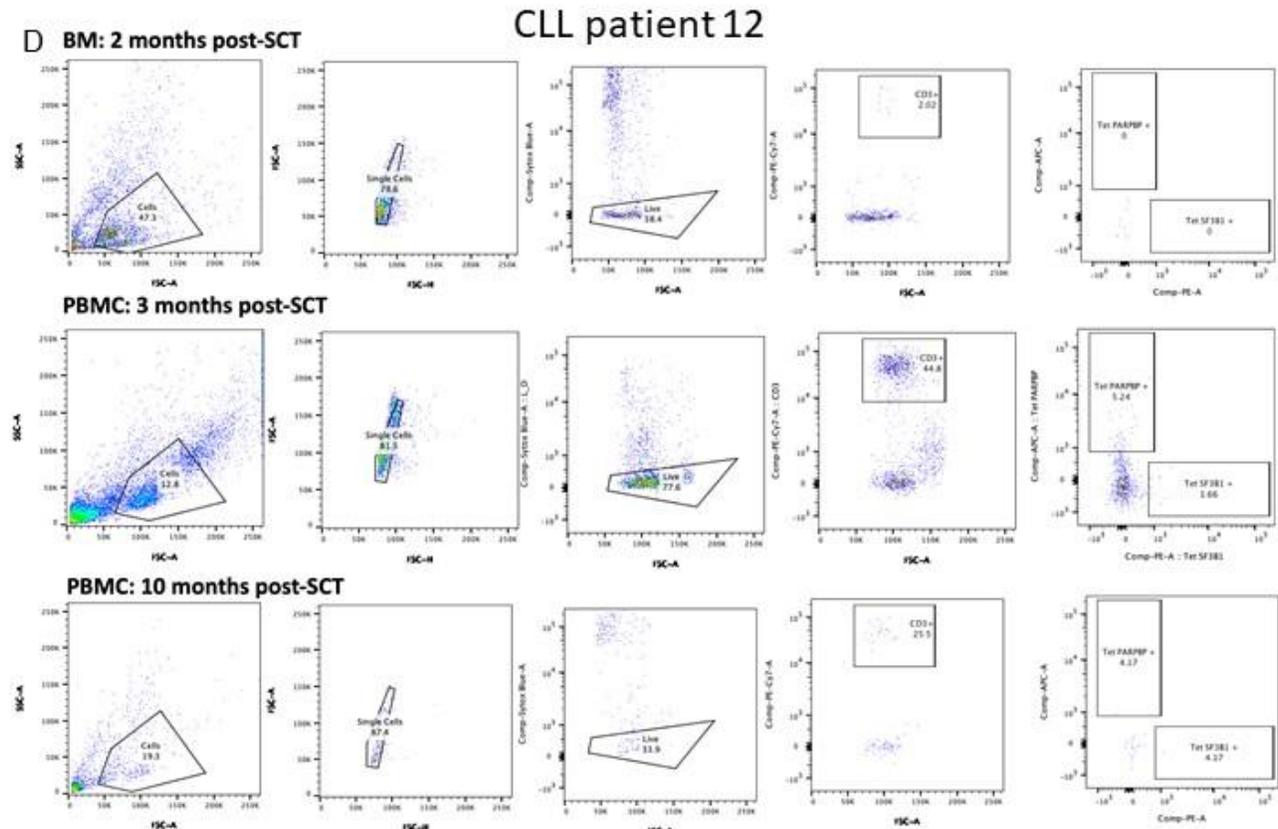
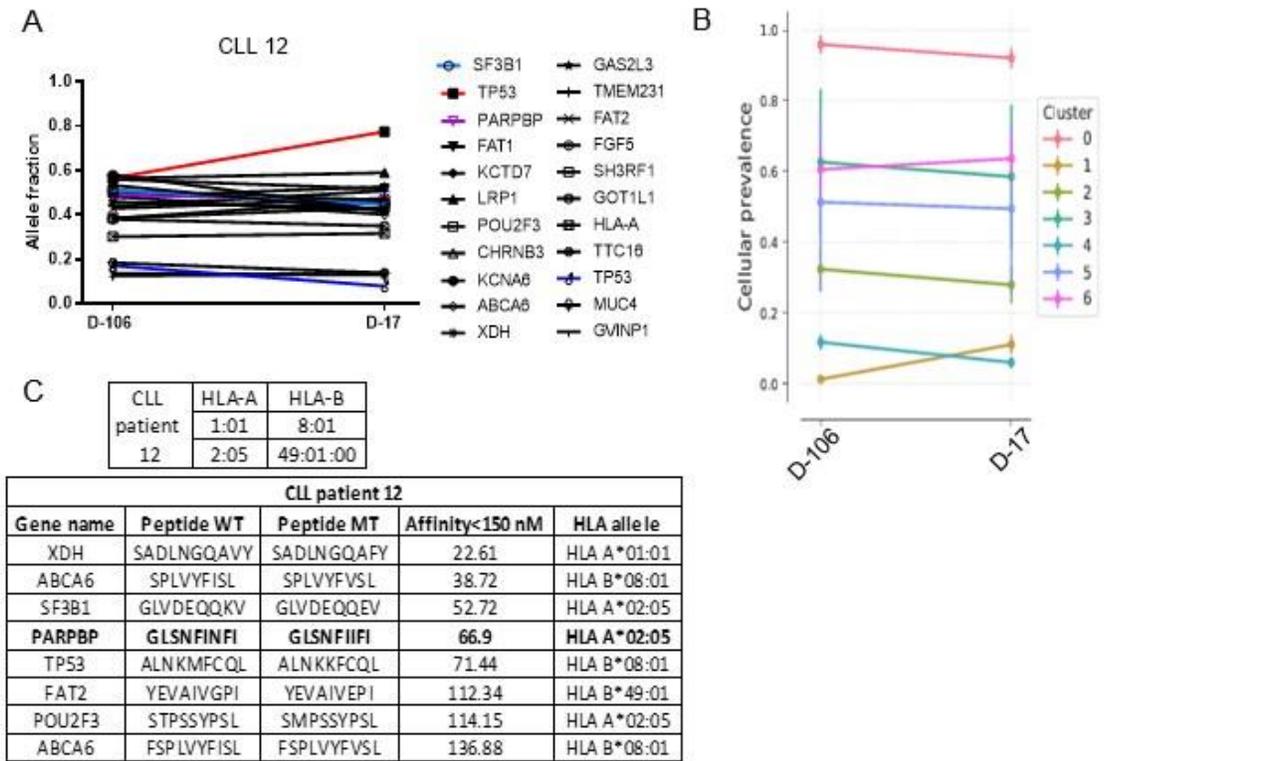


Figure 20. Neoantigen-specific T cells in CLL patient 12

(A) WES profile of purified CLL from 2 pre-alloSCT samples in CLL patient 12. (B) Pyclone analysis of WES data from CLL patient 12. This tool estimates the cellular prevalence of given somatic mutations. Of note, the PARPBP mutation is in cluster 0. (C) HLA type and subset of predicted CLL neoantigens for CLL patient 12. (D) Tetramers for neoantigens in SF3B1 and PARPBP (both HLA 02:05) were folded and 3 samples from CLL patient 12 were used to isolate neoantigen specific T cells via single T cell sorting.

Figure 21. Integrating neoantigen tetramer-binding TCR β sequences with longitudinal, bulk repertoire TCR β sequences for CLL patient 12

(A) TCR sequencing of mutant PARPBP, neoantigen tetramer-binding T cells revealed 5 unique CDR3 sequences. (B) Three of the 5 unique CDR3 sequences from single T cell, neoantigen tetramer sorting were found within the longitudinal bulk TCR sequencing data for CLL patient 12 and these clones are indicated by arrows. The 3 T cell clones were present at expanded frequencies during the early post-SCT period and then contracted over time. (C) Convergent DNA recombination data showing that many of the CDR3 amino acid sequences, including the PARPBP neoantigen-binding clones, are encoded by multiple, unique DNA sequences. (D) Immunophenotyping by gene expression and flow cytometry of PARPBP neoantigen tetramer binding T cells (with CDR3 CASSVTGGYNEQFF).

Figure 21

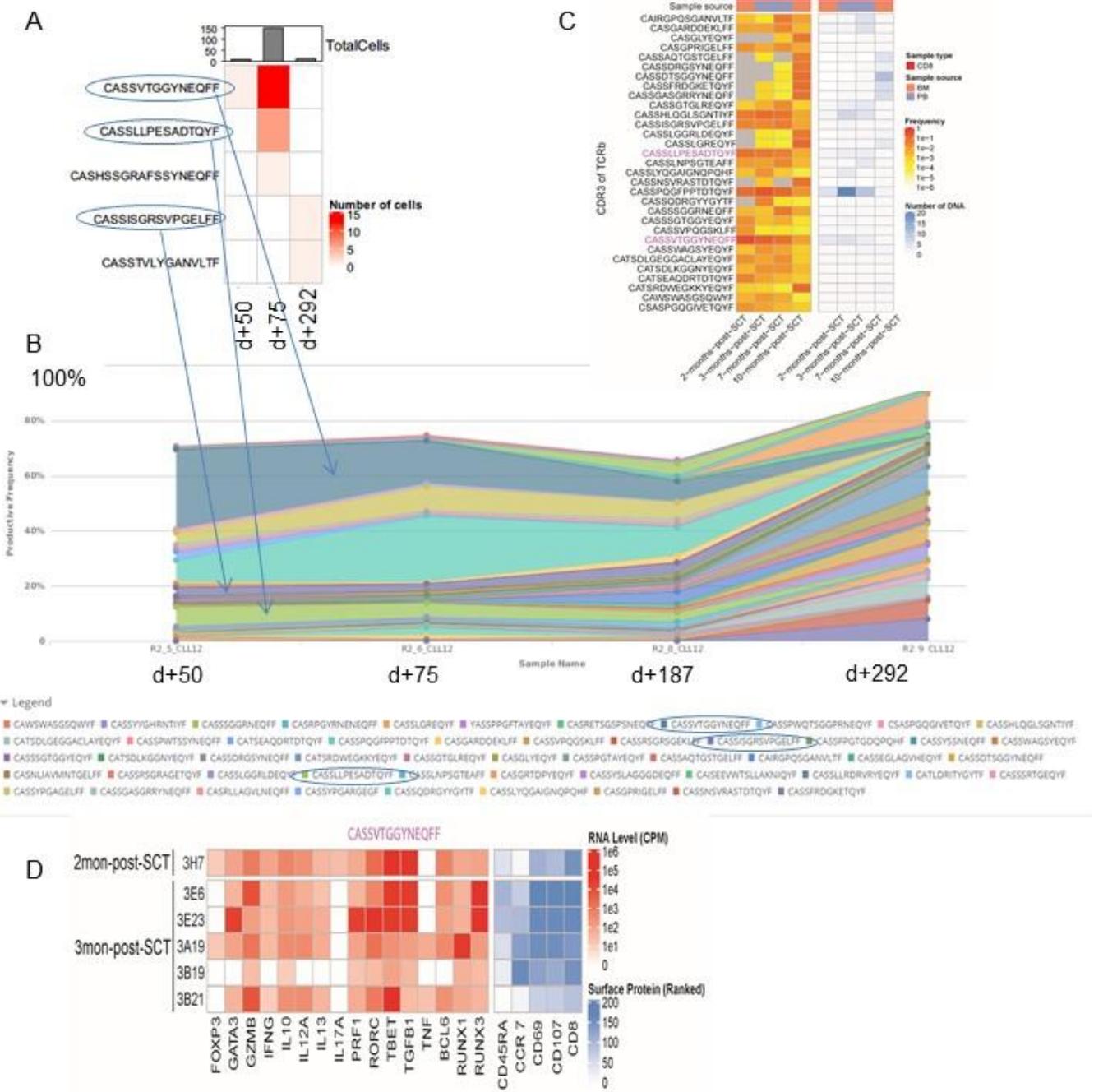


Figure 22

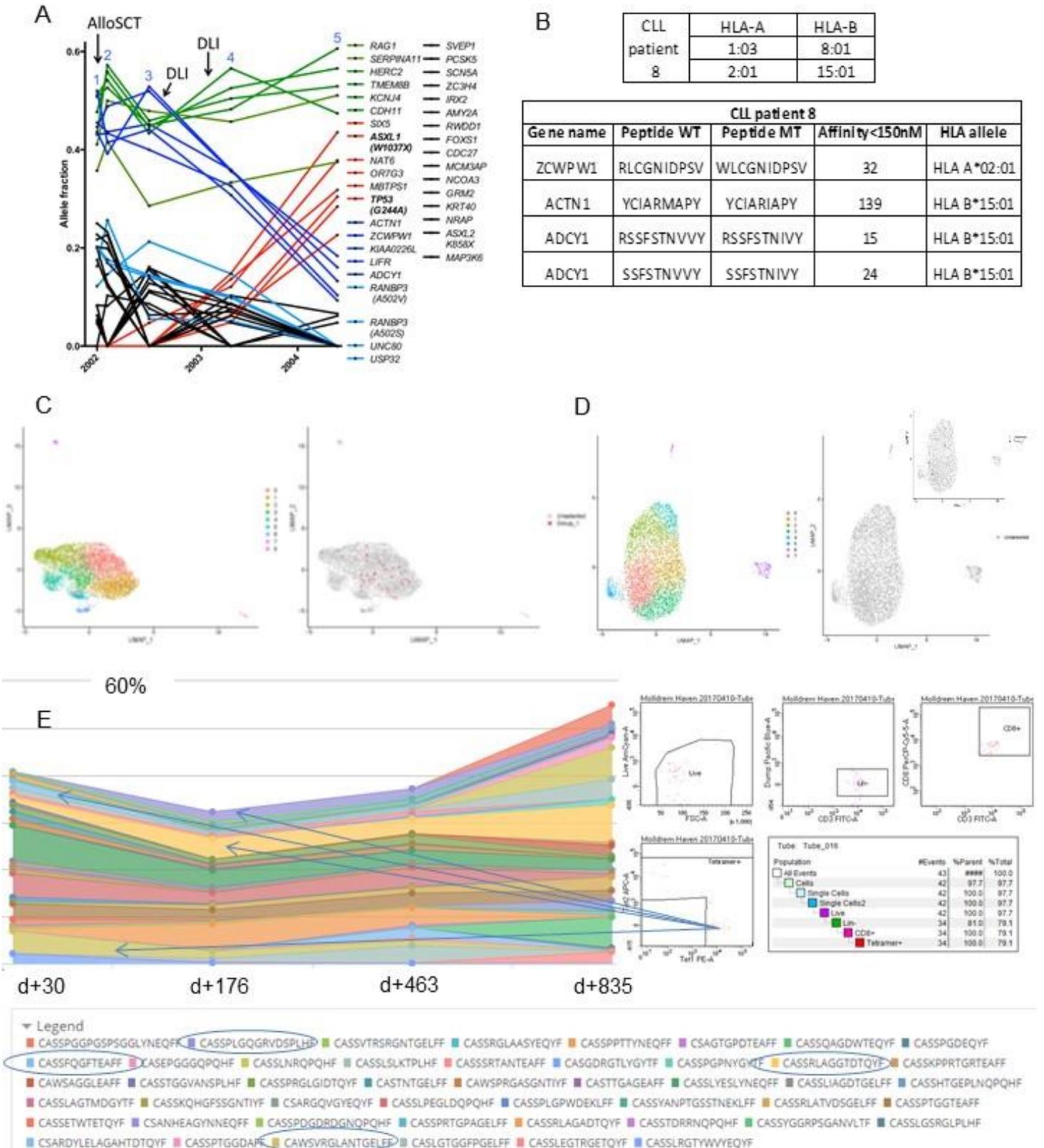


Figure 22. Neoantigen-specific T cells in CLL patient 8

(A) As in Figure 3D, recurrent, somatic exonic nonsynonymous mutations in purified leukemia from CLL patient 8 are shown across 5 time points. (B) HLA type and subset of predicted CLL neoantigens for CLL patient 8. (C) Single cell RNA-seq data from an early post-alloSCT purified leukemia sample from CLL patient 8. UMAP analysis shows 9 clusters of CLL (left panel). Expression of the mutant ACTN1 M460I transcript is found in nearly all of the clusters in agreement with the WES data (right panel, red dots). (D) Single cell RNA-seq data from a late post-alloSCT purified leukemia sample from CLL patient 8. UMAP analysis reveals 8 clusters of CLL that are distinct from those in the earlier post-alloSCT sample in C (left panel). Surprisingly, though the ACTN1 mutation was still detected in a subclone in the WES data, the altered allele was no longer expressed in the late post-alloSCT disease (right panel). The inset shows expression of the mutant TP53 allele (G244A), concordant with the WES evolution data. (E) Longitudinal bulk TCR sequencing data (left) and neoantigen tetramer binding T cells for CLL patient 8 (for ACTN1, M460I in HLA*B 15:01) (right). The arrows match a given neoantigen-binding T cell clone to its longitudinal frequency within the post-alloSCT T cell repertoire of CLL patient 8 using TCR beta chain sequences. Of note, significant subclonal evolution in the patient's disease was observed between d+176 and d+463 (time points 3 and 4 in the WES plot).

Chapter 7: Discussion and Future Directions

Discussion: Our study of patients who received a matched, nonmyeloablative alloSCT for chemorefractory CLL revealed marked subclonal leukemic evolution after transplant. The majority of patients (8 of 11) with recurrent or persistent disease post-alloSCT experienced shifts in the molecular composition of their disease, which encompassed somatic point mutations and CNAs. Branched and linear patterns of CLL evolution were observed post-alloSCT, and these changes included recognized CLL driver lesions in every case.

In two patients (CLL patients 5 and 8), branched subclonal evolution coincided with DLI, a treatment that consists of billions of donor lymphocytes that are infused in the absence of conditioning chemotherapy. This is strong evidence that allogeneic T cells were responsible for the changes in CLL disease architecture in these patients. For the remaining 6 patients who experienced post-alloSCT disease evolution but who either didn't receive a DLI or for whom the timing of disease evolution in relation to DLI was less clear, it is important to emphasize that 5 of those 6 patients received nonmyeloablative conditioning with alloSCT (CLL patient 9 received reduced-intensity conditioning). Nonmyeloablative regimens were designed to optimize donor T engraftment, relying on T cells to eradicate disease via the GVL effect rather than cytotoxic chemotherapy [111]. Moreover, the nonmyeloablative regimen that these patients received was either FC or FCR chemotherapy, a regimen all of the patients had received previously that proved ineffective for their CLL. In this context, we expect that any observed post-alloSCT disease evolution is attributable, at least in part, to selective pressure from the graft rather than to chemotherapy alone.

In addition, our study provides a strategy to disentangle donor and host exomes post-alloSCT. The resultant, reproducible longitudinal CLL exome and copy number data enabled us to investigate a unique pattern of post-alloSCT CLL branched evolution in CLL patient 1. This led to the discovery of a novel candidate driver mutation in the protein kinase, Chk2 (c.275T,

p.P92L). Co-immunoprecipitation and ultracentrifugation techniques showed that the clonal *CHEK2* somatic mutation found in CLL patient 1 strengthened both homo- and heterodimerization of the kinase. These data suggest impairment in Chk2 P92L function that may be relevant in leukemogenesis and highlight the power of studying a single patient's disease longitudinally. We identified *CHEK2* as a candidate driver in CLL through in-depth study of one patient compared to the hundreds of patients it required using a bioinformatics approach.

Thirteen patients with a CR to alloSCT were added to the cohort to enrich the T cell analyses and to enable comparison of CLL molecular features between responders and nonresponders. The 24 patients in the final cohort had CLL with aggressive molecular features, including enrichment for recognized CLL drivers (e.g. *TP53* mutation/loss) and an unmutated IGHV gene. This was expected since all 24 patients had chemorefractory leukemia. We did not detect any significant differences between responders and nonresponders with respect to the number/composition of exonic mutations, copy number changes, or the leukemic neoantigen burden. However, a difference was observed in CLL structural heterogeneity between response groups. Complete responders to alloSCT tended to have disease that was more structurally clonal than nonresponders, meaning the copy number gains and losses detected in CR patients tended to be homogeneously distributed throughout the CLL population. In contrast, alloSCT nonresponders had CLL with one or more subclones containing private CNAs that were not shared by the larger CLL population. This suggests that alloSCT may be more effective for patients with structurally pure CLL and less potent in patients with CLL that is structurally heterogeneous, a finding that warrants further study.

Next, we focused on the post-alloSCT donor T cells to assess for evolution of the T cell repertoire in response to host CLL. Of the 24 patients with CLL WES data, 19 had sufficient samples for TCR analyses. T cell reconstitution post-alloSCT is immensely complex and is its own field of study. Multiple factors are at play including conditioning type, graft type (related vs. unrelated vs. umbilical cord and matched vs. haploidentical), infection (bacterial, fungal, and

viral), viral reactivation (EBV, CMV, etc.), GVHD, tolerance, post-alloSCT immunosuppression, and thymic output. The TCR repertoire diversity in the CLL alloSCT cohort was markedly oligoclonal compared to healthy donors. For many patients, this defect in the T cell repertoire persisted for years.

TCR diversity was more dynamic over time in CLL transplant recipients than in healthy donors; however, there was no clear pattern of clonality change in alloSCT responders versus nonresponders. Among patients with disease response post-alloSCT, those with chronic GVHD had a higher clonality (i.e. a more oligoclonal T cell repertoire) compared to patients without chronic GVHD. Altogether, though, the sample size was too limited to identify a signature of the GVL effect within the bulk, longitudinal TCR repertoire data. As discussed below, resolving a pattern will require large numbers of patients to account for important clinical variables like viral reactivation and GVHD.

Lastly, we focused on individual T cell clones within the expansive T cell repertoires to test our hypothesis of coevolving host leukemia and donor T cell clones. Personalized neoantigens were predicted for each patient using the WES data. Single tumor cell RNA sequencing confirmed expression of neoantigens and corroborated subclonal leukemic evolution post-alloSCT. We then adapted an established single T cell TCR sequencing method to identify CLL neoantigen tetramer-binding CD8 T cell clones by their TCR beta chain sequence. In two patients, these clones were mapped onto the respective patient longitudinal bulk T cell repertoires. The candidate GVL T cell clones were present at expanded frequencies within the post-alloSCT T cell populations during periods of disease response (CLL 14) and disease evolution (CLL 8). However, tumor cytotoxicity assays are needed to confirm anti-CLL activity. In summary, this work provides a robust method to link detailed information about a single T cell, including its antigen specificity, TCR sequence, and immunophenotype, to its position in the larger T cell repertoire. Importantly, the strategy permits tracking of a given clone through its contraction and/or expansion within the repertoire during periods of leukemia

response or in the context of other clinical parameters (e.g. GVHD, viral reactivation, disease progression, etc.).

There is a growing body of literature with which to put these findings in context. In 1976, Dr. Peter Nowell described a model of the evolution of heterogeneous tumor cell populations within a patient's malignancy. In this report, he called tumor subclones tumor 'sublines' and warned that even personalized therapies could be thwarted by the emergence of treatment-resistance sublines [112]. Recent advances in genetic sequencing and single tumor cell analyses have revealed a clearer picture of the intratumoral heterogeneity that he predicted and confirm that clonal evolution is precipitated by therapeutic intervention in both liquid and solid tumors. For example, investigators studied serial samples of 14 patients with myelodysplastic syndrome treated with lenalidomide and identified examples of both branching and linear subclonal disease evolution post-treatment. In most patients, lenalidomide effectively eliminated clones containing deletion 5q; however, related subclones emerged to compose the relapsed disease and harbored mutations in characteristic MDS drivers including *TET2*, *ASXL1*, and *DNMT3A* [113]. A separate study utilizing WES in endocrine-resistant breast cancer demonstrated clonal evolution of the disease after treatment with palbociclib and fulvestrant in 12 of 14 patients [114]. Genetic alterations that emerged after progression on therapy included mutations in *RB1* and *FGFR2*. Targeted sequencing of circulating tumor DNA in an additional 195 patients pre- and post-treatment revealed outgrowth of tumor clones containing mutations in *PIK3CA* (16/195 patients) and *ESR1* (25/195 patients, particularly Y537S).

With respect to subclonal evolution in CLL, Dr. Dan Landau and Dr. Catherine Wu lead the field. In their initial article, an algorithm called ABSOLUTE was used to estimate the cancer cell fraction of candidate CLL driver alterations detected in 149 CLL patient samples (80% of patients were untreated, 20% had received prior chemotherapy) [56]. This analysis revealed that some CLL drivers, including *MYD88* mutations, trisomy 12, and deletion 13q, tended to be clonal lesions implying they were acquired early in leukemogenesis, while others, like *TP53*

loss/mutation, tended to be found at subclonal fractions, suggesting they were acquired late in the CLL disease course. In addition, they conducted WES of serial CLL samples (median 3.5 years apart) in 12 patients receiving interval chemotherapy (most FC or FCR) and 6 CLL patients receiving no intervening therapy. Of the 12 patients who received chemotherapy, 10 showed significant CLL clonal evolution including outgrowth of subclones containing del11q, *SF3B1* mutations, and *TP53* alterations. In contrast, clonal evolution was seen in only 1 of the 6 patients who did not receive treatment in between the 2 time points. Of the 11 patients with disease evolution, 5 had a branched pattern and 6 had a linear pattern. Lastly, they showed that the presence of a subclonal driver alteration in CLL was associated with a shorter FFS post-chemotherapy compared to patients without detectable subclonal drivers.

In a follow-up report, Landau et al. reported WES results from 538 CLL patients, including 278 untreated patients from the CLL8 trial that proved the superiority of FCR therapy over FC [17, 115]. Again, the ABSOLUTE algorithm was employed to categorize clonal versus subclonal CLL somatic alterations and in this cohort, chemotherapy was associated with the emergence of subclones containing alterations in *TP53*, *BIRC3*, del17p, del11q, *DDX3X*, and *MAP2K1*. They also studied paired pre-treatment and relapse samples for 59 patients on the CLL8 trial. Remarkably, clonal evolution was observed in 57 of the 59 patients after chemotherapy. The clonal fractions of leukemic populations containing *TP53* mutations/del17p or *IKZF3* mutations tended to increase at relapse post-chemotherapy though there was substantial diversity among the emergent subclonal drivers. Subclones containing alterations in *SF3B1* or *ATM* were as likely to expand as they were to contract after chemotherapy while predicted early events like trisomy 12, del13q, and del11q tended to remain clonal. Importantly, the drivers in dominant post-chemotherapy subclones were detectable at low levels within the pre-treatment disease in 42% of cases via WES or targeted sequencing implying that the observed evolution represented true outgrowth of a pre-existing subclone rather than acquisition of new lesions from genotoxic chemotherapy.

Clonal evolution of CLL also occurs after targeted therapy with the Bruton's tyrosine kinase (BTK) inhibitor, ibrutinib [20, 116]. In this context, subclonal genetic diversity represents the primary mechanism of treatment resistance. The vast majority of the CLL population is susceptible to the drug; however, rare subclones harboring mutations in *BTK* (C481S) and *PLCG2*, which encodes a protein downstream of BTK, expand under therapeutic pressure and cause disease relapse. Notably, *TP53* mutant CLL is susceptible to ibrutinib and, distinct from the clonal evolution seen after chemotherapy, subclones with *TP53* mutations were as likely to increase as decrease during ibrutinib therapy [116].

There are no published reports of CLL clonal evolution in the context of alloSCT, but there are 3 recent studies that describe clonal evolution of AML after alloSCT. In the first report, targeted sequencing of 35 AML-specific genes was conducted in 15 patients at three longitudinal time points: diagnosis, pre-alloSCT, and post-alloSCT relapse [117]. Considerable heterogeneity was observed among the longitudinal mutation profiles. The authors concluded that post-alloSCT relapse was characterized by reexpansion of the original disease clone in 4 of 9 patients and by expansion of a rare AML subclone in 5 of 9 patients. AML subclones that expanded at relapse carried mutations in *NRAS*, *DNMT3A*, *TET2*, and *TP53*.

In a separate study, 15 paired AML samples from diagnosis and post-alloSCT relapse were studied using WES [118]. There was heterogeneity in the landscape of somatic SNVs and CNAs in post-alloSCT AML, and no shared driver mutations/CNAs were found to associate with post-alloSCT relapse. In addition, there were no recurrent somatic mutations or CNAs in immune related genes in the relapsed post-alloSCT AML. In general, the somatic alterations in the post-alloSCT relapsed disease resembled those found in post-chemotherapy relapsed disease. The major finding of the study came from longitudinal RNA sequencing of AML blasts from 7 patients with post-alloSCT relapse. In this group, 221 genes were found to be differentially expressed between the diagnosis and post-alloSCT relapse samples. Pathway analysis indicated that 19 of the 30 highly enriched pathways involved the innate and adaptive

immune response. The most obvious link to immunity was the downregulation of MHC class II genes (HLA- DPA1, DPB1, DQB1, and DRB1) in 6 of 7 patients in the post-alloSCT sample compared to the diagnosis sample. This finding was verified using flow cytometry, IHC, and single AML cell RNAseq in these and additional AML patients.

Post-alloSCT downregulation of MHC Class II molecules in AML was corroborated in a separate report [119]. Twelve patients had AML samples available at diagnosis and post-alloSCT relapse, and SNP profiling of these cases demonstrated clonal evolution of CNAs in 9 of 12 patients. These CNAs were common AML driver events and were not linked to distinct immune hotspots. GEP by microarray was then performed in 9 patients at diagnosis and at post-alloSCT relapse. Deregulated pathways included those in T cell costimulation and antigen processing/presentation via HLA class II molecules. Subsequent flow cytometry analysis revealed loss of surface expression of HLA-DR and HLA-DP in 28 of 69 (40%) patients between diagnosis and post-alloSCT relapse. In both reports, a small subset of post-alloSCT AML blasts were treated with IFN-gamma *in vitro* and MHC class II protein expression was restored within 72 hours.

Clonal evolution of cancer in the context of a different form of immunotherapy (i.e. not alloSCT) has also been studied. A recent report described clonal evolution in 68 melanoma patients after immunotherapy with nivolumab, an anti-PD-1 checkpoint inhibitor [120]. Approximately 50% of the patients studied had progressed on prior ipilimumab, an anti-CTLA-4 checkpoint inhibitor. In the 50% of patients who had not received prior immunotherapy, higher tumor mutation load (and higher clonal tumor mutation load) was associated with improved OS. No single gene SNVs or CNAs were associated with either response or resistance to nivolumab. In 41 patients, WES was performed on paired pre- and on-nivolumab melanoma samples. In responders, after adjusting for tumor purity, there was a reduction in the melanoma mutation burden and neoantigen load after the initiation of nivolumab compared to nonresponders in whom these metrics did not change. In addition, all patients with CR/PR experienced

contraction of 1 or more subclones on therapy while patients with SD/PD experienced expansions of subclones on nivolumab. RNA-seq analysis showed differentially expressed genes in the pre-treatment and on-treatment samples in several immune pathways in responding versus nonresponding patients, including in HLA class II expression, checkpoint-related genes (OX40, TIGIT, VISTA), and genes involved in lymphocyte activation and cytolytic activity. For patients with PD, there was no unifying expression signature of tumor-intrinsic immune evasion. Finally, the investigators performed TCR sequencing through Adaptive Biotechnologies on melanoma tumor-infiltrating lymphocytes (TIL). In ipilimumab-naïve patients, a decrease in the evenness metric (i.e. an increasingly oligoclonal repertoire) from the pre- to on-nivolumab biopsy was associated with treatment benefit (CR/PR/SD vs. PD).

Several mechanisms of primary or therapy-acquired resistance to immunotherapy have been described [121]. Some are rooted in intratumoral heterogeneity and the expansion of subclonal immune escape variants while others are adaptive, such as tumor cells expressing PD-L1 in response to interferon-gamma exposure. Immunotherapy resistance mechanisms include but are not limited to the expression of immune checkpoint proteins [122-125], defects in interferon-receptor signaling [126, 127], loss of Beta 2 microglobulin expression [128], somatic HLA class I mutations [57, 129], tumor intrinsic beta-catenin signaling [130], tumor-intrinsic MAPK signaling or loss of PTEN [131, 132], immunoediting [133, 134], and an immune-suppressive tissue microenvironment [135-137].

When we consider our CLL WES data through the lens of these important publications, several similarities and differences emerge. Our first question was whether the CLL population evolved post-alloSCT. The answer was unequivocally yes. Analogous to patients in the studies referenced above, the majority of CLL patients experienced clonal shifts in their disease after alloSCT manifested by expansions/contractions of subclones containing distinct exonic SNVs and CNAs. Also similar to the other studies, we observed mixed patterns of branched and linear evolution after alloSCT treatment. This observation is not trivial because the patients we studied

were heavily pretreated with chemotherapy. Their CLL had likely already undergone multiple subclonal sweeps between diagnosis and alloSCT. This serves as a reminder that a patient's clinical cytogenetic and NGS disease profile largely reflects the characteristics of the dominant disease clone at a single time point but misses the subclonal heterogeneity that exists below the limit of detection. The post-treatment disease may differ substantially from the pre-treatment disease regardless of the number of prior therapies administered and how 'refractory' the disease has become. Longitudinal profiling throughout a patient's disease course is required to understand the cancer being treated.

Next, we hypothesized that any observed subclonal evolution post-alloSCT would be the result of selective immune pressure from donor T cells. We studied refractory and relapsed post-transplant CLL expecting to find immune escape variant subclones containing somatic alterations in genes and pathways implicated in immunotherapy resistance (interferon signaling, antigen presentation, immune checkpoints, etc.). Within the WES data from alloSCT-resistant disease, we did not find evidence of convergence towards altered immune pathways. The exonic alterations seen in the relapsed CLL subclones included known recurrent CLL driver lesions (similar to those that emerged post-chemotherapy), likely passenger mutations, and a diverse set of private novel mutations that were not shared by other patients. This was the same conclusion reached in the referenced AML alloSCT and melanoma nivolumab cohorts. There are multiple potential explanations for this result. It is possible that the CLL nonresponder cohort was too small to enable detection of an immune evasion signature within the WES data, especially given the considerable interpatient CLL molecular heterogeneity. Alternatively, in light of the RNA-seq data from the AML alloSCT studies [118, 119], it is possible that alloSCT resistance in CLL is mediated by changes in gene expression and epigenetic mechanisms rather than through somatic exonic alterations. For example, post-transplant downregulation of MHC class II molecules may also be implicated in CLL, a possibility that has not been investigated. This does not diminish the importance of CLL subclonal evolution as a mechanism

of transplant resistance, but it does highlight the potential importance of transcriptomic and epigenomic subclonal heterogeneity. Notably, the Landau lab has recently established a framework for this line of investigation [138]. It is possible that the imprint of selective immune pressure from grafted allogeneic T cells exists within the transcriptome/epigenome of relapsed CLL rather than in the exome. Lastly, a third possible explanation is that the leukemic evolution we observed was merely a result of “mass extinction” after alloSCT and that the driving force behind the emergence of new subclones was not immune selective pressure but rather competitive release of stochastic residual leukemic cells [139].

AlloSCT complete responders were added to the CLL cohort to increase the power to detect differences in CLL/T cell subclonal architecture that could predict transplant sensitivity/resistance. Unlike in melanoma, urothelial, and non-small cell lung cancers, tumor mutation burden and clonal neoantigen burden did not predict response to alloSCT immunotherapy. Here, it is important to point out the differences between the autologous and allogeneic settings. mHAs result from polymorphisms between the transplant recipient and donor genomes, the majority of which are SNPs. The SNPs within host CLL cells (and normal host cells) can be processed and presented within HLA-bound peptides and recognized as ‘non-self’ by donor T cells. Our inability to detect significant differences in the neoantigen load between response groups was unsurprising given that the mutation load of leukemia is low when compared to solid tumors [140]. In addition, there are more than 100 times as many mHAs, ‘allogeneic neoantigens,’ compared to somatic mutation-derived CLL neoantigens in a given related donor-recipient pair (> 200 times as many for MUDs) [141]. Possibly for the same reason, we did not observe a reduction in CLL neoantigens between pre- and post-alloSCT disease, an example of immunoediting that has been observed in solid tumors [120, 133, 142].

Differences in structural heterogeneity were observed between CLL alloSCT responders and nonresponders. Responders tended to have more structurally ‘pure’ or clonal disease as estimated by the Sequenza bioinformatics tool [87]. This result is preliminary and merits further

investigation. The studies of post-alloSCT AML did not report data regarding structural heterogeneity. If supported by additional evidence, this finding is counterintuitive. In general, alloSCT is a treatment that is reserved for the most aggressive, refractory lymphomas and leukemias. For example, it is the only treatment known to induce durable remissions in poor-risk AML, therapy-related AML, and Richter transformation of CLL [143-145]. Additionally, *TP53* alterations, which function to increase genome instability [146], predict worse outcomes in the context of chemotherapy for CLL [147]; however, alloSCT success is independent of *TP53* status (a finding supported by data from our cohort) [50]. Moreover, the Pyclone algorithm [148], which estimates tumor clonal structure from somatic mutation data, detected no difference in the number of pre-transplant CLL subclones within our cohort between responders and nonresponders. Further investigation of structural heterogeneity is needed to determine its relevance to alloSCT outcomes. In summary, there is ongoing genetic subclonal evolution of CLL post-alloSCT; however, we did not observe convergence towards any somatically altered immune genes or pathways within the exomes of transplant relapsed/refractory subclones.

We then focused on the post-transplant T cell repertoires. Post-transplant T cell reconstitution is a fascinating and complex process that has been studied for decades. The original methods used to study post-alloSCT T cell repertoires were TCR CDR3 spectratyping or flow cytometry with Vbeta chain antibodies [149-158]. Over the past decade, NGS combined with either 5'RACE or multiplex nested PCR has become the preferred technique as it yields far greater repertoire resolution, including the precise TCR alpha/beta chain clonotype sequences owing to high sequencing fidelity at the CDR3 region [93-95, 159-165]. The peripheral blood TCR repertoire in a healthy adult is vast comprising $\sim 10^7$ unique TCR beta chains on $\sim 10^{12}$ circulating T cells [166]. This has led to sampling bias and computational challenges, yielding conflicting results but several common themes have emerged.

First, studies using old and new methods agree that the post-alloSCT T cell repertoire is severely restricted in its diversity. Despite tremendous interpatient heterogeneity, for nearly all

patients, the post-transplant repertoire remains oligoclonal one-year post-alloSCT and, for many patients (greater than 50% in one study), the T cell repertoire remains abnormal for years after transplant [154, 160, 162, 163, 167]. Second, by absolute number, T cells recover relatively quickly post-transplant (by ~ 30 days), but this is driven by CD8 T cells and CD4 T cells lag behind. The normal ~ 2:1 ratio of CD4:CD8 T cells is sharply reduced and even inverted in many patients post-alloSCT [95, 160, 164]. Interestingly, the CD4 T cell compartment in healthy adults is more diverse than the CD8 compartment, a difference that is more pronounced in alloSCT recipients [93, 117, 168]. Though the post-alloSCT T cell repertoire has not been studied in CLL, we observed similar repertoire characteristics in the patients from our cohort including very restricted repertoires up to 5 years post-transplant and increased diversity among CD4 versus CD8 T cells.

There is less agreement among studies about the evolution of the repertoire post-transplant and its overlap with the input donor repertoire apart from the wide variation seen between patients. In general, the diversity and composition of the recipient repertoire matches that of the donor very early post-transplant (~ day +15) and then becomes more oligoclonal and less similar to the donor until day +100 [161]. From there, there is slow improvement in repertoire diversity in a subset of patients between 6-12 months and especially after 12 months [93, 95, 154]. Some studies argue that the overlap between the input repertoire and recipient repertoire is minimal at the 1 year mark post-transplant [94] while others report that the donor/recipient repertoires have peak similarity between 6-12 months with a decline thereafter [160]. The mechanisms responsible for the post-transplant T cell repertoires are homeostatic peripheral expansion (HPE) and thymopoiesis [103, 169]. HPE predominates in the first year after transplant and consists of proliferating mature donor T cells that survive the conditioning regimen. There is disagreement as to whether the cells immediately post-transplant derive from the donor naïve or memory compartments; however, there is greater consensus that the donor T cells that ultimately persist in the recipient originate from the donor memory compartment

[160, 164, 170]. This is likely influenced by the type of GVHD prophylaxis administered and both groups of cells have been identified post-transplant. The phenotype that those donor-derived cells assume in the recipient and their participation in GVL/GVHD has also been investigated [171-173]. Reconstituting T cells are skewed towards an antigen-experienced, effector memory phenotype [110]. Thymopoiesis slowly ramps up between 6-12 months post-alloSCT and can generate a new, diverse repertoire of naïve T cells; however, T-Cell Receptor Excision Circle (TREC) analyses demonstrate that thymic output is affected by many factors including donor/recipient age, GVHD, conditioning regimen, and CMV reactivation [153, 160, 163, 174]. Due to the retrospective nature of our study, we were unable to assess post-alloSCT T cell reconstitution at uniform time points; however, with some notable exceptions, we did not observe an increase in T cell repertoire diversity over time in our CLL cohort. Possible explanations include poor thymic output, CMV reactivation, and GVHD (discussed below).

Major factors that consistently affect post-alloSCT T cell reconstitution include CMV reactivation and GVHD – both acute and chronic – though chronic GVHD is far less studied in this context. CMV serostatus and reactivation have been a particular focus of multiple elegant studies. CMV is known to have a unique and seemingly oversized impact on even the healthy adult T cell repertoire [100, 175, 176]. In the alloSCT setting, CMV reactivation critically affects the recipient repertoire, possibly even resetting the trajectory of CD8 T cell reconstitution indefinitely [94, 95, 160, 177]. CMV reactivation triggers massive clonal expansions of donor T cells within the CD8 T effector memory subset (in both public and private TCR clonotypes) thereby reducing the diversity of the CD8 T cell compartment [95]. Acute GVHD, too, functions to impair T cell reconstitution, in part due to thymic damage. Donor CD8 T cell clonal expansions are also seen in acute GVHD and most studies associate acute GVHD with reduced repertoire diversity [149, 150, 154, 159, 161]. Importantly, though, T cells that mediate GVHD traffic to tissue, and so the T cell clones undergoing contemporaneous expansion in the peripheral blood may be distinct from those causing damage in the GVHD target organs [35].

Within the CLL cohort, we did observe reduced T cell repertoire diversity among treatment responsive patients with chronic GVHD compared to those without chronic GVHD. B cells are also known to be important in the pathogenesis of chronic GVHD, however, we did not assess B cell diversity in our study [178]. Within the CLL cohort, 14/24 patients (7 alloSCT nonresponders and 7 responders) experienced CMV reactivation. In our cohort, T cell samples were limited in patients who did not experience CMV reactivation; however, we did not observe an association between repertoire clonality and CMV reactivation, likely due to insufficient sample size.

Finally, several of the most elegant T cell reconstitution studies have failed to address the relationship between repertoire reconstitution kinetics and leukemia relapse. With the exception of one report, which associated alloSCT response with higher T cell repertoire diversity among non-GVHD patients (n=10) [159], those investigators who have examined this relationship have found no association [94, 161]. We, too, did not identify a signature of the GVL effect within the post-alloSCT T cell repertoires of CLL patients in our cohort. A study of immunosequencing in CMV [100] cautions that many more patients will be needed to reach sufficient power to detect a GVL signal in the context of important variables like GVHD and CMV reactivation. In detecting GVL characteristics, we also lack certain advantages inherent to other studies: (i) tumor antigens are less potent than viral antigens [108] and (ii) since leukemia occupies the same space as the bone marrow/peripheral blood, there is not a separate solid tumor microenvironment to enrich for tumor-specific T cells [179]. More sophisticated computational techniques than those utilized in our CLL cohort will also be required. While it is unlikely that large numbers of public TCRs exist for leukemic antigens [180, 181], this author hypothesizes that distinct features (e.g. pre-existing frequency/phenotype in the input donor repertoire, min/max frequency in recipient, timing of expansion/contraction in relation to clinical response/remission, stereotyped TCR motifs across HLA matches, pre-/post-response T cell phenotype, etc.) of the GVL effect are discernible in the context of a dynamic, restricted post-transplant T cell repertoire, a hypothesis discussed in the 'Future Directions' paragraph below.

The last phase of this project involved identifying individual candidate GVL clones. Here, we provide a strategy and proof of concept. Tumor-specific neoantigens are centrally important in anti-tumor immunity [44, 133]. Breathtaking cancer responses in patients with advanced, refractory solid tumors have been achieved through adoptive transfer of neoantigen-specific T cells [129, 182, 183]. The potency of checkpoint blockade, too, is partially rooted in autologous T cell recognition of neoantigens [184, 185]. The vast majority of neoantigens are patient-specific since they derive from a unique set of driver/passenger exonic somatic mutations in the context of each patient's array of MHC molecules; however, a subset of neoantigens are shared, including the HLA*02:01 neoepitope, HMTEVVRHC, in *TP53* (p.R175H) [186]. Improvements in neoantigen prediction algorithms and techniques used to isolate neoantigen-specific T cells continue to accelerate discovery [187]. To-date, only a small number of studies have focused on leukemic neoantigens though this area is likely to grow.

There was one report that considered neoantigens in CLL. Rajasagi et al. studied 2 CLL patients that experienced durable remissions post-alloSCT. They identified candidate GVL T cells specific for neoepitopes in *ALMS1* and *C6ORF89* (patient 1) and *FNDC3B* (patient 2) after *in vitro* T cell restimulation with peptide pools [39]. In patient 2, a longitudinal analysis was performed and candidate neoantigen T cells (*FNDC3B* (VVMSWAPPV) in HLA*02:01) were identified at a time point 6 months post-alloSCT that coincided with disease molecular remission. The frequency of these cells in the bulk repertoire at that time point is difficult to infer because the T cells underwent a 2-week *in vitro* restimulation prior to analysis; however, after restimulation they were present in 0.05% of CD8 T cells by ELISPOT. In ALL, investigators found that pediatric patients generated robust autologous anti-neoantigen immunity, including against a shared neoantigen from the *ETV6-RUNX1* fusion [188]. Finally in AML, candidate anti-leukemic T cells targeting neoepitopes in nucleophosmin 1 (*NPM1*) [41, 189] and in the *CBFB-MYH11* fusion protein [40] have been identified. A major limitation of the AML studies is that the investigators focused on expanding and testing anti-neoantigen T cells from healthy donors.

The relevance of these antigens *in vivo* in either the autologous or the allogeneic settings remains to be seen. There is considerably more evidence supporting the relevance of LAAs and mHAs in GVL. For example, post-transplant leukemia responses have been associated with T cells targeting the LAAs PR1 and WT1 [190-193] and with T cells targeting mHAs, including UTA2-1 [194-196].

Our study of CLL neoantigen-specific T cells was limited by patient number and sample availability. Despite this, we were able to confidently map candidate neoantigen-specific T cell clonotypes onto the longitudinal post-alloSCT T cell repertoires of CLL patients 8 and 12, a novel strategy. In addition, we found putative anti-PARPBP and anti-ACTN1 T cells present in expanded frequencies at the time of clinical leukemia response (CLL patient 12) and subclonal contraction of ACTN1-containing tumor cells (CLL patient 8). Given the paucity of existing data for neoantigen-specific T cell responses in the post-alloSCT setting, it is challenging to put our results in context. There are many questions to answer regarding post-transplant neoantigen-specific T cell responses: (i) what degree of expansion is expected and when? (ii) how many neoantigens mediate clinical response? (iii) what is the relative importance of leukemic neoantigens vs. mHAs vs. LAAs? (iv) how quickly does the population contract after total leukemic or subclonal leukemic response? (v) at what frequency do the clonotypes persist long-term? (vi) what is the phenotype of the cells over time (donor into recipient) and how many are recruited to memory? (vii) do they become tolerized or exhausted over time? (viii) how many T cell clonotypes (Vbeta sequences) participate in any given anti-neoantigen response and should we expect T cell clonal dominance for a strong leukemic neoantigen? Variability in anti-tumor immune responses is likely the rule, but other studies offer important insight.

For example, Chapuis et al. treated 11 relapsed or high-risk AML patients with WT1-specific (RMFPNAPYL/HLA-A*0201) post-transplant T cell adoptive therapy and achieved promising clinical results [197]. The median peak frequency of WT1-specific cells (as a percentage of CD8 T cells) between 24-72h post-infusion was 3.1% in patients supplemented

with IL-2 and was 1.5% at 14 days. The cells preferentially trafficked to the bone marrow. In a subset of patients treated with IL-21 cytokine support, antigen-specific cells persisted beyond 1 year at a frequency greater than 0.05% and acquired a memory phenotype. A more recent clinical trial further supports the clinical promise of WT1-specific T cells [192]. In a separate study of adult T-cell leukemia, which is associated with the human T-cell lymphotropic virus type 1 (HTLV-1), T cells specific for an epitope (SFHSLHLLF/HLA-A*2402) within an HTLV-1 protein, known as HTLV-1 Tax, were investigated [198]. In post-alloSCT samples, donor-derived anti-Tax T cells were present at frequencies of 0.09%, 0.03%, and 6.35% in patients who experienced durable remissions. In addition, the anti-Tax T cell clonotypes contained a CDR3 stereotyped motif, either 'P-D' or 'P-R'.

Finally, Dr. Jeffrey J. Molldrem et al. conducted an elegant study of antigen-specific T cells in patients with CML [193]. The presence of T cells specific for the LAA PR1 (VLQELNVTV/HLA-A*0201) correlated strongly with response to both alloSCT and interferon treatment. Of the 9 alloSCT patients studied, 6 had detectable PR1-specific CD8 T cells present at frequencies between 0.38–12.8% (median 1.29% of CD8 subset). The cells persisted at detectable frequencies in multiple patients and were functional against CML blasts. These studies highlight the variability in the leukemia-antigen specific T cell response between patients and antigens but demonstrate that expansions to > 1% and even >10% of the CD8 T cell repertoire, as seen in CLL patients 8 and 12, are possible as is long-lived memory. Last, the major liability of our otherwise robust single cell TCR seq and mapping strategy is the tetramer. Other investigators have shared this concern [199]. The frequency of a given T cell clonotype in a tetramer sort (as a percentage of CD8 T cells) and in the bulk TCR repertoire (as a percentage of CD8 or even CD3 T cells) should match or at least correlate closely. For some of the singly sorted neoantigen tetramer-binding clones, the frequency in the bulk repertoire was 10 times higher than that in the tetramer gate. This raises the strong possibility that some of the cells in the tetramer gate were not antigen-specific. Potential solutions to this problem include

the use of dual tetramers (same antigen in 2 colors), a method used by other groups [200], and novel technologies including tetramer display on nanoparticles [187].

Our study has several other limitations. The primary limitation of all tumor immune studies is that, at present, it is not possible to determine the cognate antigen from the sequence or structure of a given T-cell clonotype. Many investigators are working from various angles towards this advance, which promises to revolutionize the field [201, 202]. Second, gene expression differences that contributed to the altered relapse phenotypes were outside the scope of our study. This is especially vexing given the fascinating data regarding MHC class II downregulation in post-alloSCT AML. One low-tech solution would be to analyze MHC class II expression by IHC in pre- and post-transplant FFPE bone marrow specimens from CLL patients in this cohort since most of the cryopreserved leukemia samples have been depleted. Third, our study was retrospective in nature but ambitious in scope. Our small cohort size limited our ability to detect differences between transplant responders and nonresponders, especially in light of the significant interpatient heterogeneity in CLL. Moreover, outcome endpoints in alloSCT (OS, FFS, etc.) are notoriously difficult [203]. Sample availability and viability at desired time points also posed challenges. Intratumoral heterogeneity is a critical feature of tumors, but there is not yet a precise way to measure it. Similarly, more sophisticated tools for GVL-specific T cell repertoire analyses are needed. The GVL relevant fraction of the repertoire is likely in the low- to mid-frequency range, but current methods give higher weight to the most and least frequent clonotypes. Computational approaches that focus on discrete clonal expansions rather than the overall repertoire shape (i.e. clonality) may be more useful [108]. Last, functional assays of our candidate GVL T cell clones are needed. Nevertheless, the strengths of our study included purified CLL samples, longitudinal time points, simultaneous malignant/immune compartment analyses, available germline and alloSCT donor samples, reproducibility of tumor and T cell sequencing techniques, detailed clinical information, and a robust TCR mapping strategy. These

were sufficient to define genetic changes in the leukemic subclonal architecture after transplant and to propose a novel method to investigate the GVL effect.

There are many paths to take from here. There are currently no published studies addressing the impact of intratumoral heterogeneity, genetic or epigenetic, on alloSCT efficacy. This is an important knowledge gap to fill and the Sequenza data from our small patient cohort may offer a clue. Next, we hypothesize that there is a signature of GVL within longitudinal TCR repertoire data of responding alloSCT patients. If true, the challenge lies in unraveling it from confounders like GVHD and moderators like CMV. Moreover, GVHD is linked to GVL through mHAs, though as with neoantigens, the quality of mHAs is likely as important as the quantity [195]. Large numbers of longitudinal post-alloSCT T cell repertoires will be required, stratified by at minimum, GVL vs. no GVL, GVHD vs. no GVHD (and grade), and CMV reactivation vs. no CMV reactivation (and peak viral load) with accurate clinical annotation. All analyses must account for HLA type. It would be helpful to have the input donor repertoire since any GVL clone would be expected to derive from the low frequency, diverse naïve T cell pool. Though they were not a focus of our study, the post-alloSCT MHC class II downregulation data (class I expression did not change) remind us of the potential importance of GVL CD4 T cells. CD4 T cells participate in tumor immunity in the autologous setting through both direct and indirect mechanisms [204]. With regard to potential CD4 T cell neoantigens, class II antigen prediction algorithms are notoriously inaccurate, but revision is ongoing [205]. Reconstitution of the post-alloSCT CD8 and CD4 compartments differ with CD4 T cells exhibiting greater diversity but poorer overall recovery. In light of this, at least one longitudinal sample per patient should be sort-purified so that CD4/8 T cells can be separately analyzed. In this study, we primarily utilized the bulk repertoires as a map on which to place our tetramer-sorted T cell clonotypes. A reverse strategy could also be useful. For example, candidate GVL clonotypes identified from the bulk repertoire analyses could be singly sorted from companion samples using Vbeta antibodies if the clones exist at a reasonable frequency. The sorted cells will be polyclonal, but the clone of

interest can be uniquely identified by its Vbeta sequence. In this way, the phenotype of a relevant clone can be ascertained along with its Valpha sequence for TCR transduction and *in vitro* testing. In addition, mass spectrometry and RNA-seq data could be incorporated into a filter for ranking candidate neoantigens [187]. In this context, if samples allow, studying post-alloSCT T cell repertoires in CML patients may be informative because response rates are high owing to potent GVL and leukemia antigens have been described (PR1,BCR-ABL) [193, 206]. This would again invite the criticism that alloSCT in CML is largely historical, but there is much to learn about GVL and model systems are important. Finally, TCRs from important GVL clones could have therapeutic potential in adoptive T cell therapy.

In conclusion, our study showed that there is ongoing CLL subclonal genetic evolution after HLA-matched, nonmyeloablative alloSCT, including in known CLL drivers. The significant molecular shifts that occurred coincident with DLI and in the context of a previously administered, nonmyeloablative conditioning regimen argue that immune selective pressure underlies the branched and linear patterns of disease evolution that we observed. Despite this, we did not detect an exonic immunoevasive signature in the post-transplant expanded subclones, which is consistent with studies of post-transplant relapsed AML that suggest an epigenetic mechanism of escape. Our parallel study of post-transplant T cell repertoires in the same CLL cohort confirms that the input healthy donor repertoire becomes dramatically restricted upon reconstitution in the host though it remains dynamic. Additionally, we provide a strategy to track candidate GVL T cell clones through their expansion and contraction within the T cell repertoire during periods of leukemia response. A deeper understanding of the mechanisms by which heterogeneous leukemic subclones are susceptible or resistant to allogeneic T cell killing may improve patient stratification for alloSCT/DLI and account for the uniquely durable transplant responses that occur in chemorefractory patients.

Bibliography

1. Greaves, M. and C.C. Maley, *Clonal evolution in cancer*. Nature, 2012. **481**(7381): p. 306-13.
2. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation*. Cell, 2011. **144**(5): p. 646-74.
3. Gerlinger, M. and C. Swanton, *How Darwinian models inform therapeutic failure initiated by clonal heterogeneity in cancer medicine*. Br J Cancer, 2010. **103**(8): p. 1139-43.
4. Merlo, L.M., J.W. Pepper, B.J. Reid, and C.C. Maley, *Cancer as an evolutionary and ecological process*. Nat Rev Cancer, 2006. **6**(12): p. 924-35.
5. Sottoriva, A., I. Spiteri, S.G. Piccirillo, A. Touloumis, V.P. Collins, J.C. Marioni, C. Curtis, C. Watts, and S. Tavaré, *Intratumor heterogeneity in human glioblastoma reflects cancer evolutionary dynamics*. Proc Natl Acad Sci U S A, 2013. **110**(10): p. 4009-14.
6. Kim, H., S. Zheng, S.S. Amini, S.M. Virk, T. Mikkelsen, D.J. Brat, J. Grimsby, C. Sougnez, F. Muller, J. Hu, A.E. Sloan, M.L. Cohen, E.G. Van Meir, L. Scarpance, P.W. Laird, J.N. Weinstein, E.S. Lander, S. Gabriel, G. Getz, M. Meyerson, L. Chin, J.S. Barnholtz-Sloan, and R.G. Verhaak, *Whole-genome and multisector exome sequencing of primary and post-treatment glioblastoma reveals patterns of tumor evolution*. Genome Res, 2015. **25**(3): p. 316-27.
7. Ding, L., T.J. Ley, D.E. Larson, C.A. Miller, D.C. Koboldt, J.S. Welch, J.K. Ritchey, M.A. Young, T. Lamprecht, M.D. McLellan, J.F. McMichael, J.W. Wallis, C. Lu, D. Shen, C.C. Harris, D.J. Dooling, R.S. Fulton, L.L. Fulton, K. Chen, H. Schmidt, J. Kalicki-Veizer, V.J. Magrini, L. Cook, S.D. McGrath, T.L. Vickery, M.C. Wendl, S. Heath, M.A. Watson, D.C. Link, M.H. Tomasson, W.D. Shannon, J.E. Payton, S. Kulkarni, P. Westervelt, M.J. Walter, T.A. Graubert, E.R. Mardis, R.K. Wilson, and J.F. DiPersio, *Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing*. Nature, 2012. **481**(7382): p. 506-10.

8. Sotillo, E., D.M. Barrett, K.L. Black, A. Bagashev, D. Oldridge, G. Wu, R. Sussman, C. Lanauze, M. Ruella, M.R. Gazzara, N.M. Martinez, C.T. Harrington, E.Y. Chung, J. Perazzelli, T.J. Hofmann, S.L. Maude, P. Raman, A. Barrera, S. Gill, S.F. Lacey, J.J. Melenhorst, D. Allman, E. Jacoby, T. Fry, C. Mackall, Y. Barash, K.W. Lynch, J.M. Maris, S.A. Grupp, and A. Thomas-Tikhonenko, *Convergence of Acquired Mutations and Alternative Splicing of CD19 Enables Resistance to CART-19 Immunotherapy*. *Cancer Discov*, 2015. **5**(12): p. 1282-95.
9. Diaz, L.A., Jr., R.T. Williams, J. Wu, I. Kinde, J.R. Hecht, J. Berlin, B. Allen, I. Bozic, J.G. Reiter, M.A. Nowak, K.W. Kinzler, K.S. Oliner, and B. Vogelstein, *The molecular evolution of acquired resistance to targeted EGFR blockade in colorectal cancers*. *Nature*, 2012. **486**(7404): p. 537-40.
10. Strati, P. and T.D. Shanafelt, *Monoclonal B-cell lymphocytosis and early-stage chronic lymphocytic leukemia: diagnosis, natural history, and risk stratification*. *Blood*, 2015. **126**(4): p. 454-62.
11. Seifert, M., L. Sellmann, J. Bloehdorn, F. Wein, S. Stilgenbauer, J. Durig, and R. Kupperts, *Cellular origin and pathophysiology of chronic lymphocytic leukemia*. *J Exp Med*, 2012. **209**(12): p. 2183-98.
12. Kikushige, Y., F. Ishikawa, T. Miyamoto, T. Shima, S. Urata, G. Yoshimoto, Y. Mori, T. Iino, T. Yamauchi, T. Eto, H. Niino, H. Iwasaki, K. Takenaka, and K. Akashi, *Self-renewing hematopoietic stem cell is the primary target in pathogenesis of human chronic lymphocytic leukemia*. *Cancer Cell*, 2011. **20**(2): p. 246-59.
13. Hamblin, T.J., Z. Davis, A. Gardiner, D.G. Oscier, and F.K. Stevenson, *Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia*. *Blood*, 1999. **94**(6): p. 1848-54.
14. Munir, T., J.R. Brown, S. O'Brien, J.C. Barrientos, P.M. Barr, N.M. Reddy, S. Coutre, C.S. Tam, S.P. Mulligan, U. Jaeger, T.J. Kipps, C. Moreno, M. Montillo, J.A. Burger, J.C. Byrd,

- P. Hillmen, S. Dai, A. Szoke, J.P. Dean, and J.A. Woyach, *Final analysis from RESONATE: Up to six years of follow-up on ibrutinib in patients with previously treated chronic lymphocytic leukemia or small lymphocytic lymphoma*. *Am J Hematol*, 2019. **94**(12): p. 1353-1363.
15. Badoux, X.C., M.J. Keating, and W.G. Wierda, *What is the best frontline therapy for patients with CLL and 17p deletion?* *Curr Hematol Malig Rep*, 2011. **6**(1): p. 36-46.
 16. Jain, P., M.J. Keating, W.G. Wierda, M. Sivina, P.A. Thompson, A. Ferrajoli, Z. Estrov, H. Kantarjian, S. O'Brien, and J.A. Burger, *Long-term Follow-up of Treatment with Ibrutinib and Rituximab in Patients with High-Risk Chronic Lymphocytic Leukemia*. *Clin Cancer Res*, 2017. **23**(9): p. 2154-2158.
 17. Landau, D.A., E. Tausch, A.N. Taylor-Weiner, C. Stewart, J.G. Reiter, J. Bahlo, S. Kluth, I. Bozic, M. Lawrence, S. Bottcher, S.L. Carter, K. Cibulskis, D. Mertens, C.L. Sougnez, M. Rosenberg, J.M. Hess, J. Edelman, S. Kless, M. Kneba, M. Ritgen, A. Fink, K. Fischer, S. Gabriel, E.S. Lander, M.A. Nowak, H. Dohner, M. Hallek, D. Neuberg, G. Getz, S. Stilgenbauer, and C.J. Wu, *Mutations driving CLL and their evolution in progression and relapse*. *Nature*, 2015. **526**(7574): p. 525-30.
 18. *Chemotherapeutic options in chronic lymphocytic leukemia: a meta-analysis of the randomized trials*. *CLL Trialists' Collaborative Group*. *J Natl Cancer Inst*, 1999. **91**(10): p. 861-8.
 19. Burger, J.A., P.M. Barr, T. Robak, C. Owen, P. Ghia, A. Tedeschi, O. Bairey, P. Hillmen, S.E. Coutre, S. Devereux, S. Grosicki, H. McCarthy, D. Simpson, F. Offner, C. Moreno, S. Dai, I. Lal, J.P. Dean, and T.J. Kipps, *Long-term efficacy and safety of first-line ibrutinib treatment for patients with CLL/SLL: 5 years of follow-up from the phase 3 RESONATE-2 study*. *Leukemia*, 2020. **34**(3): p. 787-798.
 20. Burger, J.A., D.A. Landau, A. Taylor-Weiner, I. Bozic, H. Zhang, K. Sarosiek, L. Wang, C. Stewart, J. Fan, J. Hoellenriegel, M. Sivina, A.M. Dubuc, C. Fraser, Y. Han, S. Li, K.J.

- Livak, L. Zou, Y. Wan, S. Konoplev, C. Sougnez, J.R. Brown, L.V. Abruzzo, S.L. Carter, M.J. Keating, M.S. Davids, W.G. Wierda, K. Cibulskis, T. Zenz, L. Werner, P. Dal Cin, P. Kharchenko, D. Neuberg, H. Kantarjian, E. Lander, S. Gabriel, S. O'Brien, A. Letai, D.A. Weitz, M.A. Nowak, G. Getz, and C.J. Wu, *Clonal evolution in patients with chronic lymphocytic leukaemia developing resistance to BTK inhibition*. Nat Commun, 2016. **7**: p. 11589.
21. Singh, A.K. and J.P. McGuirk, *Allogeneic Stem Cell Transplantation: A Historical and Scientific Overview*. Cancer Res, 2016. **76**(22): p. 6445-6451.
 22. Thomas, E.D., C.D. Buckner, R.H. Rudolph, A. Fefer, R. Storb, P.E. Neiman, J.I. Bryant, R.L. Chard, R.A. Clift, R.B. Epstein, P.J. Fialkow, D.D. Funk, E.R. Giblett, K.G. Lerner, F.A. Reynolds, and S. Slichter, *Allogeneic marrow grafting for hematologic malignancy using HL-A matched donor-recipient sibling pairs*. Blood, 1971. **38**(3): p. 267-87.
 23. Weiden, P.L., N. Flournoy, E.D. Thomas, R. Prentice, A. Fefer, C.D. Buckner, and R. Storb, *Antileukemic effect of graft-versus-host disease in human recipients of allogeneic-marrow grafts*. N Engl J Med, 1979. **300**(19): p. 1068-73.
 24. Collins, R.H., Jr., Z.R. Rogers, M. Bennett, V. Kumar, A. Nikein, and J.W. Fay, *Hematologic relapse of chronic myelogenous leukemia following allogeneic bone marrow transplantation: apparent graft-versus-leukemia effect following abrupt discontinuation of immunosuppression*. Bone Marrow Transplant, 1992. **10**(4): p. 391-5.
 25. Elmaagacli, A.H., D.W. Beelen, G. Trenn, O. Schmidt, M. Nahler, and U.W. Schaefer, *Induction of a graft-versus-leukemia reaction by cyclosporin A withdrawal as immunotherapy for leukemia relapsing after allogeneic bone marrow transplantation*. Bone Marrow Transplant, 1999. **23**(8): p. 771-7.
 26. Higano, C.S., M. Brixey, E.M. Bryant, D.M. Durnam, K. Doney, K.M. Sullivan, and J.W. Singer, *Durable complete remission of acute nonlymphocytic leukemia associated with discontinuation of immunosuppression following relapse after allogeneic bone marrow*

- transplantation. A case report of a probable graft-versus-leukemia effect. Transplantation, 1990. 50(1): p. 175-7.*
27. Apperley, J.F., L. Jones, G. Hale, H. Waldmann, J. Hows, Y. Rombos, C. Tsatalas, R.E. Marcus, A.W. Goolden, E.C. Gordon-Smith, and et al., *Bone marrow transplantation for patients with chronic myeloid leukaemia: T-cell depletion with Campath-1 reduces the incidence of graft-versus-host disease but may increase the risk of leukaemic relapse. Bone Marrow Transplant, 1986. 1(1): p. 53-66.*
 28. Horowitz, M.M., R.P. Gale, P.M. Sondel, J.M. Goldman, J. Kersey, H.J. Kolb, A.A. Rimm, O. Ringden, C. Rozman, B. Speck, and et al., *Graft-versus-leukemia reactions after bone marrow transplantation. Blood, 1990. 75(3): p. 555-62.*
 29. Wagner, J.E., J.S. Thompson, S.L. Carter, N.A. Kernan, and T. Unrelated Donor Marrow Transplantation, *Effect of graft-versus-host disease prophylaxis on 3-year disease-free survival in recipients of unrelated donor bone marrow (T-cell Depletion Trial): a multi-centre, randomised phase II-III trial. Lancet, 2005. 366(9487): p. 733-41.*
 30. Marmont, A.M., M.M. Horowitz, R.P. Gale, K. Sobocinski, R.C. Ash, D.W. van Bekkum, R.E. Champlin, K.A. Dicke, J.M. Goldman, R.A. Good, and et al., *T-cell depletion of HLA-identical transplants in leukemia. Blood, 1991. 78(8): p. 2120-30.*
 31. Tomblyn, M. and H.M. Lazarus, *Donor lymphocyte infusions: the long and winding road: how should it be traveled? Bone Marrow Transplant, 2008. 42(9): p. 569-79.*
 32. Kolb, H.J., *Graft-versus-leukemia effects of transplantation and donor lymphocytes. Blood, 2008. 112(12): p. 4371-83.*
 33. Collins, R.H., Jr., S. Goldstein, S. Giralt, J. Levine, D. Porter, W. Drobyski, J. Barrett, M. Johnson, A. Kirk, M. Horowitz, and P. Parker, *Donor leukocyte infusions in acute lymphocytic leukemia. Bone Marrow Transplant, 2000. 26(5): p. 511-6.*
 34. Collins, R.H., Jr., O. Shpilberg, W.R. Drobyski, D.L. Porter, S. Giralt, R. Champlin, S.A. Goodman, S.N. Wolff, W. Hu, C. Verfaillie, A. List, W. Dalton, N. Ognoskie, A. Chetrit, J.H.

- Antin, and J. Nemunaitis, *Donor leukocyte infusions in 140 patients with relapsed malignancy after allogeneic bone marrow transplantation*. J Clin Oncol, 1997. **15**(2): p. 433-44.
35. Kolb, H.J., A. Schattenberg, J.M. Goldman, B. Hertenstein, N. Jacobsen, W. Arcese, P. Ljungman, A. Ferrant, L. Verdonck, D. Niederwieser, F. van Rhee, J. Mittermueller, T. de Witte, E. Holler, H. Ansari, B. European Group for, and L. Marrow Transplantation Working Party Chronic, *Graft-versus-leukemia effect of donor lymphocyte transfusions in marrow grafted patients*. Blood, 1995. **86**(5): p. 2041-50.
36. Khouri, I.F., M.S. Lee, R.M. Saliba, B. Andersson, P. Anderlini, D. Couriel, C. Hosing, S. Giralt, M. Korbling, J. McMannis, M.J. Keating, and R.E. Champlin, *Nonablative allogeneic stem cell transplantation for chronic lymphocytic leukemia: impact of rituximab on immunomodulation and survival*. Exp Hematol, 2004. **32**(1): p. 28-35.
37. Levine, J.E., T. Braun, S.L. Penza, P. Beatty, K. Cornetta, R. Martino, W.R. Drobyski, A.J. Barrett, D.L. Porter, S. Giralt, J. Leis, H.E. Holmes, M. Johnson, M. Horowitz, and R.H. Collins, Jr., *Prospective trial of chemotherapy and donor leukocyte infusions for relapse of advanced myeloid malignancies after allogeneic stem-cell transplantation*. J Clin Oncol, 2002. **20**(2): p. 405-12.
38. Bleakley, M. and S.R. Riddell, *Molecules and mechanisms of the graft-versus-leukaemia effect*. Nat Rev Cancer, 2004. **4**(5): p. 371-80.
39. Rajasagi, M., S.A. Shukla, E.F. Fritsch, D.B. Keskin, D. DeLuca, E. Carmona, W. Zhang, C. Sougnez, K. Cibulskis, J. Sidney, K. Stevenson, J. Ritz, D. Neuberg, V. Brusic, S. Gabriel, E.S. Lander, G. Getz, N. Hacohen, and C.J. Wu, *Systematic identification of personal tumor-specific neoantigens in chronic lymphocytic leukemia*. Blood, 2014. **124**(3): p. 453-62.
40. Biernacki, M.A., K.A. Foster, K.B. Woodward, M.E. Coon, C. Cummings, T.M. Cunningham, R.G. Dossa, M. Brault, J. Stokke, T.M. Olsen, K. Gardner, E. Estey, S.

- Meshinchi, A. Rongvaux, and M. Bleakley, *CBFB-MYH11 fusion neoantigen enables T cell recognition and killing of acute myeloid leukemia*. J Clin Invest, 2020.
41. van der Lee, D.I., R.M. Reijmers, M.W. Honders, R.S. Hagedoorn, R.C. de Jong, M.G. Kester, D.M. van der Steen, A.H. de Ru, C. Kweekel, H.M. Bijen, I. Jedema, H. Veelken, P.A. van Veelen, M.H. Heemskerk, J.H.F. Falkenburg, and M. Griffioen, *Mutated nucleophosmin 1 as immunotherapy target in acute myeloid leukemia*. J Clin Invest, 2019. **129**(2): p. 774-785.
42. Anguille, S., V.F. Van Tendeloo, and Z.N. Berneman, *Leukemia-associated antigens and their relevance to the immunotherapy of acute myeloid leukemia*. Leukemia, 2012. **26**(10): p. 2186-96.
43. Biernacki, M.A. and M. Bleakley, *Neoantigens in Hematologic Malignancies*. Front Immunol, 2020. **11**: p. 121.
44. Schumacher, T.N. and R.D. Schreiber, *Neoantigens in cancer immunotherapy*. Science, 2015. **348**(6230): p. 69-74.
45. Chen, D.S. and I. Mellman, *Oncology meets immunology: the cancer-immunity cycle*. Immunity, 2013. **39**(1): p. 1-10.
46. Boyiadzis, M.M., M.V. Dhodapkar, R.J. Brentjens, J.N. Kochenderfer, S.S. Neelapu, M.V. Maus, D.L. Porter, D.G. Maloney, S.A. Grupp, C.L. Mackall, C.H. June, and M.R. Bishop, *Chimeric antigen receptor (CAR) T therapies for the treatment of hematologic malignancies: clinical perspective and significance*. J Immunother Cancer, 2018. **6**(1): p. 137.
47. Topalian, S.L., C.G. Drake, and D.M. Pardoll, *Immune checkpoint blockade: a common denominator approach to cancer therapy*. Cancer Cell, 2015. **27**(4): p. 450-61.
48. Roeker, L.E., P. Dreger, J.R. Brown, O.B. Lahoud, T.A. Eyre, D.M. Brander, A. Skarbnik, C.C. Coombs, H.T. Kim, M. Davids, S.T. Manchini, G. George, N. Shah, T.J. Voorhees, K.H. Orchard, H.S. Walter, A.K. Arumainathan, A. Sitlinger, J.H. Park, M.B. Geyer, A.D.

- Zelenetz, C.S. Sauter, S.A. Giralt, M.A. Perales, and A.R. Mato, *Allogeneic stem cell transplantation for chronic lymphocytic leukemia in the era of novel agents*. *Blood Adv*, 2020. **4**(16): p. 3977-3989.
49. Khouri, I.F., R. Bassett, N. Poindexter, S. O'Brien, C.E. Bueso-Ramos, Y. Hsu, A. Ferrajoli, M.J. Keating, R. Champlin, and M. Fernandez-Vina, *Nonmyeloablative allogeneic stem cell transplantation in relapsed/refractory chronic lymphocytic leukemia: long-term follow-up, prognostic factors, and effect of human leukocyte histocompatibility antigen subtype on outcome*. *Cancer*, 2011. **117**(20): p. 4679-88.
50. Dreger, P., A. Schnaiter, T. Zenz, S. Bottcher, M. Rossi, P. Paschka, A. Buhler, S. Dietrich, R. Busch, M. Ritgen, D. Bunjes, M. Zeis, M. Stadler, L. Uharek, C. Scheid, U. Hegenbart, M. Hallek, M. Kneba, N. Schmitz, H. Dohner, and S. Stilgenbauer, *TP53, SF3B1, and NOTCH1 mutations and outcome of allotransplantation for chronic lymphocytic leukemia: six-year follow-up of the GCLLSG CLL3X trial*. *Blood*, 2013. **121**(16): p. 3284-8.
51. Sorrow, M.L., B.E. Storer, B.M. Sandmaier, M. Maris, J. Shizuru, R. Maziarz, E. Agura, T.R. Chauncey, M.A. Pulsipher, P.A. McSweeney, J.C. Wade, B. Bruno, A. Langston, J. Radich, D. Niederwieser, K.G. Blume, R. Storb, and D.G. Maloney, *Five-year follow-up of patients with advanced chronic lymphocytic leukemia treated with allogeneic hematopoietic cell transplantation after nonmyeloablative conditioning*. *J Clin Oncol*, 2008. **26**(30): p. 4912-20.
52. Brown, J.R., H.T. Kim, P. Armand, C. Cutler, D.C. Fisher, V. Ho, J. Koreth, J. Ritz, C. Wu, J.H. Antin, R.J. Soiffer, J.G. Gribben, and E.P. Alyea, *Long-term follow-up of reduced-intensity allogeneic stem cell transplantation for chronic lymphocytic leukemia: prognostic model to predict outcome*. *Leukemia*, 2013. **27**(2): p. 362-9.
53. Parkin, B., P. Ouillette, Y. Li, J. Keller, C. Lam, D. Roulston, C. Li, K. Shedden, and S.N. Malek, *Clonal evolution and devolution after chemotherapy in adult acute myelogenous leukemia*. *Blood*, 2013. **121**(2): p. 369-77.

54. Welch, J.S., T.J. Ley, D.C. Link, C.A. Miller, D.E. Larson, D.C. Koboldt, L.D. Wartman, T.L. Lamprecht, F. Liu, J. Xia, C. Kandoth, R.S. Fulton, M.D. McLellan, D.J. Dooling, J.W. Wallis, K. Chen, C.C. Harris, H.K. Schmidt, J.M. Kalicki-Veizer, C. Lu, Q. Zhang, L. Lin, M.D. O'Laughlin, J.F. McMichael, K.D. Delehaunty, L.A. Fulton, V.J. Magrini, S.D. McGrath, R.T. Demeter, T.L. Vickery, J. Hundal, L.L. Cook, G.W. Swift, J.P. Reed, P.A. Alldredge, T.N. Wylie, J.R. Walker, M.A. Watson, S.E. Heath, W.D. Shannon, N. Varghese, R. Nagarajan, J.E. Payton, J.D. Baty, S. Kulkarni, J.M. Klco, M.H. Tomasson, P. Westervelt, M.J. Walter, T.A. Graubert, J.F. DiPersio, L. Ding, E.R. Mardis, and R.K. Wilson, *The origin and evolution of mutations in acute myeloid leukemia*. *Cell*, 2012. **150**(2): p. 264-78.
55. Smith, C.C., Q. Wang, C.S. Chin, S. Salerno, L.E. Damon, M.J. Levis, A.E. Perl, K.J. Travers, S. Wang, J.P. Hunt, P.P. Zarrinkar, E.E. Schadt, A. Kasarskis, J. Kuriyan, and N.P. Shah, *Validation of ITD mutations in FLT3 as a therapeutic target in human acute myeloid leukaemia*. *Nature*, 2012. **485**(7397): p. 260-3.
56. Landau, D.A., S.L. Carter, P. Stojanov, A. McKenna, K. Stevenson, M.S. Lawrence, C. Sougnez, C. Stewart, A. Sivachenko, L. Wang, Y. Wan, W. Zhang, S.A. Shukla, A. Vartanov, S.M. Fernandes, G. Saksena, K. Cibulskis, B. Tesar, S. Gabriel, N. Hacohen, M. Meyerson, E.S. Lander, D. Neuberg, J.R. Brown, G. Getz, and C.J. Wu, *Evolution and impact of subclonal mutations in chronic lymphocytic leukemia*. *Cell*, 2013. **152**(4): p. 714-26.
57. Shukla, S.A., M.S. Rooney, M. Rajasagi, G. Tiao, P.M. Dixon, M.S. Lawrence, J. Stevens, W.J. Lane, J.L. Dellagatta, S. Steelman, C. Sougnez, K. Cibulskis, A. Kiezun, N. Hacohen, V. Brusica, C.J. Wu, and G. Getz, *Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes*. *Nat Biotechnol*, 2015. **33**(11): p. 1152-8.
58. Li, H. and R. Durbin, *Fast and accurate long-read alignment with Burrows-Wheeler transform*. *Bioinformatics*, 2010. **26**(5): p. 589-95.

59. Cibulskis, K., M.S. Lawrence, S.L. Carter, A. Sivachenko, D. Jaffe, C. Sougnez, S. Gabriel, M. Meyerson, E.S. Lander, and G. Getz, *Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples*. Nat Biotechnol, 2013. **31**(3): p. 213-9.
60. Gobbin, E., M. Ezzalfani, V. Dieras, T. Bachelot, E. Brain, M. Debled, W. Jacot, M.A. Mouret-Reynier, A. Goncalves, F. Dalenc, A. Patsouris, J.M. Ferrero, C. Levy, V. Lorgis, L. Vanlemmens, C. Lefeuvre-Plesse, S. Mathoulin-Pelissier, T. Petit, L. Uwer, C. Jouannaud, M. Leheurteur, M. Lacroix-Triki, A.L. Cleaud, M. Robain, C. Courtinard, C. Cailliot, D. Perol, and S. Delalogue, *Time trends of overall survival among metastatic breast cancer patients in the real-life ESME cohort*. Eur J Cancer, 2018. **96**: p. 17-24.
61. Genomes Project, C., A. Auton, L.D. Brooks, R.M. Durbin, E.P. Garrison, H.M. Kang, J.O. Korb, J.L. Marchini, S. McCarthy, G.A. McVean, and G.R. Abecasis, *A global reference for human genetic variation*. Nature, 2015. **526**(7571): p. 68-74.
62. Ye, K., M.H. Schulz, Q. Long, R. Apweiler, and Z. Ning, *Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads*. Bioinformatics, 2009. **25**(21): p. 2865-71.
63. Puente, X.S., S. Bea, R. Valdes-Mas, N. Villamor, J. Gutierrez-Abril, J.I. Martin-Subero, M. Munar, C. Rubio-Perez, P. Jares, M. Aymerich, T. Baumann, R. Beekman, L. Belver, A. Carrio, G. Castellano, G. Clot, E. Colado, D. Colomer, D. Costa, J. Delgado, A. Enjuanes, X. Estivill, A.A. Ferrando, J.L. Gelpi, B. Gonzalez, S. Gonzalez, M. Gonzalez, M. Gut, J.M. Hernandez-Rivas, M. Lopez-Guerra, D. Martin-Garcia, A. Navarro, P. Nicolas, M. Orozco, A.R. Payer, M. Pinyol, D.G. Pisano, D.A. Puente, A.C. Queiros, V. Quesada, C.M. Romeo-Casabona, C. Royo, R. Royo, M. Rozman, N. Russinol, I. Salaverria, K. Stamatopoulos, H.G. Stunnenberg, D. Tamborero, M.J. Terol, A. Valencia, N. Lopez-Bigas, D. Torrents, I. Gut, A. Lopez-Guillermo, C. Lopez-Otin, and E. Campo, *Non-coding recurrent mutations in chronic lymphocytic leukaemia*. Nature, 2015. **526**(7574): p. 519-24.

64. Robinson, J.T., H. Thorvaldsdottir, W. Winckler, M. Guttman, E.S. Lander, G. Getz, and J.P. Mesirov, *Integrative genomics viewer*. Nat Biotechnol, 2011. **29**(1): p. 24-6.
65. Wang, K., M. Li, and H. Hakonarson, *ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data*. Nucleic Acids Res, 2010. **38**(16): p. e164.
66. Rimmer, A., H. Phan, I. Mathieson, Z. Iqbal, S.R. Twigg, W.G.S. Consortium, A.O. Wilkie, G. McVean, and G. Lunter, *Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications*. Nat Genet, 2014. **46**(8): p. 912-8.
67. Olshen, A.B., E.S. Venkatraman, R. Lucito, and M. Wigler, *Circular binary segmentation for the analysis of array-based DNA copy number data*. Biostatistics, 2004. **5**(4): p. 557-72.
68. Sherry, S.T., M.H. Ward, M. Kholodov, J. Baker, L. Phan, E.M. Smigielski, and K. Sirotkin, *dbSNP: the NCBI database of genetic variation*. Nucleic Acids Res, 2001. **29**(1): p. 308-11.
69. Forbes, S.A., D. Beare, P. Gunasekaran, K. Leung, N. Bindal, H. Boutselakis, M. Ding, S. Bamford, C. Cole, S. Ward, C.Y. Kok, M. Jia, T. De, J.W. Teague, M.R. Stratton, U. McDermott, and P.J. Campbell, *COSMIC: exploring the world's knowledge of somatic mutations in human cancer*. Nucleic Acids Res, 2015. **43**(Database issue): p. D805-11.
70. Clifford, R., T. Louis, P. Robbe, S. Ackroyd, A. Burns, A.T. Timbs, G. Wright Colopy, H. Dreau, F. Sigaux, J.G. Judde, M. Rotger, A. Telenti, Y.L. Lin, P. Pasero, J. Maelfait, M. Titsias, D.R. Cohen, S.J. Henderson, M.T. Ross, D. Bentley, P. Hillmen, A. Pettitt, J. Rehwinkel, S.J. Knight, J.C. Taylor, Y.J. Crow, M. Benkirane, and A. Schuh, *SAMHD1 is mutated recurrently in chronic lymphocytic leukemia and is involved in response to DNA damage*. Blood, 2014. **123**(7): p. 1021-31.
71. Wang, L., M.S. Lawrence, Y. Wan, P. Stojanov, C. Sougnez, K. Stevenson, L. Werner, A. Sivachenko, D.S. DeLuca, L. Zhang, W. Zhang, A.R. Vartanov, S.M. Fernandes, N.R.

- Goldstein, E.G. Folco, K. Cibulskis, B. Tesar, Q.L. Sievers, E. Shefler, S. Gabriel, N. Hacohen, R. Reed, M. Meyerson, T.R. Golub, E.S. Lander, D. Neuberg, J.R. Brown, G. Getz, and C.J. Wu, *SF3B1 and other novel cancer genes in chronic lymphocytic leukemia*. N Engl J Med, 2011. **365**(26): p. 2497-506.
72. Quesada, V., L. Conde, N. Villamor, G.R. Ordonez, P. Jares, L. Bassaganyas, A.J. Ramsay, S. Bea, M. Pinyol, A. Martinez-Trillos, M. Lopez-Guerra, D. Colomer, A. Navarro, T. Baumann, M. Aymerich, M. Rozman, J. Delgado, E. Gine, J.M. Hernandez, M. Gonzalez-Diaz, D.A. Puente, G. Velasco, J.M. Freije, J.M. Tubio, R. Royo, J.L. Gelpi, M. Orozco, D.G. Pisano, J. Zamora, M. Vazquez, A. Valencia, H. Himmelbauer, M. Bayes, S. Heath, M. Gut, I. Gut, X. Estivill, A. Lopez-Guillermo, X.S. Puente, E. Campo, and C. Lopez-Otin, *Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia*. Nat Genet, 2011. **44**(1): p. 47-52.
73. Fabbri, G., S. Rasi, D. Rossi, V. Trifonov, H. Khiabani, J. Ma, A. Grunn, M. Fangazio, D. Capello, S. Monti, S. Cresta, E. Gargiulo, F. Forconi, A. Guarini, L. Arcaini, M. Paulli, L. Laurenti, L.M. Larocca, R. Marasca, V. Gattei, D. Oscier, F. Bertoni, C.G. Mullighan, R. Foa, L. Pasqualucci, R. Rabadan, R. Dalla-Favera, and G. Gaidano, *Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation*. J Exp Med, 2011. **208**(7): p. 1389-401.
74. Puente, X.S., M. Pinyol, V. Quesada, L. Conde, G.R. Ordonez, N. Villamor, G. Escaramis, P. Jares, S. Bea, M. Gonzalez-Diaz, L. Bassaganyas, T. Baumann, M. Juan, M. Lopez-Guerra, D. Colomer, J.M. Tubio, C. Lopez, A. Navarro, C. Tornador, M. Aymerich, M. Rozman, J.M. Hernandez, D.A. Puente, J.M. Freije, G. Velasco, A. Gutierrez-Fernandez, D. Costa, A. Carrio, S. Guijarro, A. Enjuanes, L. Hernandez, J. Yague, P. Nicolas, C.M. Romeo-Casabona, H. Himmelbauer, E. Castillo, J.C. Dohm, S. de Sanjose, M.A. Piris, E. de Alava, J. San Miguel, R. Royo, J.L. Gelpi, D. Torrents, M. Orozco, D.G. Pisano, A. Valencia, R. Guigo, M. Bayes, S. Heath, M. Gut, P. Klatt, J. Marshall, K. Raine, L.A.

- Stebbins, P.A. Futreal, M.R. Stratton, P.J. Campbell, I. Gut, A. Lopez-Guillermo, X. Estivill, E. Montserrat, C. Lopez-Otin, and E. Campo, *Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia*. *Nature*, 2011. **475**(7354): p. 101-5.
75. Laue, T.M. and S.L. Pelletier, *Interactive Computer-Aided Interpretation of Sedimentation Data*. *Biophysical Journal*, 1986. **49**(2): p. A274-A274.
76. Schuck, P., *On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation*. *Anal Biochem*, 2003. **320**(1): p. 104-24.
77. Vistica, J., J. Dam, A. Balbo, E. Yikilmaz, R.A. Mariuzza, T.A. Rouault, and P. Schuck, *Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition*. *Anal Biochem*, 2004. **326**(2): p. 234-56.
78. Schuster, T.M. and T.M. Laue, *Modern analytical ultracentrifugation : acquisition and interpretation of data for biological and synthetic polymer systems*. Emerging biochemical and biophysical techniques. 1994, Boston: Birkhäuser. xiv, 351 p.
79. Zannini, L., D. Delia, and G. Buscemi, *CHK2 kinase in the DNA damage response and beyond*. *J Mol Cell Biol*, 2014. **6**(6): p. 442-57.
80. Antoni, L., N. Sodha, I. Collins, and M.D. Garrett, *CHK2 kinase: cancer susceptibility and cancer therapy - two sides of the same coin?* *Nat Rev Cancer*, 2007. **7**(12): p. 925-36.
81. Schwarz, J.K., C.M. Lovly, and H. Piwnica-Worms, *Regulation of the Chk2 protein kinase by oligomerization-mediated cis- and trans-phosphorylation*. *Mol Cancer Res*, 2003. **1**(8): p. 598-609.
82. Cai, Z., N.H. Chehab, and N.P. Pavletich, *Structure and activation mechanism of the CHK2 DNA damage checkpoint kinase*. *Mol Cell*, 2009. **35**(6): p. 818-29.
83. Bell, D.W., J.M. Varley, T.E. Szydlo, D.H. Kang, D.C. Wahrer, K.E. Shannon, M. Lubratovich, S.J. Verselis, K.J. Isselbacher, J.F. Fraumeni, J.M. Birch, F.P. Li, J.E. Garber,

- and D.A. Haber, *Heterozygous germ line hCHK2 mutations in Li-Fraumeni syndrome*. Science, 1999. **286**(5449): p. 2528-31.
84. Giudicelli, V., X. Brochet, and M.P. Lefranc, *IMGT/V-QUEST: IMGT standardized analysis of the immunoglobulin (IG) and T cell receptor (TR) nucleotide sequences*. Cold Spring Harb Protoc, 2011. **2011**(6): p. 695-715.
85. Snyder, A. and T.A. Chan, *Immunogenic peptide discovery in cancer genomes*. Curr Opin Genet Dev, 2015. **30**: p. 7-16.
86. Nielsen, M., C. Lundegaard, T. Blicher, K. Lamberth, M. Harndahl, S. Justesen, G. Roder, B. Peters, A. Sette, O. Lund, and S. Buus, *NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence*. PLoS One, 2007. **2**(8): p. e796.
87. Favero, F., T. Joshi, A.M. Marquard, N.J. Birkbak, M. Krzystanek, Q. Li, Z. Szallasi, and A.C. Eklund, *Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data*. Ann Oncol, 2015. **26**(1): p. 64-70.
88. Parikh, S.A., K.G. Rabe, N.E. Kay, T.G. Call, W. Ding, S.M. Schwager, D.A. Bowen, M. Conte, D.F. Jelinek, S.L. Slager, and T.D. Shanafelt, *Chronic lymphocytic leukemia in young (<= 55 years) patients: a comprehensive analysis of prognostic factors and outcomes*. Haematologica, 2014. **99**(1): p. 140-7.
89. Pflug, N., J. Bahlo, T.D. Shanafelt, B.F. Eichhorst, M.A. Bergmann, T. Elter, K. Bauer, G. Malchau, K.G. Rabe, S. Stilgenbauer, H. Dohner, U. Jager, M.J. Eckart, G. Hopfinger, R. Busch, A.M. Fink, C.M. Wendtner, K. Fischer, N.E. Kay, and M. Hallek, *Development of a comprehensive prognostic index for patients with chronic lymphocytic leukemia*. Blood, 2014. **124**(1): p. 49-62.
90. Nadeu, F., J. Delgado, C. Royo, T. Baumann, T. Stankovic, M. Pinyol, P. Jares, A. Navarro, D. Martin-Garcia, S. Bea, I. Salaverria, C. Oldreive, M. Aymerich, H. Suarez-Cisneros, M. Rozman, N. Villamor, D. Colomer, A. Lopez-Guillermo, M. Gonzalez, M.

- Alcoceba, M.J. Terol, E. Colado, X.S. Puente, C. Lopez-Otin, A. Enjuanes, and E. Campo, *Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia*. *Blood*, 2016. **127**(17): p. 2122-30.
91. Young, E., D. Noerenberg, L. Mansouri, V. Ljungstrom, M. Frick, L.A. Sutton, S.J. Blakemore, J. Galan-Sousa, K. Plevova, P. Baliakas, D. Rossi, R. Clifford, D. Roos-Weil, V. Navrkalova, B. Dorken, C.A. Schmitt, K.E. Smedby, G. Juliusson, B. Giacomelli, J.S. Blachly, C. Belessi, P. Panagiotidis, N. Chiorazzi, F. Davi, A.W. Langerak, D. Oscier, A. Schuh, G. Gaidano, P. Ghia, W. Xu, L. Fan, O.A. Bernard, F. Nguyen-Khac, L. Rassenti, J. Li, T.J. Kipps, K. Stamatopoulos, S. Pospisilova, T. Zenz, C.C. Oakes, J.C. Strefford, R. Rosenquist, and F. Damm, *EGR2 mutations define a new clinically aggressive subgroup of chronic lymphocytic leukemia*. *Leukemia*, 2017.
92. Guieze, R., P. Robbe, R. Clifford, S. de Guibert, B. Pereira, A. Timbs, M.S. Dilhuydy, M. Cabes, L. Ysebaert, A. Burns, F. Nguyen-Khac, F. Davi, L. Veronese, P. Combes, M. Le Garff-Tavernier, V. Leblond, H. Merle-Beral, R. Alsolami, A. Hamblin, J. Mason, A. Pettitt, P. Hillmen, J. Taylor, S.J. Knight, O. Tournilhac, and A. Schuh, *Presence of multiple recurrent mutations confers poor trial outcome of relapsed/refractory CLL*. *Blood*, 2015. **126**(18): p. 2110-7.
93. van Heijst, J.W., I. Ceberio, L.B. Lipuma, D.W. Samilo, G.D. Wasilewski, A.M. Gonzales, J.L. Nieves, M.R. van den Brink, M.A. Perales, and E.G. Pamer, *Quantitative assessment of T cell repertoire recovery after hematopoietic stem cell transplantation*. *Nat Med*, 2013. **19**(3): p. 372-7.
94. Buhler, S., F. Bettens, C. Dantin, S. Ferrari-Lacraz, M. Ansari, A.C. Mamez, S. Masouridi-Levrat, Y. Chalandon, and J. Villard, *Genetic T-cell receptor diversity at 1 year following allogeneic hematopoietic stem cell transplantation*. *Leukemia*, 2020. **34**(5): p. 1422-1432.
95. Suessmuth, Y., R. Mukherjee, B. Watkins, D.T. Koura, K. Finstermeier, C. Desmarais, L. Stempora, J.T. Horan, A. Langston, M. Qayed, H.J. Khoury, A. Grizzle, J.A. Cheeseman,

- J.A. Conger, J. Robertson, A. Garrett, A.D. Kirk, E.K. Waller, B.R. Blazar, A.K. Mehta, H.S. Robins, and L.S. Kean, *CMV reactivation drives posttransplant T-cell reconstitution and results in defects in the underlying TCRbeta repertoire*. *Blood*, 2015. **125**(25): p. 3835-50.
96. Nazarov, V.I., M.V. Pogorelyy, E.A. Komech, I.V. Zvyagin, D.A. Bolotin, M. Shugay, D.M. Chudakov, Y.B. Lebedev, and I.Z. Mamedov, *tcR: an R package for T cell receptor repertoire advanced data analysis*. *BMC Bioinformatics*, 2015. **16**: p. 175.
97. Venturi, V., K. Kedzierska, M.M. Tanaka, S.J. Turner, P.C. Doherty, and M.P. Davenport, *Method for assessing the similarity between subsets of the T cell receptor repertoire*. *J Immunol Methods*, 2008. **329**(1-2): p. 67-80.
98. Kirsch, I., M. Vignali, and H. Robins, *T-cell receptor profiling in cancer*. *Mol Oncol*, 2015. **9**(10): p. 2063-70.
99. Venturi, V., K. Kedzierska, S.J. Turner, P.C. Doherty, and M.P. Davenport, *Methods for comparing the diversity of samples of the T cell receptor repertoire*. *J Immunol Methods*, 2007. **321**(1-2): p. 182-95.
100. Emerson, R.O., W.S. DeWitt, M. Vignali, J. Gravley, J.K. Hu, E.J. Osborne, C. Desmarais, M. Klinger, C.S. Carlson, J.A. Hansen, M. Rieder, and H.S. Robins, *Immunosequencing identifies signatures of cytomegalovirus exposure history and HLA-mediated effects on the T cell repertoire*. *Nat Genet*, 2017. **49**(5): p. 659-665.
101. Lim, A., L. Trautmann, M.A. Peyrat, C. Couedel, F. Davodeau, F. Romagne, P. Kourilsky, and M. Bonneville, *Frequent contribution of T cell clonotypes with public TCR features to the chronic response against a dominant EBV-derived epitope: application to direct detection of their molecular imprint on the human peripheral T cell repertoire*. *J Immunol*, 2000. **165**(4): p. 2001-11.
102. Zhang, C., J. Lou, N. Li, I. Todorov, C.L. Lin, Y.A. Cao, C.H. Contag, F. Kandeel, S. Forman, and D. Zeng, *Donor CD8+ T cells mediate graft-versus-leukemia activity without*

- clinical signs of graft-versus-host disease in recipients conditioned with anti-CD3 monoclonal antibody.* J Immunol, 2007. **178**(2): p. 838-50.
103. Fry, T.J. and C.L. Mackall, *Immune reconstitution following hematopoietic progenitor cell transplantation: challenges for the future.* Bone Marrow Transplant, 2005. **35 Suppl 1**: p. S53-7.
104. Han, A., J. Glanville, L. Hansmann, and M.M. Davis, *Linking T-cell receptor sequence to functional phenotype at the single-cell level.* Nat Biotechnol, 2014. **32**(7): p. 684-92.
105. Glanville, J., W. Zhai, J. Berka, D. Telman, G. Huerta, G.R. Mehta, I. Ni, L. Mei, P.D. Sundar, G.M. Day, D. Cox, A. Rajpal, and J. Pons, *Precise determination of the diversity of a combinatorial antibody library gives insight into the human immunoglobulin repertoire.* Proc Natl Acad Sci U S A, 2009. **106**(48): p. 20216-21.
106. Butler, A., P. Hoffman, P. Smibert, E. Papalexi, and R. Satija, *Integrating single-cell transcriptomic data across different conditions, technologies, and species.* Nat Biotechnol, 2018. **36**(5): p. 411-420.
107. Stuart, T., A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W.M. Mauck, 3rd, Y. Hao, M. Stoeckius, P. Smibert, and R. Satija, *Comprehensive Integration of Single-Cell Data.* Cell, 2019. **177**(7): p. 1888-1902 e21.
108. DeWitt, W.S., R.O. Emerson, P. Lindau, M. Vignali, T.M. Snyder, C. Desmarais, C. Sanders, H. Utsugi, E.H. Warren, J. McElrath, K.W. Makar, A. Wald, and H.S. Robins, *Dynamics of the cytotoxic T cell response to a model of acute viral infection.* J Virol, 2015. **89**(8): p. 4517-26.
109. Venturi, V., K. Kedzierska, D.A. Price, P.C. Doherty, D.C. Douek, S.J. Turner, and M.P. Davenport, *Sharing of T cell receptors in antigen-specific responses is driven by convergent recombination.* Proc Natl Acad Sci U S A, 2006. **103**(49): p. 18691-6.
110. Alho, A.C., H.T. Kim, M.J. Chammas, C.G. Reynolds, T.R. Matos, E. Forcade, J. Whangbo, S. Nikiforow, C.S. Cutler, J. Koreth, V.T. Ho, P. Armand, J.H. Antin, E.P. Alyea,

- J.F. Lacerda, R.J. Soiffer, and J. Ritz, *Unbalanced recovery of regulatory and effector T cells after allogeneic stem cell transplantation contributes to chronic GVHD*. *Blood*, 2016. **127**(5): p. 646-57.
111. Storb, R. and B.M. Sandmaier, *Nonmyeloablative allogeneic hematopoietic cell transplantation*. *Haematologica*, 2016. **101**(5): p. 521-30.
112. Nowell, P.C., *The clonal evolution of tumor cell populations*. *Science*, 1976. **194**(4260): p. 23-8.
113. Mossner, M., J.C. Jann, J. Wittig, F. Nolte, S. Fey, V. Nowak, J. Oblander, J. Pressler, I. Palme, C. Xanthopoulos, T. Boch, G. Metzgeroth, H. Rohl, S.H. Witt, H. Dukal, C. Klein, S. Schmitt, P. Gelss, U. Platzbecker, E. Balaian, A. Fabarius, H. Blum, T.J. Schulze, M. Meggendorfer, C. Haferlach, A. Trumpp, W.K. Hofmann, H. Medyouf, and D. Nowak, *Mutational hierarchies in myelodysplastic syndromes dynamically adapt and evolve upon therapy response and failure*. *Blood*, 2016. **128**(9): p. 1246-59.
114. O'Leary, B., R.J. Cutts, Y. Liu, S. Hrebien, X. Huang, K. Fenwick, F. Andre, S. Loibl, S. Loi, I. Garcia-Murillas, M. Cristofanilli, C. Huang Bartlett, and N.C. Turner, *The Genetic Landscape and Clonal Evolution of Breast Cancer Resistance to Palbociclib plus Fulvestrant in the PALOMA-3 Trial*. *Cancer Discov*, 2018. **8**(11): p. 1390-1403.
115. Hallek, M., K. Fischer, G. Fingerle-Rowson, A.M. Fink, R. Busch, J. Mayer, M. Hensel, G. Hopfinger, G. Hess, U. von Grunhagen, M. Bergmann, J. Catalano, P.L. Zinzani, F. Caligaris-Cappio, J.F. Seymour, A. Berrebi, U. Jager, B. Cazin, M. Trneny, A. Westermann, C.M. Wendtner, B.F. Eichhorst, P. Staib, A. Buhler, D. Winkler, T. Zenz, S. Bottcher, M. Ritgen, M. Mendila, M. Kneba, H. Dohner, S. Stilgenbauer, I. International Group of, and G. German Chronic Lymphocytic Leukaemia Study, *Addition of rituximab to fludarabine and cyclophosphamide in patients with chronic lymphocytic leukaemia: a randomised, open-label, phase 3 trial*. *Lancet*, 2010. **376**(9747): p. 1164-74.

116. Landau, D.A., C. Sun, D. Rosebrock, S.E.M. Herman, J. Fein, M. Sivina, C. Underbayev, D. Liu, J. Hoellenriegel, S. Ravichandran, M.Z.H. Farooqui, W. Zhang, C. Cibulskis, A. Zviran, D.S. Neuberg, D. Livitz, I. Bozic, I. Leshchiner, G. Getz, J.A. Burger, A. Wiestner, and C.J. Wu, *The evolutionary landscape of chronic lymphocytic leukemia treated with ibrutinib targeted therapy*. Nat Commun, 2017. **8**(1): p. 2185.
117. Quek, L., P. Ferguson, M. Metzner, I. Ahmed, A. Kennedy, C. Garnett, S. Jeffries, C. Walter, K. Piechocki, A. Timbs, R. Danby, M. Raghavan, A. Peniket, M. Griffiths, A. Bacon, J. Ward, K. Wheatley, P. Vyas, and C. Craddock, *Mutational analysis of disease relapse in patients allografted for acute myeloid leukemia*. Blood Adv, 2016. **1**(3): p. 193-204.
118. Christopher, M.J., A.A. Petti, M.P. Rettig, C.A. Miller, E. Chendamara, E.J. Duncavage, J.M. Klco, N.M. Helton, M. O'Laughlin, C.C. Fronick, R.S. Fulton, R.K. Wilson, L.D. Wartman, J.S. Welch, S.E. Heath, J.D. Baty, J.E. Payton, T.A. Graubert, D.C. Link, M.J. Walter, P. Westervelt, T.J. Ley, and J.F. DiPersio, *Immune Escape of Relapsed AML Cells after Allogeneic Transplantation*. N Engl J Med, 2018. **379**(24): p. 2330-2341.
119. Toffalori, C., L. Zito, V. Gambacorta, M. Riba, G. Oliveira, G. Bucci, M. Barcella, O. Spinelli, R. Greco, L. Crucitti, N. Cieri, M. Noviello, F. Manfredi, E. Montaldo, R. Ostuni, M.M. Naldini, B. Gentner, M. Waterhouse, R. Zeiser, J. Finke, M. Hanoun, D.W. Beelen, I. Gojo, L. Luznik, M. Onozawa, T. Teshima, R. Devillier, D. Blaise, C.J.M. Halkes, M. Griffioen, M.G. Carrabba, M. Bernardi, J. Peccatori, C. Barlassina, E. Stupka, D. Lazarevic, G. Tonon, A. Rambaldi, D. Cittaro, C. Bonini, K. Fleischhauer, F. Ciceri, and L. Vago, *Immune signature drives leukemia escape and relapse after hematopoietic cell transplantation*. Nat Med, 2019. **25**(4): p. 603-611.
120. Riaz, N., J.J. Havel, V. Makarov, A. Desrichard, W.J. Urba, J.S. Sims, F.S. Hodi, S. Martin-Algarra, R. Mandal, W.H. Sharfman, S. Bhatia, W.J. Hwu, T.F. Gajewski, C.L. Slingluff, Jr., D. Chowell, S.M. Kendall, H. Chang, R. Shah, F. Kuo, L.G.T. Morris, J.W. Sidhom,

- J.P. Schneck, C.E. Horak, N. Weinhold, and T.A. Chan, *Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab*. *Cell*, 2017. **171**(4): p. 934-949 e16.
121. Sharma, P., S. Hu-Lieskovan, J.A. Wargo, and A. Ribas, *Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy*. *Cell*, 2017. **168**(4): p. 707-723.
122. Pardoll, D.M., *The blockade of immune checkpoints in cancer immunotherapy*. *Nat Rev Cancer*, 2012. **12**(4): p. 252-64.
123. Spranger, S., R.M. Spaapen, Y. Zha, J. Williams, Y. Meng, T.T. Ha, and T.F. Gajewski, *Up-regulation of PD-L1, IDO, and T(regs) in the melanoma tumor microenvironment is driven by CD8(+) T cells*. *Sci Transl Med*, 2013. **5**(200): p. 200ra116.
124. Koyama, S., E.A. Akbay, Y.Y. Li, G.S. Herter-Sprie, K.A. Buczkowski, W.G. Richards, L. Gandhi, A.J. Redig, S.J. Rodig, H. Asahina, R.E. Jones, M.M. Kulkarni, M. Kuraguchi, S. Palakurthi, P.E. Fecci, B.E. Johnson, P.A. Janne, J.A. Engelman, S.P. Gangadharan, D.B. Costa, G.J. Freeman, R. Bueno, F.S. Hodi, G. Dranoff, K.K. Wong, and P.S. Hammerman, *Adaptive resistance to therapeutic PD-1 blockade is associated with upregulation of alternative immune checkpoints*. *Nat Commun*, 2016. **7**: p. 10501.
125. Green, M.R., S. Monti, S.J. Rodig, P. Juszczynski, T. Currie, E. O'Donnell, B. Chapuy, K. Takeyama, D. Neuberg, T.R. Golub, J.L. Kutok, and M.A. Shipp, *Integrative analysis reveals selective 9p24.1 amplification, increased PD-1 ligand expression, and further induction via JAK2 in nodular sclerosing Hodgkin lymphoma and primary mediastinal large B-cell lymphoma*. *Blood*, 2010. **116**(17): p. 3268-77.
126. Zaretsky, J.M., A. Garcia-Diaz, D.S. Shin, H. Escuin-Ordinas, W. Hugo, S. Hu-Lieskovan, D.Y. Torrejon, G. Abril-Rodriguez, S. Sandoval, L. Barthly, J. Saco, B. Homet Moreno, R. Mezzadra, B. Chmielowski, K. Ruchalski, I.P. Shintaku, P.J. Sanchez, C. Puig-Saus, G. Cherry, E. Seja, X. Kong, J. Pang, B. Berent-Maoz, B. Comin-Anduix, T.G. Graeber, P.C. Tumeh, T.N. Schumacher, R.S. Lo, and A. Ribas, *Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma*. *N Engl J Med*, 2016. **375**(9): p. 819-29.

127. Gao, J., L.Z. Shi, H. Zhao, J. Chen, L. Xiong, Q. He, T. Chen, J. Roszik, C. Bernatchez, S.E. Woodman, P.L. Chen, P. Hwu, J.P. Allison, A. Futreal, J.A. Wargo, and P. Sharma, *Loss of IFN-gamma Pathway Genes in Tumor Cells as a Mechanism of Resistance to Anti-CTLA-4 Therapy*. Cell, 2016. **167**(2): p. 397-404 e9.
128. Restifo, N.P., F.M. Marincola, Y. Kawakami, J. Taubenberger, J.R. Yannelli, and S.A. Rosenberg, *Loss of functional beta 2-microglobulin in metastatic melanomas from five patients receiving immunotherapy*. J Natl Cancer Inst, 1996. **88**(2): p. 100-8.
129. Tran, E., P.F. Robbins, Y.C. Lu, T.D. Prickett, J.J. Gartner, L. Jia, A. Pasetto, Z. Zheng, S. Ray, E.M. Groh, I.R. Kriley, and S.A. Rosenberg, *T-Cell Transfer Therapy Targeting Mutant KRAS in Cancer*. N Engl J Med, 2016. **375**(23): p. 2255-2262.
130. Spranger, S., R. Bao, and T.F. Gajewski, *Melanoma-intrinsic beta-catenin signalling prevents anti-tumour immunity*. Nature, 2015. **523**(7559): p. 231-5.
131. Liu, C., W. Peng, C. Xu, Y. Lou, M. Zhang, J.A. Wargo, J.Q. Chen, H.S. Li, S.S. Watowich, Y. Yang, D. Tompers Frederick, Z.A. Cooper, R.M. Mbofung, M. Whittington, K.T. Flaherty, S.E. Woodman, M.A. Davies, L.G. Radvanyi, W.W. Overwijk, G. Lizee, and P. Hwu, *BRAF inhibition increases tumor infiltration by T cells and enhances the antitumor activity of adoptive immunotherapy in mice*. Clin Cancer Res, 2013. **19**(2): p. 393-403.
132. Peng, W., J.Q. Chen, C. Liu, S. Malu, C. Creasy, M.T. Tetzlaff, C. Xu, J.A. McKenzie, C. Zhang, X. Liang, L.J. Williams, W. Deng, G. Chen, R. Mbofung, A.J. Lazar, C.A. Torres-Cabala, Z.A. Cooper, P.L. Chen, T.N. Tieu, S. Spranger, X. Yu, C. Bernatchez, M.A. Forget, C. Haymaker, R. Amaria, J.L. McQuade, I.C. Glitza, T. Cascone, H.S. Li, L.N. Kwong, T.P. Heffernan, J. Hu, R.L. Bassett, Jr., M.W. Bosenberg, S.E. Woodman, W.W. Overwijk, G. Lizee, J. Roszik, T.F. Gajewski, J.A. Wargo, J.E. Gershenwald, L. Radvanyi, M.A. Davies, and P. Hwu, *Loss of PTEN Promotes Resistance to T Cell-Mediated Immunotherapy*. Cancer Discov, 2016. **6**(2): p. 202-16.

133. Rooney, M.S., S.A. Shukla, C.J. Wu, G. Getz, and N. Hacohen, *Molecular and genetic properties of tumors associated with local immune cytolytic activity*. Cell, 2015. **160**(1-2): p. 48-61.
134. Matsushita, H., M.D. Vesely, D.C. Koboldt, C.G. Rickert, R. Uppaluri, V.J. Magrini, C.D. Arthur, J.M. White, Y.S. Chen, L.K. Shea, J. Hundal, M.C. Wendl, R. Demeter, T. Wylie, J.P. Allison, M.J. Smyth, L.J. Old, E.R. Mardis, and R.D. Schreiber, *Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting*. Nature, 2012. **482**(7385): p. 400-4.
135. Joyce, J.A. and D.T. Fearon, *T cell exclusion, immune privilege, and the tumor microenvironment*. Science, 2015. **348**(6230): p. 74-80.
136. Salerno, E.P., D. Bedognetti, I.S. Mauldin, D.H. Deacon, S.M. Shea, J. Pinczewski, J.M. Obeid, G. Coukos, E. Wang, T.F. Gajewski, F.M. Marincola, and C.L. Slingluff, Jr., *Human melanomas and ovarian cancers overexpressing mechanical barrier molecule genes lack immune signatures and have increased patient mortality risk*. Oncoimmunology, 2016. **5**(12): p. e1240857.
137. Gajewski, T.F., H. Schreiber, and Y.X. Fu, *Innate and adaptive immune cells in the tumor microenvironment*. Nat Immunol, 2013. **14**(10): p. 1014-22.
138. Gaiti, F., R. Chaligne, H. Gu, R.M. Brand, S. Kothen-Hill, R.C. Schulman, K. Grigorev, D. Risso, K.T. Kim, A. Pastore, K.Y. Huang, A. Alonso, C. Sheridan, N.D. Omans, E. Biederstedt, K. Clement, L. Wang, J.A. Felsenfeld, E.B. Bhavsar, M.J. Aryee, J.N. Allan, R. Furman, A. Gnirke, C.J. Wu, A. Meissner, and D.A. Landau, *Epigenetic evolution and lineage histories of chronic lymphocytic leukaemia*. Nature, 2019. **569**(7757): p. 576-580.
139. Walther, V., C.T. Hiley, D. Shibata, C. Swanton, P.E. Turner, and C.C. Maley, *Cancer oncology recapitulate paleontology? Lessons from species extinctions*. Nat Rev Clin Oncol, 2015. **12**(5): p. 273-85.

140. Alexandrov, L.B., S. Nik-Zainal, D.C. Wedge, S.A. Aparicio, S. Behjati, A.V. Biankin, G.R. Bignell, N. Bolli, A. Borg, A.L. Borresen-Dale, S. Boyault, B. Burkhardt, A.P. Butler, C. Caldas, H.R. Davies, C. Desmedt, R. Eils, J.E. Eyfjord, J.A. Foekens, M. Greaves, F. Hosoda, B. Hutter, T. Ilcic, S. Imbeaud, M. Imielinski, N. Jager, D.T. Jones, D. Jones, S. Knappskog, M. Kool, S.R. Lakhani, C. Lopez-Otin, S. Martin, N.C. Munshi, H. Nakamura, P.A. Northcott, M. Pajic, E. Papaemmanuil, A. Paradiso, J.V. Pearson, X.S. Puente, K. Raine, M. Ramakrishna, A.L. Richardson, J. Richter, P. Rosenstiel, M. Schlesner, T.N. Schumacher, P.N. Span, J.W. Teague, Y. Totoki, A.N. Tutt, R. Valdes-Mas, M.M. van Buuren, L. van 't Veer, A. Vincent-Salomon, N. Waddell, L.R. Yates, I. Australian Pancreatic Cancer Genome, I.B.C. Consortium, I.M.-S. Consortium, I. PedBrain, J. Zucman-Rossi, P.A. Futreal, U. McDermott, P. Lichter, M. Meyerson, S.M. Grimmond, R. Siebert, E. Campo, T. Shibata, S.M. Pfister, P.J. Campbell, and M.R. Stratton, *Signatures of mutational processes in human cancer*. Nature, 2013. **500**(7463): p. 415-21.
141. Jameson-Lee, M., V. Koparde, P. Griffith, A.F. Scalora, J.K. Sampson, H. Khalid, N.U. Sheth, M. Batalo, M.G. Serrano, C.H. Roberts, M.L. Hess, G.A. Buck, M.C. Neale, M.H. Manjili, and A.A. Toor, *In silico Derivation of HLA-Specific Alloreactivity Potential from Whole Exome Sequencing of Stem-Cell Transplant Donors and Recipients: Understanding the Quantitative Immunobiology of Allogeneic Transplantation*. Front Immunol, 2014. **5**: p. 529.
142. Verdegaal, E.M., N.F. de Miranda, M. Visser, T. Harryvan, M.M. van Buuren, R.S. Andersen, S.R. Hadrup, C.E. van der Minne, R. Schotte, H. Spits, J.B. Haanen, E.H. Kapiteijn, T.N. Schumacher, and S.H. van der Burg, *Neoantigen landscape dynamics during human melanoma-T cell interactions*. Nature, 2016. **536**(7614): p. 91-5.
143. Litzow, M.R., S. Tarima, W.S. Perez, B.J. Bolwell, M.S. Cairo, B.M. Camitta, C.S. Cutler, M. de Lima, J.F. Dipersio, R.P. Gale, A. Keating, H.M. Lazarus, S. Luger, D.I. Marks, R.T. Maziarz, P.L. McCarthy, M.C. Pasquini, G.L. Phillips, J.D. Rizzo, J. Sierra, M.S. Tallman,

- and D.J. Weisdorf, *Allogeneic transplantation for therapy-related myelodysplastic syndrome and acute myeloid leukemia*. Blood, 2010. **115**(9): p. 1850-7.
144. Slovak, M.L., K.J. Kopecky, P.A. Cassileth, D.H. Harrington, K.S. Theil, A. Mohamed, E. Paietta, C.L. Willman, D.R. Head, J.M. Rowe, S.J. Forman, and F.R. Appelbaum, *Karyotypic analysis predicts outcome of preremission and postremission therapy in adult acute myeloid leukemia: a Southwest Oncology Group/Eastern Cooperative Oncology Group Study*. Blood, 2000. **96**(13): p. 4075-83.
145. Kharfan-Dabaja, M.A., A. Kumar, M. Hamadani, S. Stilgenbauer, P. Ghia, C. Anasetti, P. Dreger, E. Montserrat, M.A. Perales, E.P. Alyea, F.T. Awan, E. Ayala, J.C. Barrientos, J.R. Brown, J.E. Castro, R.R. Furman, J. Gribben, B.T. Hill, M. Mohty, C. Moreno, S. O'Brien, S.Z. Pavletic, J. Pinilla-Ibarz, N.M. Reddy, M. Sorrow, C. Bredeson, P. Carpenter, and B.N. Savani, *Clinical Practice Recommendations for Use of Allogeneic Hematopoietic Cell Transplantation in Chronic Lymphocytic Leukemia on Behalf of the Guidelines Committee of the American Society for Blood and Marrow Transplantation*. Biol Blood Marrow Transplant, 2016. **22**(12): p. 2117-2125.
146. Donehower, L.A., T. Soussi, A. Korkut, Y. Liu, A. Schultz, M. Cardenas, X. Li, O. Babur, T.K. Hsu, O. Lichtarge, J.N. Weinstein, R. Akbani, and D.A. Wheeler, *Integrated Analysis of TP53 Gene and Pathway Alterations in The Cancer Genome Atlas*. Cell Rep, 2019. **28**(11): p. 3010.
147. Van Den Neste, E., V. Robin, J. Francart, A. Hagemeijer, M. Stul, P. Vandenberghe, A. Delannoy, A. Sonet, V. Deneys, S. Costantini, A. Ferrant, A. Robert, and L. Michaux, *Chromosomal translocations independently predict treatment failure, treatment-free survival and overall survival in B-cell chronic lymphocytic leukemia patients treated with cladribine*. Leukemia, 2007. **21**(8): p. 1715-22.

148. Roth, A., J. Khattra, D. Yap, A. Wan, E. Laks, J. Biele, G. Ha, S. Aparicio, A. Bouchard-Cote, and S.P. Shah, *PyClone: statistical inference of clonal population structure in cancer*. Nat Methods, 2014. **11**(4): p. 396-8.
149. Liu, C., M. He, B. Rooney, T.B. Kepler, and N.J. Chao, *Longitudinal analysis of T-cell receptor variable beta chain repertoire in patients with acute graft-versus-host disease after allogeneic stem cell transplantation*. Biol Blood Marrow Transplant, 2006. **12**(3): p. 335-45.
150. Du, J.W., J.Y. Gu, J. Liu, X.N. Cen, Y. Zhang, Y. Ou, B. Chu, and P. Zhu, *TCR spectratyping revealed T lymphocytes associated with graft-versus-host disease after allogeneic hematopoietic stem cell transplantation*. Leuk Lymphoma, 2007. **48**(8): p. 1618-27.
151. Chaudhry, M.S., E. Velardi, F. Malard, and M.R. van den Brink, *Immune Reconstitution after Allogeneic Hematopoietic Stem Cell Transplantation: Time To T Up the Thymus*. J Immunol, 2017. **198**(1): p. 40-46.
152. Clave, E., M. Busson, C. Douay, R. Peffault de Latour, J. Berrou, C. Rabian, M. Carmagnat, V. Rocha, D. Charron, G. Socie, and A. Toubert, *Acute graft-versus-host disease transiently impairs thymic output in young patients after allogeneic hematopoietic stem cell transplantation*. Blood, 2009. **113**(25): p. 6477-84.
153. Small, T.N., E.B. Papadopoulos, F. Boulad, P. Black, H. Castro-Malaspina, B.H. Childs, N. Collins, A. Gillio, D. George, A. Jakubowski, G. Heller, M. Fazzari, N. Kernan, S. MacKinnon, P. Szabolcs, J.W. Young, and R.J. O'Reilly, *Comparison of immune reconstitution after unrelated and related T-cell-depleted bone marrow transplantation: effect of patient age and donor leukocyte infusions*. Blood, 1999. **93**(2): p. 467-80.
154. Verfuether, S., K. Peggs, P. Vyas, L. Barnett, R.J. O'Reilly, and S. Mackinnon, *Longitudinal monitoring of immune reconstitution by CDR3 size spectratyping after T-cell-depleted*

- allogeneic bone marrow transplant and the effect of donor lymphocyte infusions on T-cell repertoire.* Blood, 2000. **95**(12): p. 3990-5.
155. Wu, C.J., A. Chillemi, E.P. Alyea, E. Orsini, D. Neuberg, R.J. Soiffer, and J. Ritz, *Reconstitution of T-cell receptor repertoire diversity following T-cell depleted allogeneic bone marrow transplantation is related to hematopoietic chimerism.* Blood, 2000. **95**(1): p. 352-9.
156. Cochet, M., C. Pannetier, A. Regnault, S. Darche, C. Leclerc, and P. Kourilsky, *Molecular detection and in vivo analysis of the specific T cell response to a protein antigen.* Eur J Immunol, 1992. **22**(10): p. 2639-47.
157. Gorski, J., M. Yassai, X. Zhu, B. Kissela, B. Kissella, C. Keever, and N. Flomenberg, *Circulating T cell repertoire complexity in normal individuals and bone marrow recipients analyzed by CDR3 size spectratyping. Correlation with immune status.* J Immunol, 1994. **152**(10): p. 5109-19.
158. Langerak, A.W., R. van Den Beemd, I.L. Wolvers-Tettero, P.P. Boor, E.G. van Lochem, H. Hooijkaas, and J.J. van Dongen, *Molecular and flow cytometric analysis of the Vbeta repertoire for clonality assessment in mature TCRalpha/beta T-cell proliferations.* Blood, 2001. **98**(1): p. 165-73.
159. Yew, P.Y., H. Alachkar, R. Yamaguchi, K. Kiyotani, H. Fang, K.L. Yap, H.T. Liu, A. Wickrema, A. Artz, K. van Besien, S. Imoto, S. Miyano, M.R. Bishop, W. Stock, and Y. Nakamura, *Quantitative characterization of T-cell repertoire in allogeneic hematopoietic stem cell transplant recipients.* Bone Marrow Transplant, 2015. **50**(9): p. 1227-34.
160. Kanakry, C.G., D.G. Coffey, A.M. Towler, A. Vulic, B.E. Storer, J. Chou, C.C. Yeung, C.D. Gocke, H.S. Robins, P.V. O'Donnell, L. Luznik, and E.H. Warren, *Origin and evolution of the T cell repertoire after posttransplantation cyclophosphamide.* JCI Insight, 2016. **1**(5).
161. Leick, M., R.M. Gittelman, E. Yusko, C. Sanders, H. Robins, Z. DeFilipp, S. Nikiforow, J. Ritz, and Y.B. Chen, *T Cell Clonal Dynamics Determined by High-Resolution TCR-beta*

- Sequencing in Recipients after Allogeneic Hematopoietic Cell Transplantation*. Biol Blood Marrow Transplant, 2020.
162. Meier, J.A., M. Haque, M. Fawaz, H. Abdeen, D. Coffey, A. Towler, A. Abdeen, A. Toor, E. Warren, J. Reed, C.G. Kanakry, A. Keating, L. Luznik, and A.A. Toor, *T Cell Repertoire Evolution after Allogeneic Bone Marrow Transplantation: An Organizational Perspective*. Biol Blood Marrow Transplant, 2019. **25**(5): p. 868-882.
163. Kanda, J., L.W. Chiou, P. Szabolcs, G.D. Sempowski, D.A. Rizzieri, G.D. Long, K.M. Sullivan, C. Gasparetto, J.P. Chute, A. Morris, J. McPherson, J. Hale, J.A. Livingston, G. Broadwater, D. Niedzwiecki, N.J. Chao, and M.E. Horwitz, *Immune recovery in adult patients after myeloablative dual umbilical cord blood, matched sibling, and matched unrelated donor hematopoietic cell transplantation*. Biol Blood Marrow Transplant, 2012. **18**(11): p. 1664-1676 e1.
164. Link-Rachner, C.S., A. Eugster, E. Rucker-Braun, F. Heidenreich, U. Oelschlagel, A. Dahl, C. Klesse, M. Kuhn, J.M. Middeke, M. Bornhauser, E. Bonifacio, and J. Schetelig, *T-cell receptor-alpha repertoire of CD8+ T cells following allogeneic stem cell transplantation using next-generation sequencing*. Haematologica, 2019. **104**(3): p. 622-631.
165. Koyama, D., M. Murata, R. Hanajiri, T. Akashi, S. Okuno, S. Kamoshita, J. Julamanee, E. Takagi, K. Miyao, R. Sakemura, T. Goto, S. Terakura, T. Nishida, and H. Kiyoi, *Quantitative Assessment of T Cell Clonotypes in Human Acute Graft-versus-Host Disease Tissues*. Biol Blood Marrow Transplant, 2019. **25**(3): p. 417-423.
166. Robins, H.S., S.K. Srivastava, P.V. Campregher, C.J. Turtle, J. Andriesen, S.R. Riddell, C.S. Carlson, and E.H. Warren, *Overlap and effective size of the human CD8+ T cell receptor repertoire*. Sci Transl Med, 2010. **2**(47): p. 47ra64.
167. Maury, S., J.Y. Mary, C. Rabian, M. Schwarzinger, A. Toubert, C. Scieux, M. Carmagnat, H. Esperou, P. Ribaud, A. Devergie, P. Guardiola, P. Vexiau, D. Charron, E. Gluckman,

- and G. Socie, *Prolonged immune deficiency following allogeneic stem cell transplantation: risk factors and complications in adult patients*. Br J Haematol, 2001. **115**(3): p. 630-41.
168. Qi, Q., Y. Liu, Y. Cheng, J. Glanville, D. Zhang, J.Y. Lee, R.A. Olshen, C.M. Weyand, S.D. Boyd, and J.J. Goronzy, *Diversity and clonal selection in the human T-cell repertoire*. Proc Natl Acad Sci U S A, 2014. **111**(36): p. 13139-44.
169. Gress, R.E., S.G. Emerson, and W.R. Drobyski, *Immune reconstitution: how it should work, what's broken, and why it matters*. Biol Blood Marrow Transplant, 2010. **16**(1 Suppl): p. S133-7.
170. Roberto, A., L. Castagna, V. Zanon, S. Bramanti, R. Crocchiolo, J.E. McLaren, S. Gandolfi, P. Tentorio, B. Sarina, I. Timofeeva, A. Santoro, C. Carlo-Stella, B. Bruno, C. Carniti, P. Corradini, E. Gostick, K. Ladell, D.A. Price, M. Roederer, D. Mavilio, and E. Lugli, *Role of naive-derived T memory stem cells in T-cell reconstitution following allogeneic transplantation*. Blood, 2015. **125**(18): p. 2855-64.
171. Anderson, B.E., J. McNiff, J. Yan, H. Doyle, M. Mamula, M.J. Shlomchik, and W.D. Shlomchik, *Memory CD4+ T cells do not induce graft-versus-host disease*. J Clin Invest, 2003. **112**(1): p. 101-8.
172. Chen, B.J., D. Deoliveira, X. Cui, N.T. Le, J. Son, J.F. Whitesides, and N.J. Chao, *Inability of memory T cells to induce graft-versus-host disease is a result of an abortive alloresponse*. Blood, 2007. **109**(7): p. 3115-23.
173. Zheng, H., C. Matte-Martone, H. Li, B.E. Anderson, S. Venketesan, H. Sheng Tan, D. Jain, J. McNiff, and W.D. Shlomchik, *Effector memory CD4+ T cells mediate graft-versus-leukemia without inducing graft-versus-host disease*. Blood, 2008. **111**(4): p. 2476-84.
174. Weinberg, K., B.R. Blazar, J.E. Wagner, E. Agura, B.J. Hill, M. Smogorzewska, R.A. Koup, M.R. Betts, R.H. Collins, and D.C. Douek, *Factors affecting thymic function after allogeneic hematopoietic stem cell transplantation*. Blood, 2001. **97**(5): p. 1458-66.

175. Chidrawar, S., N. Khan, W. Wei, A. McLarnon, N. Smith, L. Nayak, and P. Moss, *Cytomegalovirus-seropositivity has a profound influence on the magnitude of major lymphoid subsets within healthy individuals*. Clin Exp Immunol, 2009. **155**(3): p. 423-32.
176. Khan, N., N. Shariff, M. Cobbold, R. Bruton, J.A. Ainsworth, A.J. Sinclair, L. Nayak, and P.A. Moss, *Cytomegalovirus seropositivity drives the CD8 T cell repertoire toward greater clonality in healthy elderly individuals*. J Immunol, 2002. **169**(4): p. 1984-92.
177. Lugthart, G., M.M. van Ostaijen-Ten Dam, C.M. Jol-van der Zijde, T.C. van Holten, M.G. Kester, M.H. Heemskerk, R.G. Bredius, M.J. van Tol, and A.C. Lankester, *Early cytomegalovirus reactivation leaves a specific and dynamic imprint on the reconstituting T cell compartment long-term after hematopoietic stem cell transplantation*. Biol Blood Marrow Transplant, 2014. **20**(5): p. 655-61.
178. Sarantopoulos, S. and J. Ritz, *Aberrant B-cell homeostasis in chronic GVHD*. Blood, 2015. **125**(11): p. 1703-7.
179. Wu, T.D., S. Madireddi, P.E. de Almeida, R. Banchereau, Y.J. Chen, A.S. Chitre, E.Y. Chiang, H. Iftikhar, W.E. O'Gorman, A. Au-Yeung, C. Takahashi, L.D. Goldstein, C. Poon, S. Keerthivasan, D.E. de Almeida Nagata, X. Du, H.M. Lee, K.L. Banta, S. Mariathasan, M. Das Thakur, M.A. Huseni, M. Ballinger, I. Estay, P. Caplazi, Z. Modrusan, L. Delamarre, I. Mellman, R. Bourgon, and J.L. Grogan, *Peripheral T cell expansion predicts tumour infiltration and clinical response*. Nature, 2020. **579**(7798): p. 274-278.
180. Li, H., C. Ye, G. Ji, and J. Han, *Determinants of public T cell responses*. Cell Res, 2012. **22**(1): p. 33-42.
181. Venturi, V., D.A. Price, D.C. Douek, and M.P. Davenport, *The molecular basis for public T-cell responses?* Nat Rev Immunol, 2008. **8**(3): p. 231-8.
182. Zacharakis, N., H. Chinnasamy, M. Black, H. Xu, Y.C. Lu, Z. Zheng, A. Pasetto, M. Langan, T. Shelton, T. Prickett, J. Gartner, L. Jia, K. Trebska-McGowan, R.P. Somerville, P.F. Robbins, S.A. Rosenberg, S.L. Goff, and S.A. Feldman, *Immune recognition of*

- somatic mutations leading to complete durable regression in metastatic breast cancer. Nat Med, 2018. 24(6): p. 724-730.*
183. Tran, E., M. Ahmadzadeh, Y.C. Lu, A. Gros, S. Turcotte, P.F. Robbins, J.J. Gartner, Z. Zheng, Y.F. Li, S. Ray, J.R. Wunderlich, R.P. Somerville, and S.A. Rosenberg, *Immunogenicity of somatic mutations in human gastrointestinal cancers. Science, 2015. 350(6266): p. 1387-90.*
184. Rizvi, N.A., M.D. Hellmann, A. Snyder, P. Kvistborg, V. Makarov, J.J. Havel, W. Lee, J. Yuan, P. Wong, T.S. Ho, M.L. Miller, N. Rekhtman, A.L. Moreira, F. Ibrahim, C. Bruggeman, B. Gasmi, R. Zappasodi, Y. Maeda, C. Sander, E.B. Garon, T. Merghoub, J.D. Wolchok, T.N. Schumacher, and T.A. Chan, *Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science, 2015. 348(6230): p. 124-8.*
185. Snyder, A., V. Makarov, T. Merghoub, J. Yuan, J.M. Zaretsky, A. Desrichard, L.A. Walsh, M.A. Postow, P. Wong, T.S. Ho, T.J. Hollmann, C. Bruggeman, K. Kannan, Y. Li, C. Elipenahli, C. Liu, C.T. Harbison, L. Wang, A. Ribas, J.D. Wolchok, and T.A. Chan, *Genetic basis for clinical response to CTLA-4 blockade in melanoma. N Engl J Med, 2014. 371(23): p. 2189-2199.*
186. Chumsri, S., Z. Li, D.J. Serie, A. Mashadi-Hosseini, G. Colon-Otero, N. Song, K.L. Pogue-Geile, P.G. Gavin, S. Paik, A. Moreno-Aspitia, E.A. Perez, and E.A. Thompson, *Incidence of Late Relapses in Patients With HER2-Positive Breast Cancer Receiving Adjuvant Trastuzumab: Combined Analysis of NCCTG N9831 (Alliance) and NRG Oncology/NSABP B-31. J Clin Oncol, 2019. 37(35): p. 3425-3435.*
187. Peng, S., J.M. Zaretsky, A.H.C. Ng, W. Chour, M.T. Bethune, J. Choi, A. Hsu, E. Holman, X. Ding, K. Guo, J. Kim, A.M. Xu, J.E. Heath, W.J. Noh, J. Zhou, Y. Su, Y. Lu, J. McLaughlin, D. Cheng, O.N. Witte, D. Baltimore, A. Ribas, and J.R. Heath, *Sensitive*

- Detection and Analysis of Neoantigen-Specific T Cell Populations from Tumors and Blood.* Cell Rep, 2019. **28**(10): p. 2728-2738 e7.
188. Zamora, A.E., J.C. Crawford, E.K. Allen, X.J. Guo, J. Bakke, R.A. Carter, H.A. Abdelsamed, A. Moustaki, Y. Li, T.C. Chang, W. Awad, M.H. Dallas, C.G. Mullighan, J.R. Downing, T.L. Geiger, T. Chen, D.R. Green, B.A. Youngblood, J. Zhang, and P.G. Thomas, *Pediatric patients with acute lymphoblastic leukemia generate abundant and functional neoantigen-specific CD8(+) T cell responses.* Sci Transl Med, 2019. **11**(498).
189. Greiner, J., Y. Ono, S. Hofmann, A. Schmitt, E. Mehring, M. Gotz, P. Guillaume, K. Dohner, J. Mytilineos, H. Dohner, and M. Schmitt, *Mutated regions of nucleophosmin 1 elicit both CD4(+) and CD8(+) T-cell responses in patients with acute myeloid leukemia.* Blood, 2012. **120**(6): p. 1282-9.
190. Rezvani, K., A.S. Yong, S. Mielke, B.N. Savani, L. Musse, J. Superata, B. Jafarpour, C. Boss, and A.J. Barrett, *Leukemia-associated antigen-specific T-cell responses following combined PR1 and WT1 peptide vaccination in patients with myeloid malignancies.* Blood, 2008. **111**(1): p. 236-42.
191. Rezvani, K., A.S. Yong, B.N. Savani, S. Mielke, K. Keyvanfar, E. Gostick, D.A. Price, D.C. Douek, and A.J. Barrett, *Graft-versus-leukemia effects associated with detectable Wilms tumor-1 specific T lymphocytes after allogeneic stem-cell transplantation for acute lymphoblastic leukemia.* Blood, 2007. **110**(6): p. 1924-32.
192. Chapuis, A.G., D.N. Egan, M. Bar, T.M. Schmitt, M.S. McAfee, K.G. Paulson, V. Voillet, R. Gottardo, G.B. Ragnarsson, M. Bleakley, C.C. Yeung, P. Muhlhauser, H.N. Nguyen, L.A. Kropp, L. Castelli, F. Wagener, D. Hunter, M. Lindberg, K. Cohen, A. Seese, M.J. McElrath, N. Duerkopp, T.A. Gooley, and P.D. Greenberg, *T cell receptor gene therapy targeting WT1 prevents acute myeloid leukemia relapse post-transplant.* Nat Med, 2019. **25**(7): p. 1064-1072.

193. Molldrem, J.J., P.P. Lee, C. Wang, K. Felio, H.M. Kantarjian, R.E. Champlin, and M.M. Davis, *Evidence that specific T lymphocytes may participate in the elimination of chronic myelogenous leukemia*. Nat Med, 2000. **6**(9): p. 1018-23.
194. Summers, C., V.S. Sheth, and M. Bleakley, *Minor Histocompatibility Antigen-Specific T Cells*. Front Pediatr, 2020. **8**: p. 284.
195. Lansford, J.L., U. Dharmasiri, S. Chai, S.A. Hunsucker, D.S. Bortone, J.E. Keating, I.M. Schlup, G.L. Glish, E.J. Collins, G. Alatrash, J.J. Molldrem, P.M. Armistead, and B.G. Vincent, *Computational modeling and confirmation of leukemia-associated minor histocompatibility antigens*. Blood Adv, 2018. **2**(16): p. 2052-2062.
196. Oostvogels, R., M.C. Minnema, M. van Elk, R.M. Spaapen, G.D. te Raa, B. Giovannone, A. Buijs, D. van Baarle, A.P. Kater, M. Griffioen, E. Spierings, H.M. Lokhorst, and T. Mutis, *Towards effective and safe immunotherapy after allogeneic stem cell transplantation: identification of hematopoietic-specific minor histocompatibility antigen UTA2-1*. Leukemia, 2013. **27**(3): p. 642-9.
197. Chapuis, A.G., G.B. Ragnarsson, H.N. Nguyen, C.N. Chaney, J.S. Pufnock, T.M. Schmitt, N. Duerkopp, I.M. Roberts, G.L. Pogosov, W.Y. Ho, S. Ochsenreither, M. Wolfli, M. Bar, J.P. Radich, C. Yee, and P.D. Greenberg, *Transferred WT1-reactive CD8+ T cells can mediate antileukemic activity and persist in post-transplant patients*. Sci Transl Med, 2013. **5**(174): p. 174ra27.
198. Tanaka, Y., H. Nakasone, R. Yamazaki, K. Sato, M. Sato, K. Terasako, S. Kimura, S. Okuda, S. Kako, K. Oshima, A. Tanihara, J. Nishida, T. Yoshikawa, T. Nakatsura, H. Sugiyama, and Y. Kanda, *Single-cell analysis of T-cell receptor repertoire of HTLV-1 Tax-specific cytotoxic T cells in allogeneic transplant recipients with adult T-cell leukemia/lymphoma*. Cancer Res, 2010. **70**(15): p. 6181-92.
199. Hunsucker, S.A., C.S. McGary, B.G. Vincent, A.A. Enyenihi, J.P. Waugh, K.P. McKinnon, L.M. Bixby, P.A. Ropp, J.M. Coghill, W.A. Wood, D.A. Gabriel, S. Sarantopoulos, T.C.

- Shea, J.S. Serody, G. Alatrash, T. Rodriguez-Cruz, G. Lizee, A.S. Buntzman, J.A. Frelinger, G.L. Glish, and P.M. Armistead, *Peptide/MHC tetramer-based sorting of CD8(+) T cells to a leukemia antigen yields clonotypes drawn nonspecifically from an underlying restricted repertoire*. *Cancer Immunol Res*, 2015. **3**(3): p. 228-35.
200. Yu, W., N. Jiang, P.J. Ebert, B.A. Kidd, S. Muller, P.J. Lund, J. Juang, K. Adachi, T. Tse, M.E. Birnbaum, E.W. Newell, D.M. Wilson, G.M. Grotenbreg, S. Valitutti, S.R. Quake, and M.M. Davis, *Clonal Deletion Prunes but Does Not Eliminate Self-Specific alphabeta CD8(+) T Lymphocytes*. *Immunity*, 2015. **42**(5): p. 929-41.
201. Smith, S.N., D.T. Harris, and D.M. Kranz, *T Cell Receptor Engineering and Analysis Using the Yeast Display Platform*. *Methods Mol Biol*, 2015. **1319**: p. 95-141.
202. Spindler, M.J., A.L. Nelson, E.K. Wagner, N. Oppermans, J.S. Bridgeman, J.M. Heather, A.S. Adler, M.A. Asensio, R.C. Edgar, Y.W. Lim, E.H. Meyer, R.E. Hawkins, M. Cobbold, and D.S. Johnson, *Massively parallel interrogation and mining of natively paired human TCRalphabeta repertoires*. *Nat Biotechnol*, 2020. **38**(5): p. 609-619.
203. Holtan, S.G., T.E. DeFor, A. Lazaryan, N. Bejanyan, M. Arora, C.G. Brunstein, B.R. Blazar, M.L. MacMillan, and D.J. Weisdorf, *Composite end point of graft-versus-host disease-free, relapse-free survival after allogeneic hematopoietic cell transplantation*. *Blood*, 2015. **125**(8): p. 1333-8.
204. Malandro, N., S. Budhu, N.F. Kuhn, C. Liu, J.T. Murphy, C. Cortez, H. Zhong, X. Yang, G. Rizzuto, G. Altan-Bonnet, T. Merghoub, and J.D. Wolchok, *Clonal Abundance of Tumor-Specific CD4(+) T Cells Potentiates Efficacy and Alters Susceptibility to Exhaustion*. *Immunity*, 2016. **44**(1): p. 179-193.
205. Chen, B., M.S. Khodadoust, N. Olsson, L.E. Wagar, E. Fast, C.L. Liu, Y. Muftuoglu, B.J. Sworder, M. Diehn, R. Levy, M.M. Davis, J.E. Elias, R.B. Altman, and A.A. Alizadeh, *Predicting HLA class II antigen presentation through integrated deep learning*. *Nat Biotechnol*, 2019. **37**(11): p. 1332-1343.

206. Comoli, P., S. Basso, G. Riva, P. Barozzi, I. Guido, A. Gurrado, G. Quartuccio, L. Rubert, I. Lagreca, D. Vallerini, F. Forghieri, M. Morselli, P. Bresciani, A. Cuoghi, A. Paolini, E. Colaci, R. Marasca, A. Cuneo, L. Iughetti, T. Trenti, F. Narni, R. Foa, M. Zecca, M. Luppi, and L. Potenza, *BCR-ABL-specific T-cell therapy in Ph+ ALL patients on tyrosine-kinase inhibitors*. *Blood*, 2017. **129**(5): p. 582-586.

Vita

Haven Rebecca Garber was born in Greensboro, North Carolina, the daughter of Keith and Janice Garber. She attended public school in Greensboro NC (Alamance Elementary), Tupelo MS (Martin Luther King Jr. Intermediate School), and Pickerington OH (Pickerington Junior High and High School). She received the degree of Bachelor of Science in Engineering (B.S.E) with a major in Biomedical Engineering and a minor in Chemistry from Duke University, graduating cum laude. She attended medical school at The Ohio State University College of Medicine in Columbus, Ohio and graduated magna cum laude. She completed Internal Medicine Residency at Beth Israel Deaconess Medical Center in Boston, Massachusetts and Hematology and Oncology Fellowship at The University of Texas MD Anderson Cancer Center in Houston, Texas. She received the 2020 Lyndon B. Johnson 3rd Year Fellow of the Year Award for her clinical oncology work at the community hospital.