

T-Cell Clonality And Genetic Heterogeneity In Cutaneous T-cell Lymphoma

By

Aishwarya Iyer

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Medicine
University of Alberta

© Aishwarya Iyer, 2019

Abstract

Cutaneous T-cell lymphoma (CTCL) is heterogeneous form of T-cell lymphoma of which, Mycosis fungoides (MF) is the most common entity. In early stages MF presents on skin as a scaly plaques that in advanced stages develop into tumors and may disseminate to other sites such as lymph node and central nervous system. There no prognostic markers of disease progression. The disease is incurable, currently available therapies providing temporary remissions but the disease inadvertently relapses.

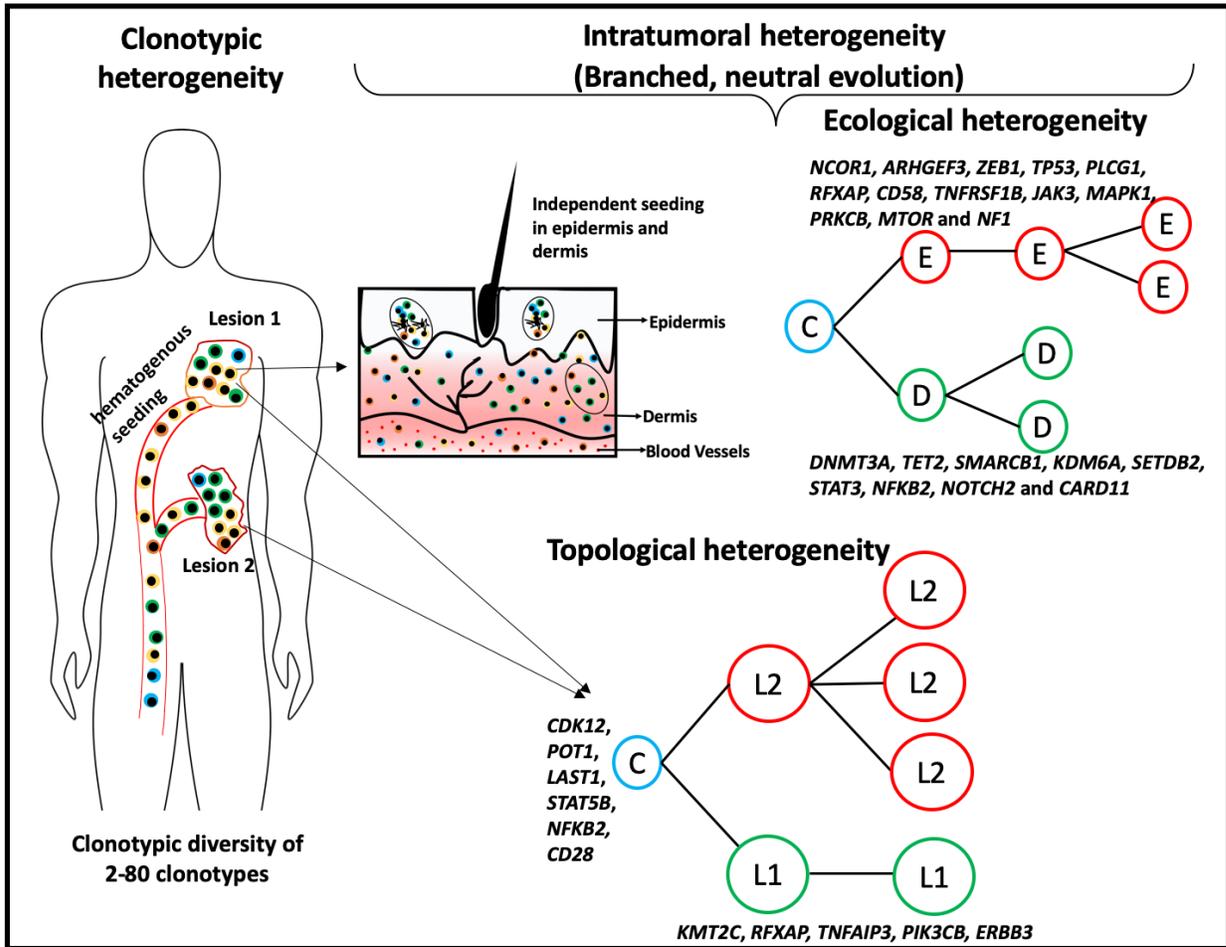
The long-standing dogma in MF is that it develops from a mature, skin resident T-cell and progresses in a linear fashion from plaques to tumors and finally disseminated to extracutaneous sites. The purpose of this thesis was to critically test this assumption. Based on previous studies and clinical observations we formulated a working hypothesis that MF is a genetically and clonally heterogeneous disease. The testable consequence of this hypothesis was that MF does not develop from the skin-resident T-cell but must originate by seeding of transformed cells from the circulation to the skin leading to a multifocal development.

We decided to use whole exome sequencing (WES) as the primary research method taking advantage from the fact that WES allows to quantify the tumor cell fraction, identify spectrum of genomic mutations, cancer subclones and T-cell receptor (TCR) repertoire. We were able to show conclusively that MF is not a monoclonal lymphoproliferation but comprised multiple TCR- α , β , and γ clonotypes indicating presence of clonal heterogeneity. Next, we compared the TCR sequences from skin with those in the circulating blood and identified that neoplastic TCR clonotypes from skin are present in the circulation, even in the early stages of MF. Additional sampling of skin lesions from different areas on the body and longitudinal sampling presented a model of MF, in which the lesions do not arise from skin-resident T-

cells but are initiated by seeding of transformed cells to the skin from the circulation. Further analysis of the molecular architecture of MF, we catalogued genetic abnormalities (somatic variants and copy number aberrations) and used this information to determine the fraction of genetic aberrations which are clonal (common in most of cancer cells) and that are subclonal (present in a subpopulation of malignant cells). We observed that significant proportion of mutations were present in the subclones indicating genetic heterogeneity. The phylogenetic deconvolution of the genetic aberrations presented MF to be evolved via unrestricted branched evolution producing an increasing heterogeneity and mutational complexity. Tumor cells in different lesions and in different tissue niches within the same lesion (epidermis vs dermis) showed independent evolution with no evidence of competition or cell transfer between the compartments.

We believe that our studies meaningfully advance the knowledge on the molecular pathogenesis of MF. We propose a model for MF development, where the skin lesions originate by seeding from the pool of circulating precursor cells in the blood. Those cells proliferate in the skin, accumulate mutations and further branch into multiple genetic subclones. Our model might provide a better understanding of the mechanism of disease progression, treatment resistance and relapse. Further studies by single cell sequencing techniques would increase the accuracy of identifying clonal and subclonal populations for targeted therapies and immunotherapies.

Visual Abstract



Heterogeneity of Mycosis fungoides (MF).

Skin lesions in MF are seeded with malignant T-cells via the circulation, leading to heterogeneous mixture of T-cells. Analysis of genomic alternations within these T-cells groups them in a number of distinct subclones that present neutral divergent evolution.

Preface

This thesis contains both published and unpublished work, which is the original work done by me. All the work listed in the following chapters received ethics approval number HREBA.CC-16-0820-REN1 from Health Research Ethics Board of Alberta, Cancer Committee. In the studies mentioned below, I was responsible for the study design, data assembly and writing of manuscript under the supervision of Robert Gniadecki. Sample preparation for sequencing was done by me and Sandra O’Keefe. Dylan Hennessey and Jordan Patterson helped with the bioinformatics analysis. Weiwei Wang and Gane Ka-Shu Wong provided guidance with the experimental design and analyses.

Chapter 2 of this thesis is published as A. Iyer, D. Hennessey, S. O’Keefe, J. Patterson, W. Wang, G. Ka-Shu Wong, R. Gniadecki, Clonotypic heterogeneity in cutaneous T-cell lymphoma (mycosis fungoides) revealed by comprehensive whole-exome sequencing. *Blood Adv.* 2019;3(7):1175–1184.

Chapter 3 of this thesis is published as A. Iyer, D. Hennessey, S. O’Keefe, J. Patterson, W. Wang, G. Ka-Shu Wong, R. Gniadecki, Skin colonization by circulating neoplastic clones in cutaneous T-cell lymphoma, *Blood* (2019): blood.2019002516. Web. 13 Sept2019.

Chapter 4 of this thesis is published online as a preprint in the format of A. Iyer, D. Hennessey, S. O’Keefe, J. Patterson, W. Wang, G. Ka-Shu Wong, R. Gniadecki, Branched evolution and genomic intratumor heterogeneity in the pathogenesis of cutaneous T-cell lymphoma. *bioRxiv.* 2019;804351. This chapter will subsequently be submitted for peer review publication.

Chapter 5 of this thesis will be submitted for peer review publication as A. Iyer, D.

Hennessey, S. O'Keefe, J. Patterson, W. Wang, G. Ka-Shu Wong, R. Gniadecki, Independent evolution of cutaneous lymphoma subclones in different microenvironments of the skin.

Dedication

This thesis is dedicated to my brother, Shriram Iyer, who encouraged and supported me to pursue my doctorate through all the difficulties.

This thesis is also dedicated to my parents, Latha and Ganesh Iyer for their unconditional love and support.

Acknowledgements

First and foremost, I would like to thank my supervisor Dr. Robert Gniadecki for his incredible support and guidance. I am very thankful for his patience in dealing with my stubbornness! The work in this thesis would not have been possible without his constant encouragement and immense faith in me. I am extremely thankful for having him as my mentor. My time spent under his supervision has not only matured my knowledge in science but also prepared me for real life challenges. I am not only going to miss working with him but also the long discussions we have had on the innumerable topics. The past three years spent in his lab have been fun and a great learning experience.

I would like to thank my co-supervisor Dr. Gane Ka-Shu Wong for his valuable feedback, support and suggestions. His advice on my projects has immensely improved the quality of my work. I would like to thank my committee member Dr. Raymond Lai for his support to finish my thesis work in a timely manner.

I would also like to acknowledge the efforts of Dr. Thomas Salopek and the nursing staff at Kaye Clinic, Edmonton, Alberta for their help with sample collection. I would also like to express my gratitude towards the funding agencies Canadian Dermatology Foundation, University Hospital Foundation (University of Alberta), Bispebjerg Hospital (Copenhagen, Denmark), Danish Cancer Society and Department of Medicine, University of Alberta for their generous sponsorships that led to the successful completion of the projects in this thesis.

I would next like to thank my lab members Dylan Hennessey and Sandra O'Keefe for their timely help with my experiments and analysis. I would like to thank lab members of Dr. Wong-Weiwei Wang, Tracy Jordan and Jordan Patterson for training and helping me understand the complex protocols of nucleic acid sequencing and analysis. I would like to

thank the CEGIIR and TAGC facility for access to the instruments. I would like to mention the incredible help of Phyllis Ewasiw and Barb Thomson. This graduate work would have been an impossible feat without their valuable support! I would also like to acknowledge the support of my friends and colleagues Megan Yap and Zoulika Zak whose advice helped me make some of my most difficult and valuable choices during graduate studies.

Finally, I would like to thank my friends and family for their patience and support during this long journey. Their understanding and support are something I will never be able to repay. Especially my mom for being the example of an incredibly strong woman and my dad for helping me become independent person that I am today. Lastly, I would like to thank every unnamed person who indirectly contributed in making this thesis a possibility.

Table of contents

Abstract.....	ii
Visual Abstract.....	iv
Preface.....	v
Dedication.....	vii
Acknowledgements.....	viii
Table of contents.....	x
List of figures.....	xii
List of abbreviations.....	xiv
Key Definitions.....	xvi
Chapter 1: Introduction.....	1
1.1 Cutaneous T-Cell Lymphoma.....	1
1.1.1 Mycosis Fungoides.....	1
1.1.2 Sezary Syndrome.....	3
1.2 Relevance of T-Cell Receptor sequences in CTCL.....	3
1.3 Cytogenetics and Genomics of CTCL.....	5
1.4 Intra-Tumor Heterogeneity (ITH).....	6
1.5 Rationale and Objectives.....	7
1.6 Hypothesis.....	8
1.7 References.....	9
Chapter 2: Clonotypic heterogeneity in cutaneous T-cell lymphoma reveled by comprehensive whole exome sequencing*.....	17
2.1 Abstract.....	17
2.2 Background.....	19
2.3 Material and Methods.....	21
2.4 Results.....	23
2.5 Discussion.....	32
2.6 References.....	36
Chapter 3: Skin colonization by circulating neoplastic clones in cutaneous T-cell lymphoma*.....	41
3.1 Abstract.....	41

3.2 Background.....	43
3.3 Material and Methods.....	45
3.4 Results	48
3.5 Discussion	57
3.6 References.....	64
Chapter 4: Branched evolution and genomic intratumor heterogeneity in the pathogenesis of cutaneous T-cell lymphoma*.....	74
4.1 Abstract.....	74
4.2 Background.....	76
4.3 Material and Methods.....	78
4.4 Results	79
4.5 Discussion	89
4.6 References.....	95
Chapter 5: Independent evolution of cutaneous lymphoma subclones in different microenvironments of the skin	101
5.1 Abstract.....	101
5.2 Background.....	102
5.3 Material and Methods.....	104
5.4 Results	105
5.5 Discussion	112
5.6 References.....	116
Chapter 6: Discussion and Conclusion.....	121
6.1 General Discussion	121
6.2 Conclusion and perspectives.....	122
Appendix A	154
Appendix B	157
Appendix C.....	170
Appendix D	187

List of figures

Figure 1.1: Histology of MF.....	2
Figure 1.2: T-cell receptor (TCR) rearrangement.	5
Figure 2.1: Schematic representation of sample collection, processing, and TCR sequencing.	25
Figure 2.2: Efficiency of WES probe capture and WTS protocols in the detection of CDR3 clonotypes in MF biopsies.....	26
Figure 2.3: Relative frequency of T-cell clonotypes.	29
Figure 2.4: Clonotypic diversity of MF.....	30
Figure 2.5: Shared T-cell clonotypes.....	31
Figure 3.1: Schematic representation of sample collection and processing.	49
Figure 3.2: Clonotypic heterogeneity of skin lesions in MF.	50
Figure 3.3: Detection of the neoplastic TCR β clonotypes in the peripheral blood.	53
Figure 3.4: Dynamics of neoplastic clonotypes in the skin and the blood.	56
Figure 3.5: The hypothesis of tumor seeding in the pathogenesis of mycosis fungoides.	56
Figure 4.1: Summary of experimental methods and data analysis.	80
Figure 4.2: Mutational landscape of putative driver genes in MF.....	81
Figure 4.3: Genomic copy number changes in MF.	83
Figure 4.4: Intratumoral heterogeneity in MF.	84
Figure 4.5: Distribution of mutations in the stem and clades in MF.	87
Figure 4.6: Phylogenetic relationship between different lesions in the same patient: topologic heterogeneity in MF.	89
Figure 4.7: Proposed model of the evolution of MF.	93
Figure 5.1: Clonotypic heterogeneity and tumour cell seeding of the skin microenvironment in MF.	107
Figure 5.2: Mutational landscape of putative driver genes in anatomical layers of skin.	109
Figure 5.3: Evolutionary facets of the genetic clones in skin microenvironment.	110

Figure 5.4: Phylogenetic analysis of the neoplastic T-cells in skin microenvironment. 111

Figure 5.5: Generation of ecological heterogeneity in MF. 113

List of abbreviations

CTCL	Cutaneous T-Cell Lymphoma
MF	Mycosis Fungoides
SS	Sezary Syndrome
TCR	T Cell Receptor
CDR3	Complementary Determining Region 3
WES	Whole Exome Sequencing
WTS	Whole Transcriptome Sequencing
OCT	Optimal Cutting Temperature
PBMC	Peripheral Blood Mononuclear Cells
FBS	Fetal Bovine Serum
DMSO	DiMethyl SulfOxide
PEN	PolyEthylene Naphthalate
RPMI	Roswell Park Memorial Institute
PUVA	Psoralen UltraViolet A
LCM	Laser Capture Microdissection
SNPs	Single Nucleotide Polymorphism
SVs	Somatic Variants
SSMs	Single Somatic Mutations

CNAs	Copy Number Aberrations
NGS	Next Generation Sequencing
ESP	Early Stage Plaque
LSP	Late Stage Plaque
TMR	Tumor
E (Epi)	Epidermis
D (Der)	Dermis
P	Plaque
T	Tumor

Key Definitions

1. Clonotype- single rearranged VJ or VDJ sequences of TCR coreceptors (α , β , γ and δ).
2. Clones- The clone represents a population of cells originating from a single precursor.
The term “T-cell clone” has been defined operationally as group of lymphocytes exhibiting the same clonotype. In this thesis we also use the term “genetic clone” which is defined as genetically identical cells and the term “clonal” which in the context of mutations signifies genetic aberrations found in all cells in the tumor (as opposed to “subclonal”- see below).
3. Subclone- a term used here in the context of mutational heterogeneity of the tumor signifying a group of malignant cells presenting a common genomic mutational profile of SVs and CNAs.
4. Phylogenetic tree- representation of evolutionary relation between the genetic subclones.
5. Neutral evolution- is a pattern of mutational evolution of cancer when the generations of cancer cells harboring new mutations are under no, or minimal selection pressure.
6. Darwinian evolution- is a theory that certain genetic aberrations are more fundamental for the growth and development of the tumor and are therefore, under selection pressure over other random genetic aberrations.
7. Divergent evolution- also defined as branched evolution, is an evolutionary process in which a single genetic clone diverges to form two or more subclones.
8. Convergent evolution- an evolutionary process in which two genetic clones converge to form a single clone.

Chapter 1: Introduction

1.1 Cutaneous T-Cell Lymphoma

Primary cutaneous T-cell lymphomas are lymphoid malignancies that clinically manifest themselves in the skin. They comprise the second most common form of extranodal non-Hodgkin's lymphomas.^{1,2} The annual incident rate is 6.4 cases per million.³ Mycosis fungoides (MF) and Sezary syndrome are the most common types of CTCL⁴⁻⁷ which occur in adults with median age of 56 at diagnosis.

1.1.1 Mycosis Fungoides

Mycosis fungoides comprises approximately 50% of all CTCL.^{3,7} Skin lesions in early stages of MF are comprised of red scaly patches or plaques that in advance stages develops into tumors. MF may disseminate to other organs such as lymph nodes and the central nervous system. The TNMB staging system of the skin (T), lymph nodes (N), viscera (M) and the blood (B) classifies patients into four clinical stages (I-IV). Patients with stages I-IIA, which is characterized by the presence of patches (T1) and plaques (T2) in the skin without involvement of other organs. These are considered to be early stages because patients have favorable prognosis and the life expectancy is similar to the healthy age-matched population. The hallmark of disease progression from early to advanced stage is occurrence of skin tumors (T3), an event that increases an overall risk of systemic spread to 30% of the cases. Five-year survival in stage IIB is 5 years and further drops to 1.4 year in stage IVB.^{3,7,8} Therapy is individualized and based on patient age, extent of disease burden and risk of progression. In early stages, the treatment is limited to skin directed therapies (e.g. psoralen UVA

phototherapy (PUVA), recombinant beta-interferon, retinoids) but in late stages a combination of skin directed therapies, targeted therapies and chemotherapeutic agents is required.⁹ Even with complete eradication of skin lesions the disease eventually relapses and complete cure is not achievable.

Histology of MF

The definitive diagnosis of MF in early stages of patch or plaque is difficult as many of the clinical and histopathological markers mimic with those of skin inflammatory disease such as dermatitis.^{10,11} The distinguishable features in early stage lesions are the lymphoid infiltrates presenting with different degrees of atypical nuclei. These lymphocytes are often located in the basal layer of epidermis and associated with vacuolar degeneration of the basal epidermal layer. Pautrier microabscesses, which consists of small aggregates of atypical lymphocytes and Langerhans cells in the epidermis are a useful marker for diagnosis but are present in only 25% of the cases (**Fig 1**).^{12,13} Immunohistochemical staining usually reveal positivity for CD3 and CD4 and variable loss of common lymphocytic markers such as CD5 and CD7.

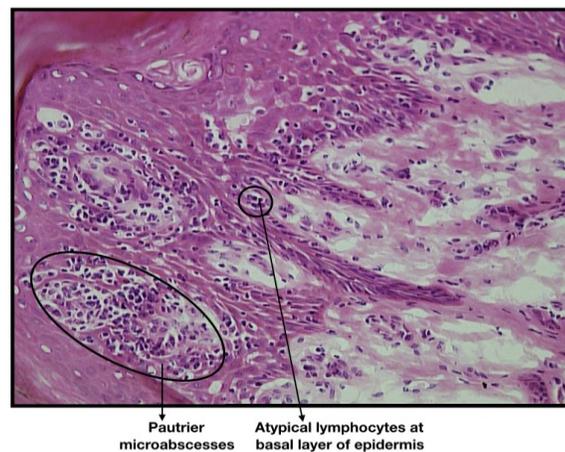


Figure 1.1: Histology of MF.

Hematoxylin and eosin staining of a section from MF skin lesion under 20x magnification.

1.1.2 Sezary Syndrome

Sezary Syndrome (SS) is a leukemic subtype of CTCL. Patients present with erythroderma, lymphadenopathy, and circulating malignant T-cells.¹⁴⁻¹⁶ It presents more aggressive clinical behaviour and poor prognosis as compared to MF and other CTCL subtypes.^{7,17} Sezary cells are large, atypical T lymphocytes with convoluted nuclei and are a marker for identification of SS. But as these cells are also found in normal people and in patients with inflammatory diseases, additional markers such as absolute Sezary cell count $\geq 1000/\mu\text{l}$, an increased number of CD4⁺, aberrant expression of pan-T-cell antigens (loss of CD7, CD26) and demonstration of T-cell clonality in the PCR-based method are considered for diagnosis of SS.^{6,18} Though SS is considered as a continuum of MF, recent studies indicate that these may be distinct diseases developing from different subset of T-cells.^{16,19-21} T-cells isolated from SS samples are known to express markers of the central memory T-cells (CCR7 and L-selectin) whereas, T-cells in MF express skin resident T-cell markers (CCR4 and CLA).

1.2 Relevance of T-Cell Receptor sequences in CTCL

TCR is a membrane bound heterodimer composed of two polypeptide chains ($\alpha\beta$ or $\gamma\delta$) linked by a disulfide bond. In the peripheral blood, most T cells express $\alpha\beta$ receptor and up to 10% express $\gamma\delta$ receptor.²²⁻²⁴ The TCR gene (*TRA*, *TRB*, *TRG* and *TRD*) loci contain many different variable (V), diversity (D) and joining (J) gene segments. Random selection of VDJ (*TRB* and *TRD*) or VJ (*TRA* and *TRG*) generate a diverse T cell repertoire and the selected V(D)J are joined using site specific recombinases RAG1 and RAG2 (**Fig 1.2A**).

Recombination of *TRD*, *TRG* and *TRB* occur sequentially during the CD4⁻ CD8⁻ double-

negative (DN)2, DN1 and DN3 stages of thymocyte development. Successful recombination of *TRD* and *TRG* promotes assembly of $\gamma\delta$ TCR, whereas successful recombination of *TRB* generates TCR β and pre-TRC α . After several rounds of proliferation and differentiation to CD4⁺CD8⁺ double positive (DP) stage these cells rearrange *TRA* genes to generate $\alpha\beta$ TCR (**Fig 1.2B**).²²⁻²⁴ Normal T-cells never coexpress $\alpha\beta$ or $\gamma\delta$ because δ locus is nested within *TRA* and excised during *TRA* gene recombination. However, TCR γ rearrangement is retained in all T-cells, even if it is not expressed. *TRB* rearrangement is usually monoallelic (allelic exclusion), but this rule does not apply to *TRG* or *TRA* loci where rearrangements can occur on both chromosomes.²²⁻²⁶ Recent studies with mathematical modeling and single cell sequencing have indicated that 20-30% of cells express multiple TCR α but only one of the sequences is expressed on the cell surface.^{27,28}

For diagnostic purposes, TCR rearrangement can be detected by polymerase chain reaction (PCR) using BIOMED-2 primers and analyzing the results by GeneScan. In GeneScan, fluorochrome-labeled PCR products of rearranged TCR genes are denatured prior to fragment analysis. Monoclonal samples will give rise to a single dominant peak, representing PCR products of identical size, whereas in the case of polyclonal samples the PCR products of numerous sizes will show distribution of peaks.²⁹ Identification of monoclonal TCR rearrangement is considered to be diagnostic for CTCL since it is believed that these neoplasms arise from mature T-cells.³⁰ Recent studies have used next generation sequencing techniques to replace the GeneScan analysis method but these studies rely heavily on sequencing of a single TCR coreceptor (γ or β) or the use of RNA to identify the rearranged sequences.

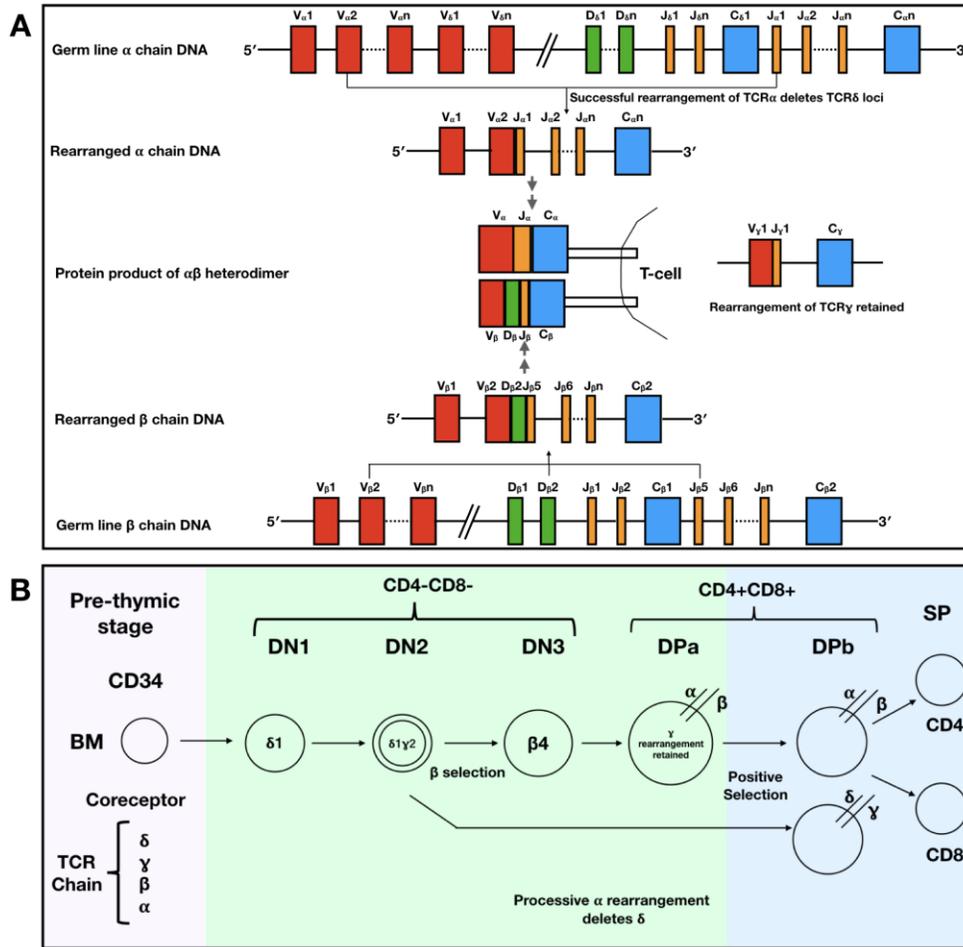


Figure 1.2: T-cell receptor (TCR) rearrangement.

T-cells during development undergo rearrangement at the coreceptor genes (α , β , γ and δ). (A) Each coreceptor is composed of multiple VJ or VDJ genes that are randomly selected and rearranged to form a functional receptor. (B) T-cells in the thymus during development undergo TCR rearrangement at coreceptor genes. The rearrangement happens sequentially starting with TCR δ , followed by TCR γ and β during DN1-DN3 stage of development. During the DP stage the T-cells that are committed to express TCR $\gamma\delta$ will undergo positive selection. The cells committed to express TCR $\alpha\beta$ chains will undergo rearrangement at the coreceptor α gene which then undergo positive selection before being released into blood circulation.

1.3 Cytogenetics and Genomics of CTCL

Multiple genomic array based studies have implicated Copy number aberrations (CNAs) as prognostically relevant in CTCL.³¹⁻⁴³ Studies have implicated the loss of chromosome 9p21 in

correlation with a decrease in overall survival.^{36,44} Recent studies in CTCL have also investigated the implications of CNAs with particular emphasis on the differences in MF and SS. Increased copy number gains was observed in chromosome 1 and 7 for MF and chromosome 8 and 17 for SS. Deletions were observed at chromosome 9 for MF and chromosome 10 for SS. These significant differences of CNAs between MF and SS implicate that the two entities have independent genetic evolution and therefore, require independent investigations.⁴⁵

Previous genomic studies in CTCL focused on SS (~84% of all sequenced cases of CTCL) and less so on MF.⁴⁶⁻⁵⁴ The few studies that included MF samples implicated genes mutated in NF-kB, JAK- STAT, T-cell activation and DNA repair pathways play an important role in T cell development and proliferation.^{32,46,47} Promoter methylation, epigenetic instability and mutations in many tumor suppressor genes, including those involved in the induction of apoptosis, appears to be commonly employed mechanisms of lymphomagenesis in CTCL.^{32,49} However, a recent meta-analysis study indicated that mutations identified in tumor suppressor genes in CTCL were largely presented in SS samples but not in MF.⁵⁰ Mutations and molecular marker(s) that are catalogued vary between different patients suggest intratumor heterogeneity in CTCL.

1.4 Intra-Tumor Heterogeneity (ITH)

Recent studies with solid tumors have documented that many malignant tumors are genetically heterogeneous. The degree of intratumor heterogeneity can be highly variable, in some instances reaching up to 8000 non-synonymous mutations in tumor subclones.⁵⁵

Intratumor heterogeneity explains the process of tumor evolution, metastasis, resistance to

treatment and factors responsible for tumor relapse. ITH implies that the tumors are composed of multiple genetic subclones with different properties to proliferate and metastasize. ITH in terms of degree of subclonality and its evolution within a tumor correlates with response to treatment is presented in lung, melanoma and prostate cancer.⁵⁶⁻⁵⁹ To this end, few studies with bulk or single cell RNA sequencing studies have implicated tumor heterogeneity in CTCL⁶⁰⁻⁶³ but there are no studies to date exploring the tumor heterogeneity at the genomic level.

1.5 Rationale and Objectives

The current model of MF pathogenesis assumes that the disease develops from a single, transformed, skin resident memory T-cell that does not recirculate and evolves in a linear fashion by stepwise progression from early patch to plaque and finally to tumors and metastases.^{16,64-66} This “mature T-cell model” implies that MF is clonal (i.e. all malignant cells comprise the same clonotype). However, it cannot easily explain why MF emerges simultaneously in many different areas of the skin rather than as a single skin lesion. It is also difficult to understand the mechanism of skin relapses which inevitably occur even after the most powerful cytoreductive therapies such as total skin electron beam radiation. It is even more puzzling that the relapses occur in multiple skin regions, not necessarily within the previously affected areas. The obvious explanation that skin lesions in MF are initiated by circulating, rather than skin-resident malignant T-cells has never been challenged experimentally.

In our research we were motivated by gaps in knowledge regarding the molecular pathogenesis of MF. Therefore, the first objective of this study was to develop a technique to

identify TCR clonotypes of malignant cells. Second, we asked whether malignant T-cell clones are present in circulating blood in early and late stage MF patients. Lastly, we decided to determine the genetic landscape of MF skin lesions to study clonality of malignant cells on the genetic level.

1.6 Hypothesis

According to the current model, MF is a clonal proliferation of malignant T-cells comprising a single clonotype, that primarily develops in the skin, and progresses in a linear fashion by accumulation of mutations.

Our alternative research hypothesis is that MF is composed of multiple neoplastic T-cell clones, spreads via hematogenous seeding and presents substantial genetic heterogeneity.

1.7 References

1. Rizvi MA. T-cell non-Hodgkin lymphoma. *Blood*. 2006;107(4):1255–1264.
2. Ansell SM. Non-Hodgkin Lymphoma: Diagnosis and Treatment. *Mayo Clinic Proceedings*. 2015;90(8):1152–1163.
3. Vd C, Ma W. Incidence of cutaneous t-cell lymphoma in the united states, 1973-2002. *Arch. Dermatol*. 2007;143(7):854–859.
4. Rodd AL, Ververis K, Karagiannis TC. Current and Emerging Therapeutics for Cutaneous T-Cell Lymphoma: Histone Deacetylase Inhibitors. *Lymphoma*. 2012;2012:1–10.
5. Bagherani N, Smoller BR. An overview of cutaneous T cell lymphomas. *F1000Research*. 2016;5:1882.
6. Wilcox RA. Cutaneous T-cell lymphoma: 2016 update on diagnosis, risk-stratification, and management. *Am. J. Hematol*. 2016;91(1):151–165.
7. Willemze R, Jaffe ES, Burg G, et al. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768 LP–3785.
8. Olsen E, Vonderheid E, Pimpinelli N, et al. Revisions to the staging and classification of mycosis fungoides and Sézary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the cutaneous lymphoma task force of the European Organization of Research and Treatment of Ca. *Blood*. 2007;110(6):1713 LP–1722.
9. Huber MA, Kunzi-Rapp K, Staib G, Scharffetter-Kochanek K. Management of refractory early-stage cutaneous T-cell lymphoma (mycosis fungoides) with a combination of oral

- bexarotene and psoralen plus ultraviolet bath therapy. *Journal of the American Academy of Dermatology*. 2004;50(3):475–476.
10. Foo SH, Shah F, Chaganti S, Stevens A, Scarisbrick JJ. Unmasking mycosis fungoides/Sézary syndrome from preceding or co-existing benign inflammatory dermatoses requiring systemic therapies: patients frequently present with advanced disease and have an aggressive clinical course. *Br. J. Dermatol*. 2016;174(4):901–904.
 11. Eklund Y, Aronsson A, Schmidtchen A, Relander T. Mycosis Fungoides: A Retrospective Study of 44 Swedish Cases. *Acta Derm. Venereol*. 2016;96(5):669–673.
 12. Song SX, Willemze R, Swerdlow SH, Kinney MC, Said JW. Mycosis fungoides: report of the 2011 Society for Hematopathology/European Association for Haematopathology workshop. *Am. J. Clin. Pathol*. 2013;139(4):466–490.
 13. Harvey NT, Spagnolo DV, Wood BA. “Could it be mycosis fungoides?”: an approach to diagnosing patch stage mycosis fungoides. *Journal of Hematopathology*. 2015;8(4):209–223.
 14. Trautinger F, Knobler R, Willemze R, et al. EORTC consensus recommendations for the treatment of mycosis fungoides/Sézary syndrome. *Eur. J. Cancer*. 2006;42(8):1014–1030.
 15. Hwang ST, Janik JE, Jaffe ES, Wilson WH. Mycosis fungoides and Sézary syndrome. *Lancet*. 2008;371(9616):945–957.
 16. Campbell JJ, Clark RA, Watanabe R, Kupper TS. Sezary syndrome and mycosis fungoides arise from distinct T-cell subsets: a biologic rationale for their distinct clinical behaviors. *Blood*. 2010;116(5):767–771.

17. Compton CC, Byrd DR, Garcia-Aguilar J, et al. AJCC Cancer Staging Atlas: A Companion to the Seventh Editions of the AJCC Cancer Staging Manual and Handbook. Springer Science & Business Media; 2012.
18. Vonderheid EC, Bernengo MG, Burg G, et al. Update on erythrodermic cutaneous T-cell lymphoma: report of the International Society for Cutaneous Lymphomas. *J. Am. Acad. Dermatol.* 2002;46(1):95–106.
19. Shin J, Monti S, Aires DJ, et al. Lesional gene expression profiling in cutaneous T-cell lymphoma reveals natural clusters associated with disease outcome. *Blood.* 2007;110(8):3015–3027.
20. Kim YH, Liu HL, Mraz-Gernhard S, Varghese A, Hoppe RT. Long-term outcome of 525 patients with mycosis fungoides and Sezary syndrome: clinical prognostic factors and risk for disease progression. *Arch. Dermatol.* 2003;139(7):857–866.
21. van Doorn R, van Kester MS, Dijkman R, et al. Oncogenomic analysis of mycosis fungoides reveals major differences with Sezary syndrome. *Blood.* 2009;113(1):127–136.
22. Bassing CH, Swat W, Alt FW. The Mechanism and Regulation of Chromosomal V(D)J Recombination. *Cell.* 2018;109(2):S45–S55.
23. Hayday AC, Pennington DJ. Key factors in the organized chaos of early T cell development. *Nature Immunology.* 2007;8(2):137–144.
24. Krangel MS. Gene segment selection in V(D)J recombination: accessibility and beyond. *Nat. Immunol.* 2003;4(7):624–630.
25. Xiong N, Kang C, Raulet DH. Redundant and Unique Roles of Two Enhancer Elements in the TCR γ Locus in Gene Regulation and $\gamma\delta$ T Cell Development. *Immunity.* 2002;16(3):453–463.

26. Cobb RM, Oestreich KJ, Osipovich OA, Oltz EMBT-. A in I. Accessibility Control of V(D)J Recombination. 2006;91:45–109.
27. Dupic T, Marcou Q, Walczak AM, Mora T. Genesis of the $\alpha\beta$ T-cell receptor. *PLoS Comput. Biol.* 2019;15(3):e1006874.
28. Han A, Glanville J, Hansmann L, Davis MM. Corrigendum: Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat. Biotechnol.* 2015;33(2):210.
29. van Dongen JJM, Langerak AW, Brüggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: Report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia.* 2003;17:2257.
30. Thurber SE, Zhang B, Kim YH, et al. T-cell clonality analysis in biopsy specimens from two different skin sites shows high specificity in the diagnosis of patients with suggested mycosis fungoides. *J. Am. Acad. Dermatol.* 2007;57(5):782–790.
31. Barba G, Matteucci C, Girolomoni G, et al. Comparative genomic hybridization identifies 17q11.2 approximately q12 duplication as an early event in cutaneous T-cell lymphomas. *Cancer Genet. Cytogenet.* 2008;184(1):48–51.
32. da Silva Almeida AC, Abate F, Khiabani H, et al. The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome. *Nat. Genet.* 2015;47(12):1465–1470.
33. Prochazkova M, Chevret E, Mainhaguet G, et al. Common chromosomal abnormalities in mycosis fungoides transformation. *Genes Chromosomes Cancer.* 2007;46(9):828–838.
34. Fischer TC, Gellrich S, Muche JM, et al. Genomic aberrations and survival in cutaneous T cell lymphomas. *J. Invest. Dermatol.* 2004;122(3):579–586.

35. Carbone A, Bernardini L, Valenzano F, et al. Array-based comparative genomic hybridization in early-stage mycosis fungoides: recurrent deletion of tumor suppressor genes BCL7A, SMAC/DIABLO, and RHOV. *Genes Chromosomes Cancer*. 2008;47(12):1067–1075.
36. Salgado R, Servitje O, Gallardo F, et al. Oligonucleotide array-CGH identifies genomic subgroups and prognostic markers for tumor stage mycosis fungoides. *J. Invest. Dermatol*. 2010;130(4):1126–1135.
37. Karenko L, Kähkönen M, Hyytinen ER, Lindlof M, Ranki A. Notable losses at specific regions of chromosomes 10q and 13q in the Sézary syndrome detected by comparative genomic hybridization. *J. Invest. Dermatol*. 1999;112(3):392–395.
38. So CC, Wong KF, Siu LL, Kwong YL. Large cell transformation of Sézary syndrome. A conventional and molecular cytogenetic study. *Am. J. Clin. Pathol*. 2000;113(6):792–797.
39. Mao X, Lillington D, Scarisbrick JJ, et al. Molecular cytogenetic analysis of cutaneous T-cell lymphomas: identification of common genetic alterations in Sézary syndrome and mycosis fungoides. *Br. J. Dermatol*. 2002;147(3):464–475.
40. Mao X, Lillington DM, Czepulkowski B, et al. Molecular cytogenetic characterization of Sézary syndrome. *Genes Chromosomes Cancer*. 2003;36(3):250–260.
41. Vermeer MH, van Doorn R, Dijkman R, et al. Novel and highly recurrent chromosomal alterations in Sézary syndrome. *Cancer Res*. 2008;68(8):2689–2698.
42. Caprini E, Cristofolletti C, Arcelli D, et al. Identification of key regions and genes important in the pathogenesis of sezary syndrome by combining genomic and expression microarrays. *Cancer Res*. 2009;69(21):8438–8446.

43. Boudjarane J, Essaydi A, Farnault L, et al. Characterization of the novel Sezary lymphoma cell line BKP1. *Exp. Dermatol.* 2015;24(1):60–62.
44. Laharanne E, Oumouhou N, Bonnet F, et al. Genome-wide analysis of cutaneous T-cell lymphomas identifies three clinically relevant classes. *J. Invest. Dermatol.* 2010;130(6):1707–1718.
45. Gug G, Huang Q, Chiticariu E, Solovan C, Baudis M. DNA copy number imbalances in primary cutaneous lymphomas. *Journal of the European Academy of Dermatology and Venereology.* 2019;33(6):1062–1075.
46. McGirt LY, Jia P, Baerenwald DA, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. *Blood.* 2015;126(4):508–519.
47. Ungewickell A, Bhaduri A, Rios E, et al. Genomic analysis of mycosis fungoides and Sézary syndrome identifies recurrent alterations in TNFR2. *Nat. Genet.* 2015;47(9):1056–1060.
48. Wang L, Ni X, Covington KR, et al. Genomic profiling of Sézary syndrome identifies alterations of key T cell signaling and differentiation genes. *Nat. Genet.* 2015;47(12):1426–1434.
49. Moffitt AB, Dave SS. Clinical Applications of the Genomic Landscape of Aggressive Non-Hodgkin Lymphoma. *J. Clin. Oncol.* 2017;35(9):955–962.
50. Park J, Yang J, Wenzel AT, et al. Genomic analysis of 220 CTCLs identifies a novel recurrent gain-of-function alteration in RLTPR (p.Q575E). *Blood.* 2017;130(12):1430 LP–1440.
51. Tensen CP. PLCG1 Gene Mutations in Cutaneous T-Cell Lymphomas Revisited. *J. Invest. Dermatol.* 2015;135(9):2153–2154.

52. Prasad A, Rabionet R, Espinet B, et al. Identification of Gene Mutations and Fusion Genes in Patients with Sézary Syndrome. *J. Invest. Dermatol.* 2016;136(7):1490–1499.
53. Kiel MJ, Sahasrabudhe AA, Rolland DCM, et al. Genomic analyses reveal recurrent mutations in epigenetic modifiers and the JAK–STAT pathway in Sézary syndrome. *Nature Communications.* 2015;6(1.):
54. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. *Nature Genetics.* 2015;47(9):1011–1019.
55. McGranahan N, Rosenthal R, Hiley CT, et al. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell.* 2017;171(6):1259–1271.e11.
56. Mardis E. Faculty of 1000 evaluation for Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *F1000 - Post-publication peer review of the biomedical literature.* 2015;
57. Zhang J, Fujimoto J, Zhang J, et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science.* 2014;346(6206):256–259.
58. Harbst K, Lauss M, Cirenajwis H, et al. Multiregion Whole-Exome Sequencing Uncovers the Genetic Evolution and Mutational Heterogeneity of Early-Stage Metastatic Melanoma. *Cancer Res.* 2016;76(16):4765–4774.
59. Gundem G, Van Loo P, Kremeyer B, et al. The evolutionary history of lethal metastatic prostate cancer. *Nature.* 2015;520(7547):353–357.
60. Buus TB, Willerslev-Olsen A, Fredholm S, et al. Single-cell heterogeneity in Sézary syndrome. *Blood Adv.* 2018;2(16):2115–2126.

61. Borchering N, Voigt AP, Liu V, et al. Single-Cell Profiling of Cutaneous T-Cell Lymphoma Reveals Underlying Heterogeneity Associated with Disease Progression. *Clin. Cancer Res.* 2019;25(10):2996–3005.
62. Litvinov IV, Tetzlaff MT, Thibault P, et al. Gene expression analysis in Cutaneous T-Cell Lymphomas (CTCL) highlights disease heterogeneity and potential diagnostic and prognostic indicators. *Oncoimmunology.* 2017;6(5):e1306618.
63. Gaydosik AM, Tabib T, Geskin LJ, et al. Single-Cell Lymphocyte Heterogeneity in Advanced Cutaneous T-cell Lymphoma Skin Tumors. *Clin. Cancer Res.* 2019;25(14):4443–4454.
64. Bradford PT, Devesa SS, Anderson WF, Toro JR. Cutaneous lymphoma incidence patterns in the United States: a population-based study of 3884 cases. *Blood.* 2009;113(21):5064 LP–5073.
65. Kirsch IR, Watanabe R, O'Malley JT, et al. TCR sequencing facilitates diagnosis and identifies mature T cells as the cell of origin in CTCL. *Sci. Transl. Med.* 2015;7(308):308ra158.
66. Masson A de, de Masson A, O'Malley JT, et al. High-throughput sequencing of the T cell receptor β gene identifies aggressive early-stage mycosis fungoides. *Science Translational Medicine.* 2018;10(440):eaar5894.

Chapter 2: Clonotypic heterogeneity in cutaneous T-cell lymphoma revealed by comprehensive whole exome sequencing*

2.1 Abstract

Mycosis fungoides (MF), the most common type of cutaneous T-cell lymphoma, is believed to represent a clonal expansion of a transformed, skin resident, memory T-cell. T-cell receptor (TCR) clonality (i.e. identical sequences of rearranged TCR α , β and γ), the key premise of this hypothesis, has been difficult to document conclusively because malignant cells are not readily distinguishable from the tumor infiltrating, reactive lymphocytes, which contribute to the TCR clonotypic repertoire of MF. Here we have successfully adopted the technique of targeted whole exome sequencing (WES) to identify the repertoire of rearranged TCR genes in tumor enriched samples from patients with MF. Though some of the investigated biopsies of MF had the expected monoclonal rearrangements of TCR γ of the frequency that are corresponding to those of tumor cells, majority of the samples presented multiple TCR- γ , - α and - β clonotypes by WES. Our findings are compatible with the model in which the initial malignant transformation in MF does not occur in mature, memory T-cells but rather at the level of T-lymphocyte progenitor before TCR β or TCR α rearrangements. We have also shown that WES can be combined with whole transcriptome sequencing (WTS) in the same sample which enables comprehensive characterization of the TCR repertoire in relation to tumor content. WES/WTS might be applicable to other types of T-cell lymphomas to determine clonal dominance and clonotypic heterogeneity in these malignancies.

*A version of this work has been published as A. Iyer, D. Hennessey, S. O’Keefe, J. Patterson, W. Wang, G. Ka-Shu Wong, R. Gniadecki, Clonotypic heterogeneity in cutaneous T-cell lymphoma (mycosis fungoides) revealed by comprehensive whole-exome sequencing. *Blood Adv.* 2019;3(7):1175–1184.

2.2 Background

Mycosis fungoides (MF) is the most prevalent form of cutaneous T-cell lymphoma (CTCL). In early stages it presents with scaly plaques on the skin which may progress into tumors and finally disseminate to lymph nodes and to other organs.¹⁻³ MF can be viewed as a model of low-grade T-cell lymphomas: it has a chronic, relapsing course, low-grade proliferation, chemotherapy resistance and 5-year mortality approaching 50%.^{1,4} MF expresses markers of memory T-cell and appears to exhibit T-cell receptor (TCR) monoclonality and is thus considered to be caused by malignant transformation of a mature T-cell residing in the skin.⁵

TCR gene sequences are excellent markers of T-cell lineage because TCR- δ , - γ , - β and - α loci become sequentially rearranged during intrathymic maturation of T-cell from diverse V, (D) and J gene segment pools, and the unique products of the rearrangements are retained (with the notable exception of TCR- δ) in all daughter cells.⁶ Complementarity- determining region 3 (CDR3) encoded by the V(D)J junction is especially useful for lineage tracing because its sequence heterogeneity is increased beyond the combinatorial V(D)J diversity by random insertions and deletions of nucleotides during segment recombination.⁷ Thus, identical TCR γ , - β and - α sequences of CDR3 in all lymphoma cells would be conclusive proof that malignant transformation took place in a mature T-cell which had completed TCR rearrangement.

However, true TCR monoclonality, as defined by a single T-cell clonotype, has not been demonstrated in CTCL. Usually, the dominant clone is accompanied by several other TCR clones thought to originate from reactive, tumor-infiltrating T-cells. Statistical methods have been used to formally determine clonality⁸ but these methods neither distinguish between

tumor clones and expanded reactive clones nor determine clonotypic heterogeneity of the tumor itself.

Determination of the clonotypic structure of CTCL is practically important, because clonality assessments are used for clinical diagnosis, prognosis and staging of CTCL.^{1,9} The most widely used method based on multiplexed PCR amplification of TCR γ and $-\beta$ and Genescan analysis¹⁰ is currently being replaced by methods based on high-throughput sequencing of PCR amplified CDR3 regions.^{9,11–13} They seem to have superior sensitivity and specificity in the detection of the T-cell clone but they cannot differentiate CDR3 sequences derived from tumor cells versus those derived from reactive T-cells and do not provide any measure of sample purity (the percentage of neoplastic cells). Moreover, the amplification step with multiplex PCR makes sequencing of the complex TCR α locus virtually impossible. Currently, sequencing of TCR α can be achieved by RNA-seq where primers binding to the invariable constant TCR segment are used but only the transcribed TCR alleles are detected and the information on other non-productive rearrangements in the genome is not captured. Unfortunately, RNA-seq results may be distorted by the presence of alternatively spliced mRNA and allele silencing, not uncommonly seen in cancer.¹¹

It has been reported that the CDR3 sequences of rearranged TCR β genes can be retrieved from the whole exome sequencing (WES).¹⁴ Based on this finding, we have developed a protocol in which samples are analysed by the probe capture WES. This allowed us to identify recombined TCR α , $-\beta$ and $-\gamma$ sequences from DNA in MF patients and compare their respective expression patterns. Since WES also allows to quantify the percentage of tumor

cells in the sample ¹⁵, we were able to reconstruct clonotypic composition of MF and provided evidence for TCR heterogeneity of this lymphoma.

2.3 Material and Methods

2.3.1 Sample collection and storage

Ethical approval was obtained from the Health Research Ethics Board of Alberta, Cancer Committee HREBA.CC-16-0820-REN1. After informed consent, 4mm punch skin biopsies were collected from patients and embedded in optimal cutting temperature (OCT) medium at -80°C. 10 ml of blood was collected and Ficoll was used to isolate peripheral blood mononuclear cells (PBMC) that were subsequently resuspended in 50% of Dulbecco's modified eagle medium (DMEM), 40% Fetal bovine serum (FBS) and 10% DMSO and frozen in liquid nitrogen until further use.

2.3.2 Cryosectioning and laser capture microdissection (LCM)

10 µm sections of the skin biopsies frozen in OCT were collected on 2 µm polyethylene naphthalate (PEN) membrane slides (cat# 11505158) (Leica Microsystems, Wetzlar, Germany). The slides were then stained using hematoxylin and eosin stains to identify the tumor cells. The microdissected tumor cell clusters were pooled together and collected in RLT buffer (cat# 79216) (Qiagen, Hilden, Germany) and used for simultaneous DNA/RNA isolation using AllPrep DNA/RNA micro kit (cat# 80284) (Qiagen, Hilden, Germany). Isolated DNA was preamplified using REPLI-g single cell kit (cat# 150343) (Qiagen, Hilden, Germany).

2.3.3 Sample preparation for whole exome sequencing (WES)

1 µg of DNA measured using QubitTM dsDNA HS assay kit (cat# Q32851) (Thermo Fisher, Massachusetts, United States) was sheared at a peak size of 200 bp using Covaris S2 focused-ultrasonicator (Covaris, Massachusetts, United States). Prior to end-repair sheared DNA for samples MF1, MF2, MF25, MF30, MF33, MF35, MF36, MF37, MF43, MF44 and MF45 were incubated with NEBNext FFPE DNA repair mix (cat#M6630S) (New England Biolabs, Massachusetts, United States) and later end-repaired, ligated with adaptors and indexed using NEBNext[®] UltraTM II DNA library prep kit for Illumina (cat# E7645S) (New England Biolabs, Massachusetts, United States). For DNA amplification 4-7 cycles were used rather than number of cycles recommended by NEB. Prepared libraries were hybridized with biotin labeled RNA baits (SSELXT Human All exon V6 +UTR) (Agilent Technologies, California, United State) at 65 °C for 2 hours. Few of the samples were also used for hybridization with customized probes designed to target V and J regions of TCR α , TCR β and TCR γ . These customized probes were combined with the current SSELXT Human All exon V6 +UTR kit (Custom + SSELXT Human All exon V6 +UTR) to improve the overall efficiency of the capture protocol in identifying TCR clonotypes. Hybridized DNA was pulled down using DynabeadsTM MyOneTM streptavidin T1 (cat# 65601) (Thermo Fisher, Massachusetts, United States). Captured DNA was re-amplified using KAPA library amplification kit with primers (cat# 07958978001) (Roche Diagnostics, Risch-Rotkreuz, Switzerland). The peak size of enriched DNA libraries verified using 2100 Bioanalyzer, (Agilent Technologies, California, United State) was average of 325bp. The DNA libraries were sequenced on an Illumina HiSeq 1500 sequencer using paired-end (PE) 150 kit (cat# PE-402-4002) (Hiseq PE rapid cluster kit V2) or NovaSeq 6000 S4 reagent kit 300 cycles (cat# 20012866).

2.3.4 Sample preparation for whole transcriptome sequencing (WTS)

10 ng of total RNA quantified using Qubit™ RNA HS assay kit (Q32852) (Thermo Fisher, Massachusetts, United States) was used for rRNA depletion (E6310) (New England Biolabs, Massachusetts, United States). rRNA depleted samples were used for cDNA synthesis and the library was built using NEBNext® Ultra™ II directional RNA library prep kit for Illumina (E7760) (New England Biolabs, Massachusetts, United States). The peak size of prepared cDNA libraries was verified using 2100 Bioanalyzer, (Agilent Technologies, California, United State). The cDNA libraries were later sequenced on an Illumina HiSeq 1500 sequencer using paired-end 150 kit (cat# PE-402-4002) (Hiseq PE rapid cluster kit V2).

2.3.5 Data analysis

The fastq files were analyzed using MiXCR to identify the TCR clonotypes.¹⁶ Short and long read alignments were included for whole-transcriptome sequencing (WTS); however, for WES data, partial reads were filtered out because they might be the captures of only V or J sequences. Reads were processed using the GATK4 generic data-preprocessing workflow¹⁷ and then analyzed with Titan¹⁵ to determine copy number aberration (CNA) and tumor purity using the hg38 human reference genome. The tcR package in R was used to calculate the inverse Simpson diversity index and identify the overlapping clones.¹⁸ VJ combination bias was analyzed using the VDJtools package in R.¹⁹

2.4 Results

2.4.1 Identification of T-cell clonotypes from whole exome sequencing (WES) and whole transcriptome sequencing (WTS)

The sequences of CDR3 regions and TCR clonotypes can be determined from WES and WTS. We performed laser capture microdissection (LCM) of the areas of atypical lymphocytic infiltrate in 33 biopsies of plaques (early lesions) and tumors (advanced lesions) of the 27 patients with MF (**Fig 2.1**) (Appendix **Table A1**). Due to lack of tumor specific markers, T-cells were identified only based on histology of the cells. We expected some contamination with reactive T-cells because histology cannot unequivocally define early stage lymphoma cells. Therefore, copy number aberration (CNA) analyzed from WES¹⁵, was used to identify the percentage of tumor cells in the LCM samples. Moreover, to compare the results of WES and the whole transcriptome sequencing (WTS) directly, we purified DNA and RNA simultaneously from the same isolated cell clusters for few of the tumor and plaque pairs (early and late lesion samples collected from the same patient). As shown in **Fig 2.2 A, B**, using the capture-based WES technique we successfully identified numerous CDR3 sequences corresponding to TCR α , TCR β and TCR γ clonotypes. With the sequencing depth of 87×10^6 reads, we were able to capture (median and range): 146 (37-471) TCR α , 40 (5-110) TCR β and 21.5 (1-98) TCR γ clonotypes. The relative excess in TCR α abundance is readily explainable by the fact that during T-cell development, the TCR β is under the strict allelic exclusion, but TCR α locus is usually rearranged on both chromosomes, sometimes in multiple rounds resulting in 2-4 TCR α rearrangements per single TCR β rearrangement.^{16, 17} This explanation is confirmed by WTS results documenting a comparable number of expressed TCR β clonotypes to the number of clonotypes identified at the DNA level (35.5 vs 40) and practically the same median number of transcribed TCR α clonotypes (n=50) in 9 of the MF samples with available

WTS data (**Fig 2.2 B**). There was no bias in V and J segment detection in the controls of peripheral blood samples (**Appendix Fig A1**) with the same WES protocol.

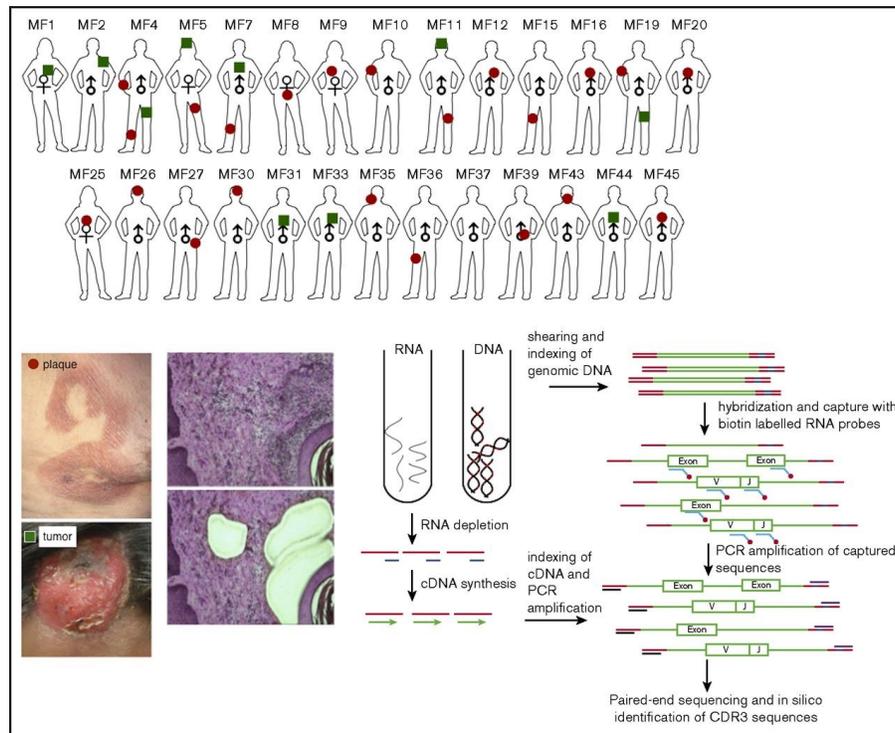


Figure 2.1: Schematic representation of sample collection, processing, and TCR sequencing.

Four-millimeter punch biopsies were collected from early lesions (plaques; red circles) or tumors (green squares) in 27 patients with MF. Biopsies were cryosectioned and laser microdissected to capture tumor cells that were pooled together. Original magnification 10x; hematoxylin and eosin staining. DNA and RNA were isolated simultaneously from the microdissected material and processed for WTS and WES. WTS data are available only for samples MF4_2T, MF4_3P, MF5_1T, MF5_2P, MF7_1T, MF7_2P, MF11T, MF11_1P, MF19_1T, and MF19_2P and a pool of normal CD4+ lymphocytes (data not shown). The gene sequence is indicated in green; the adapter sequence is indicated in red, and the index sequence is indicated in blue.

2.4.2 Efficiency of probe capture technique in identification of T-cell clonotypes

Previous protocols with probe-capture and high-throughput sequencing used TCR specific probes rather than the vast panel of probes for the entire exome.¹⁸ The drawback of that

approach is that fewer probes can paradoxically lead to decreased capture efficiency (Wong GK-S, unpublished). Since the exome capture probe set was not specifically designed to capture TCR genes, we asked whether the efficiency can be increased by adding probes targeting V and J segments of TCR α , β and γ . As shown in **Fig 2.2 C-E**, those additional probes increased the total number of identified clonotypes in 3 of the 4 samples, but the difference was not statistically significant. Therefore, for the subsequent experiments we used standard exome capture probes. We have also tested the sequencing depth on clonotype detection efficiency by sequencing two total blood samples with 400 million reads each. We observed that at a depth of 348 million read per sample (approximately 800x sequencing coverage), the capture experiments with deep sequencing did not reach saturation in identifying TCR clonotypes. The efficiency with increased sequencing depth still remained highest for TCR α and lowest for TCR γ (**Fig 2 F-H**).

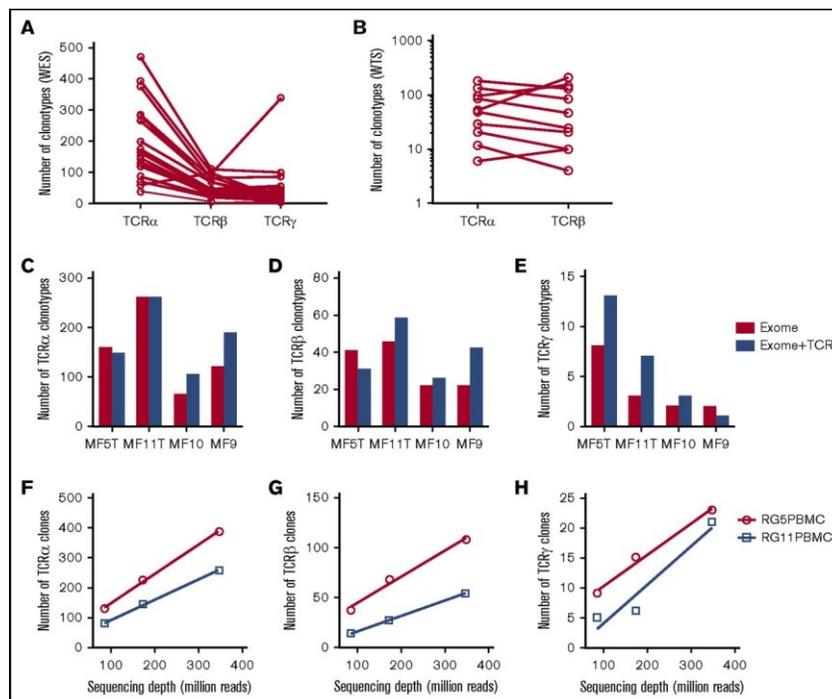


Figure 2.2: Efficiency of WES probe capture and WTS protocols in the detection of CDR3 clonotypes in MF biopsies.

All of the samples were sequenced using whole-exome probe capture (**A**) and WTS (**B**), and the number of clonotypes for TCR α , TCR β , and TCR γ was determined for each sample, as indicated. The lines connect the results for the same sample. (**C-E**) The effect of TCR-specific probes. The capture was performed in 4 samples with whole-exome probes as in panel A (Exome) or with whole-exome probes combined with specific TCR capture probes (Exome+TCR) and sequenced, and the number of unique clonotypes for TCR α (**C**), TCR β (**D**), and TCR γ (**E**) was determined, as in panel A. The addition of probes slightly increased the number of TCR γ clonotypes ($P = .024$, paired Student t test) but not the number of TCR α or TCR β clonotypes. (**F-H**) The effect of sequencing depth on clonotype detection for TCR α (**F**), TCR β (**G**), and TCR γ (**H**). Two samples of whole peripheral blood mononuclear cells were sequenced with WES at a maximum of 400 million reads, as in panel A. The samples do not reach saturation up to 348 million reads ($\sim 800\times$ sequencing depth).

2.4.3 Analysis of malignant TCR clonotypes in MF

MF is thought to develop from memory T-cells and therefore it should have the same TCR γ , - β and - α clonotypes. The concept of monoclonality of T-cell lymphoma has been well documented using multiplex/heteroduplex PCR amplification and detection by capillary electrophoresis or high throughput sequencing^{10,12} and is used as a diagnostic test in CTCL. We were therefore interested whether our WES-based method of clonotype detection could identify those TCR clones in MF samples. The biopsies always contain variable, usually unknown, amounts of reactive T-cells contributing to the repertoire of TCR clonotypes. Perhaps therefore, previous studies claimed monoclonality in samples showing dominant clonotype of frequency as low as 15%, the rest of the clonotypes (up to 85% of sample composition) being considered to represent reactive T-cells.^{9,12} As shown in Fig 3, if the 15% clonotype frequency threshold is applied, only 9 of 33 MF for TCR γ , 15 of 33 MF for TCR β and 5 of 33 MF for TCR α can be classified as monoclonal on the basis of WES.

Information from WES was used to identify CNA in cancer genome and hence calculation of the enrichment tumor cells in the LCM sample. Even in the microdissected samples the proportion of malignant cells varied between 21.1% to 98.6% (median 71.28%) and there

were no differences between the plaques and the tumors. Contrary to expectation, neither the frequency of the most abundant (dominant) clone nor diversity index (inverse Simpson index) were correlated with the proportion of tumor cells in the sample (Appendix **Fig A2**). More surprising was the finding that a single TCR β clonotype cannot account for all malignant cells in the sample (**Fig 2.4**). Even in samples with the ratio of the sum of two dominant (biallelic) TCR γ clonotypes to the proportion of tumor cells ≈ 1 (MF4_2T, MF4_3P, MF5_1T, MF5_2P, MF7_1T, MF8P, MF9P, MF11T, MF11_1P), representing samples with perfect TCR γ monoclonality, the dominant TCR β clonotype could account for a median of only 15% of tumor cells. As shown in **Fig 2.3B**, WES revealed presence of additional one to three TCR β clonotypes which together had a comparable frequency to the dominant clonotype. Intriguingly, WTS for these samples revealed single dominant TCR β and TCR α in MF4_3T and MF11T, oligoclonality in MF7_1T, MF7_2P and polyclonality for MF5_1T, MF11_1P, MF19_1T and MF19_2P (**Fig 2.3E, F**). This result illustrates that a malignant T-cell clone can rearrange multiple TCR γ , TCR β and TCR α in some instances express more than a single TCR α and $-\beta$ mRNA.

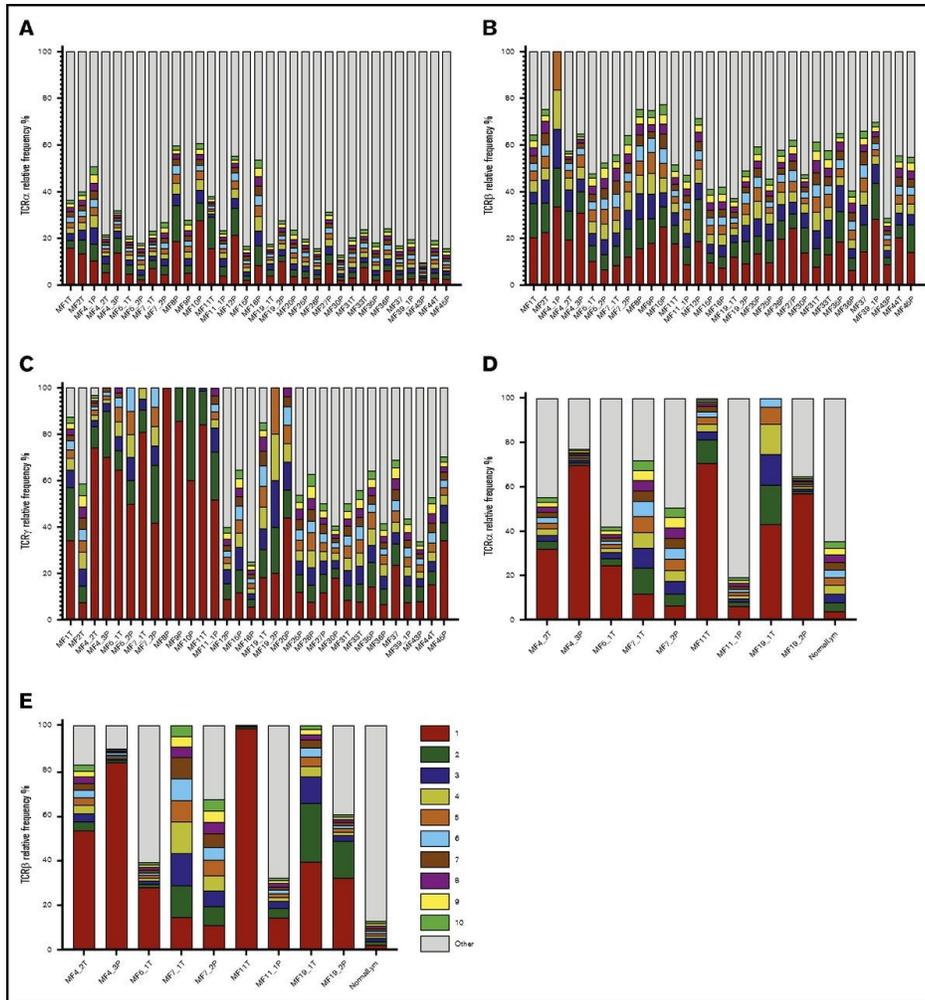


Figure 2.3: Relative frequency of T-cell clonotypes.

TCRα (A), TCRβ (B), and TCRγ (C) repertoire sequences identified from WES of MF samples. Sample ID relates to patient number, as in Figure 1, with the suffix P (plaque) or T (tumor). TCRα (D) and TCRβ (E) repertoires identified by WTS of MF samples. Each bar represents an individual CDR3 amino acid clonotype, with red and green indicating the first-ranked and tenth-ranked clonotype, respectively, in decreasing order of relative frequency. Gray bars represent the rest of the identified clonotypes in the samples. NormalLym, pooled CD4+ normal lymphocytes from 4 healthy donors.

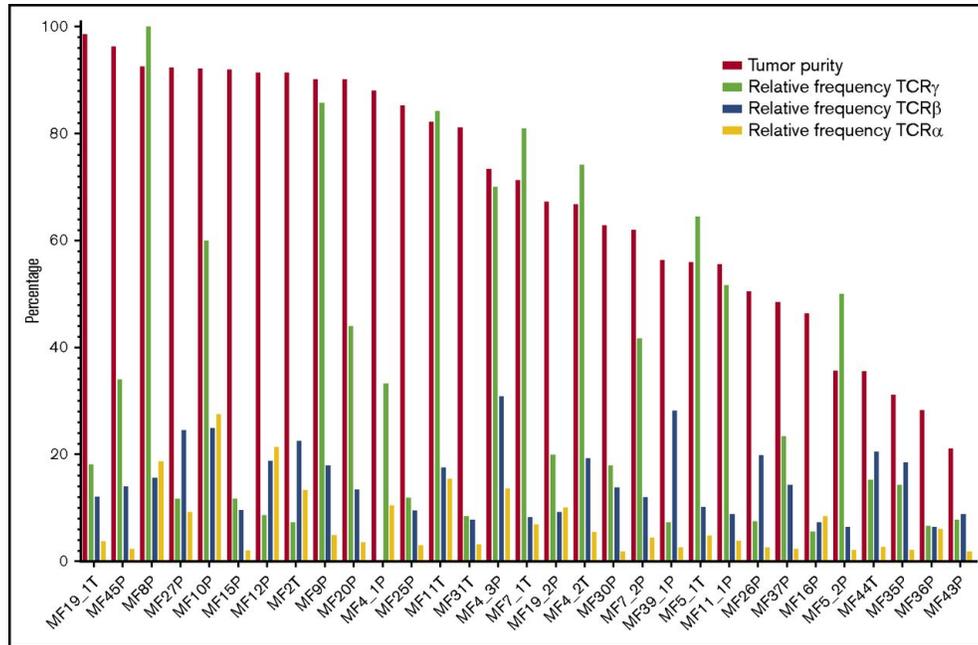


Figure 2.4: Clonotypic diversity of MF.

Contribution of the dominant clonotypes of TCR α , TCR β , and TCR γ relative to the tumor DNA enrichment of the sample. Note that for samples MF4_2T, MF4_3P, MF5_1T, MF5_2P, MF7_1T, MF7_2P, MF8P, MF9P, MF11T, and MF11_1P, the proportion of the dominant TCR γ clonotype is approximately equal to the proportion of tumor DNA in the samples, indicating that all tumor cells share the same TCR γ clonotype. However, in the same samples, the relative frequency of the dominant TCR α and TCR β clonotypes is only 15% (range, 6.52-30.89), indicating that other clonotypes are found in tumor-derived DNA.

2.4.4 Identification of shared TCR clonotypes

The monoclonal mature T-cell theory dictates that the tumor is an expansion of the clone found in early stage lesions such as patches and plaques.⁵ Therefore, high degree of overlap between clones of the tumor and the plaque samples collected from a patient at a single time point is expected. Therefore, we were interested whether clonotypic composition is the same in early (plaque) and advanced (tumor) lymphoma lesions. Due to the vast number of clonotypes and reactive T-cell contamination we focused on sharing of top 10 dominant clonotypes, most likely to represent the malignant clonotypes. For the 5 pairs in our dataset, 4 pairs (MF4_2T, MF4_3P, MF5_1T, MF5_2P, MF11T, MF11_1P, MF19_1T and MF19_2P)

shared no more than 1-3 clonotypes independently for TCR α , - β and - γ and patient MF7 (samples MF7_1T and MF7_2P) shared no clonotypes. In retrospect, given the vastness of the CDR3 repertoire it could be expected that individual clonotypes are not shared in samples from different patients. However, interindividual clonotype sharing was relatively common with the highest number of 4 of the top 10 dominant clonotypes shared between MF4_2T and MF43T for TCR α , MF30P and MF37P for TCR β and MF31T and MF44T for TCR γ (**Fig 2.5A-C**). For all clonotypes detected in a sample, the number of shared clonotypes was even higher, reaching 45 shared TCR α clonotypes, 10 TCR β and 25 TCR γ clonotypes. The V α and V β segment usage was characterized by high representation of pseudogenes (TRAV11, TRAV28, TRAV31, TRBV12-1, TRBV22-1) but otherwise did not reveal any clues as to the functional role of those clonotypes.

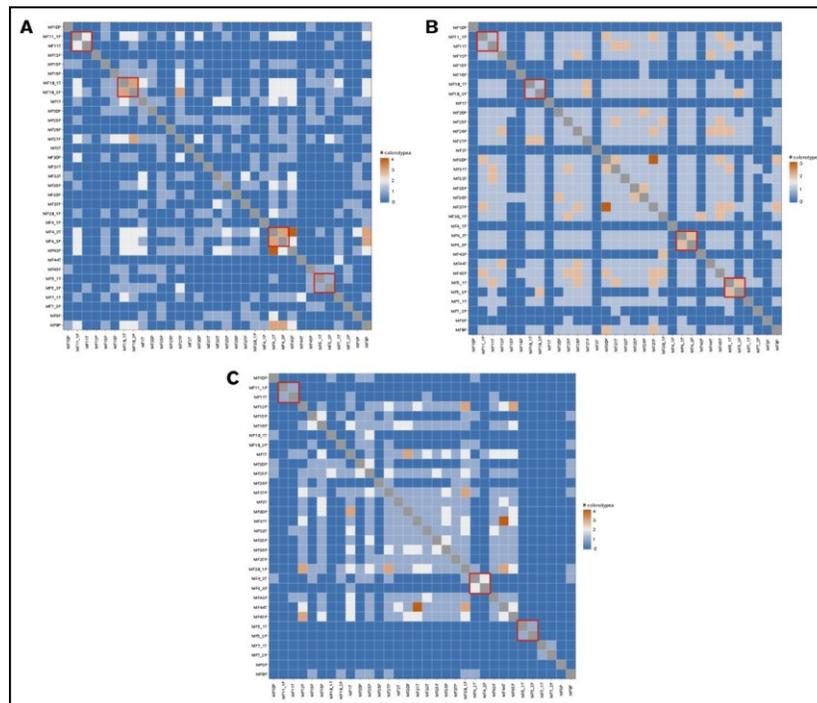


Figure 2.5: Shared T-cell clonotypes.

The 10 most frequent CDR3 sequences identified using WES were tested for overlap. The red boxes denote the TCR α (A), TCR β (B), and TCR γ (C) clonotypes shared in tumor and plaque lesion pairs collected from each individual patient.

2.5 Discussion

In this report we demonstrate that TCR repertoire in MF can be assessed by probe-capture based WES and identifies simultaneously TCR γ , - β and - α rearrangement. The method gives an advantage of identifying TCR α locus rearrangements that do not amplify reliably with multiplex PCR due to large number of V and J genes. To date, all data on TCR α were gathered with RNA sequencing^{19,20} and very little is known about the diversity of TCR α at the DNA level. Another advantage of our approach is using the exome data to estimate the percentage of malignant cells in the sample, eliminating the need of arbitrary thresholds for reactive T-cell contamination in the samples.

The drawback of our method is its lower robustness than the PCR-based methods in capturing the whole TCR repertoire in the sample. WES/WTS yielded hundreds rather than thousands of TCR α and TCR β clonotypes which although sufficient to analyse TCR rearrangement in tumor cells that does not allow for comprehensive estimation of the entire T-cell diversity. The number of detected clonotypes was linearly dependent on sequencing depth and does not reach saturation at the depth of 348 million where a maximum of 390 TCR α and 109 TCR β clonotypes could be detected in whole blood samples. It is possible that further improvements in capture probe design the robustness of the technique could be increased to study clonotypes of low frequency as well.

Analysis of TCR repertoire in MF by WES led to unexpected conclusions regarding the nature of clonal expansion of malignant cells. By comparing the proportion of tumor-derived DNA in

the sample with the relative frequencies of TCR γ , - β and - α clonotypes, we found evidence for existence of multiple, rather than single, malignant T-cell clonotypes. Especially informative were the cases where the proportion of monoclonal TCR γ rearrangement matched the proportion of tumor-derived DNA, indicating that the sample was composed of a population of malignant cells sharing identical TCR γ clonotype (e.g. cases MF4_2T, MF4_3P, MF5_1T, MF5_2P, MF7_1T, MF8P, MF9P, MF11T, MF11_1P **Fig 2.3, 2.4**). Instead of expected TCR β monoclonality, we detected 2-7 TCR β clonotypes and multiple TCR α clonotypes. This indicates that at least in some cases of MF, the initial transformation does not happen at the level of skin-resident memory T-cell, but possibly much earlier, during lymphocyte development after completion of TCR γ rearrangement, but before initiation of TCR β and - α recombination. Thus, all malignant cells inherit the identical TCR γ CDR3 sequences, but not TCR β or TCR α which would be different in the subclones descending from the same precursor. Other groups that performed TCR sequencing in CTCL also found evidence of oligoclonality.²¹ Recently, Ruggiero et al.¹⁹ using ligation-anchored PCR for mRNA amplification and sequencing of TCR α and TCR β in Sézary syndrome found oligoclonal, rather than monoclonal pattern in 4/10 patients and polyclonal TCR repertoire was reported in subgroups of patients with PTCL-NOS or AITL. Supportive evidence comes also from the studies showing multiple TCR β transcripts in CTCL with the copy number aberration of chromosome 7 containing TCR β .²² An explanation that malignant T-cells are able to recombine TCR in the periphery is unlikely, because the essential recombination activating genes RAG-1 and RAG-2 are not active in mature T-cells or in CTCL.²³ Moreover, in our WTS dataset there is no evidence of RAG1/2 expression (data not shown). Previous findings that chromosomal breaking points in CTCL contain RAG heptamer sequences, reinforce our

conclusion that initial stages of malignant transformation happens early during lymphocyte development when RAG enzymes are active.²⁴

We have also considered the possibility that the observed clonotypic heterogeneity in MF could be caused by malignant transformation of multiple cells in an inflammatory infiltrate.²⁵

Previous studies have indicated *Staphylococcus aureus* in skin microbiota provides an antigenic drive for MF. This hypothesis was supported by findings of a higher than expected usage of V β segments involved in recognition of staphylococcal superantigens (e.g. TRBV20 or TRBV5.1).^{12,18,19, 26, 27} We could not confirm those observations; on the contrary, we found that MF clonotypes including those shared between patients, contain V α and V β segments that are found at a very low frequency in peripheral blood or in the inflamed skin (e.g. pseudogenes TRAV11, TRAV28, TRAV31, TRBV12-1, TRBV22-1).^{28,29} We hypothesize that the putative increased frequency of pathogen recognizing V β usage identified in previous studies may be due to presence of reactive T-cell in the sample, which was minimized in our material by using microdissected samples enriched in neoplastic cells.

In conclusion, we have demonstrated that probe-capture based WES is a useful and straightforward approach to identify clonotypic composition in MF. Our data show that CTCL is clonotypically heterogeneous which strongly suggests that the initial malignant transformation may take place at a stage of T-cell precursor rather than the mature T-cell, as currently believed. The clinical implications of clonotypic heterogeneity for diagnosis and prognosis remain to be further investigated. It is however conceivable that clonotypic heterogeneity is a feature of a more general phenomenon of tumor heterogeneity which is

known to have profound impact on tumor prognosis and response to therapy.^{20,21} It remains to be seen whether clonotypic heterogeneity is correlated with clinical outcomes and whether this phenomenon is present in other T-cell lymphomas.

2.6 References

1. Olsen E, Vonderheid E, Pimpinelli N, et al. Revisions to the staging and classification of mycosis fungoides and Sézary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the cutaneous lymphoma task force of the European Organization of Research and Treatment of Ca. *Blood*. 2007;110(6):1713 LP-1722.
2. Kim YH, Willemze R, Pimpinelli N, et al. TNM classification system for primary cutaneous lymphomas other than mycosis fungoides and Sezary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the Cutaneous Lymphoma Task Force of the European Organization of Resear. *Blood*. 2007;110(2):479–484.
3. Willemze R, Jaffe ES, Burg G, et al. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768 LP-3785.
4. Scarisbrick JJ, Prince HM, Vermeer MH, et al. Cutaneous Lymphoma International Consortium Study of Outcome in Advanced Stages of Mycosis Fungoides and Sezary Syndrome: Effect of Specific Prognostic Markers on Survival and Development of a Prognostic Model. *J. Clin. Oncol*. 2015;33(32):3766–3773.
5. Kim EJ, Hess S, Richardson SK, et al. Immunopathogenesis and therapy of cutaneous T cell lymphoma. *J. Clin. Invest*. 2005;115(4):798–812.
6. Davis MM, Bjorkman PJ. T-cell antigen receptor genes and T-cell recognition. *Nature*. 1988;334(6181):395–402.
7. Manfras BJ, Terjung D, Boehm BO. Non-productive human TCR beta chain genes represent V-D-J diversity before selection upon function: insight into biased usage of

- TCRBD and TCRBJ genes and diversity of CDR3 region length. *Hum. Immunol.* 1999;60(11):1090–1100.
8. Kaplinsky J, Arnaout R. Robust estimates of overall immune-repertoire diversity from high-throughput measurements on samples. *Nat. Commun.* 2016;7:11881.
 9. de Masson A, O'Malley JT, Elco CP, et al. High-throughput sequencing of the T cell receptor beta gene identifies aggressive early-stage mycosis fungoides. *Sci. Transl. Med.* 2018;10(440).
 10. Assaf C, Hummel M, Dippel E, et al. High detection rate of T-cell receptor beta chain rearrangements in T-cell lymphoproliferations by family specific polymerase chain reaction in combination with the GeneScan technique and DNA sequencing. *Blood.* 2000;96(2):640–646.
 11. Hou X-L, Wang L, Ding Y-L, Xie Q, Diao H-Y. Current status and recent advances of next generation sequencing techniques in immunological repertoire. *Genes Immun.* 2016;17(3):153–164.
 12. Kirsch IR, Watanabe R, O'Malley JT, et al. TCR sequencing facilitates diagnosis and identifies mature T cells as the cell of origin in CTCL. *Sci. Transl. Med.* 2015;7(308):308ra158.
 13. Sufficool KE, Lockwood CM, Abel HJ, et al. T-cell clonality assessment by next-generation sequencing improves detection sensitivity in mycosis fungoides. *J. Am. Acad. Dermatol.* 2015;73(2):228–36.e2.
 14. Levy E, Marty R, Gárate Calderón V, et al. Immune DNA signature of T-cell infiltration in breast tumor exomes. *Sci. Rep.* 2016;6:30064.

15. Ha G, Roth A, Khattra J, et al. TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. *Genome Res.* 2014;24(11):1881–1893.
16. Bolotin DA, Poslavsky S, Mitrophanov I, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods.* 2015;12:380.
17. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297–1303.
18. Nazarov VI, Pogorelyy M V, Komech EA, et al. tcR: an R package for T cell receptor repertoire advanced data analysis. *BMC Bioinformatics.* 2015;16(1):175.
19. Shugay M, Bagaev D V., Turchaninova MA, et al. VDJtools: Unifying Post-analysis of T Cell Receptor Repertoires. *PLOS Comput. Biol.* 2015:2015;11(11):e1004503
20. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 2011;43(5):491–498.
21. Hamrouni A, Aublin A, Guillaume P, Maryanski JL. T Cell Receptor Gene Rearrangement Lineage Analysis Reveals Clues for the Origin of Highly Restricted Antigen-specific Repertoires. *J. Exp. Med.* 2003;197(5):601 LP-614.
22. Linnemann C, Heemskerk B, Kvistborg P, et al. High-throughput identification of antigen-specific TCRs by TCR gene capture. *Nat. Med.* 2013;19(11):1534–1541.
23. Ruggiero E, Nicolay JP, Fronza R, et al. High-resolution analysis of the human T-cell receptor repertoire. *Nat. Commun.* 2015;6:8081.

24. Gong Q, Wang C, Zhang W, et al. Assessment of T-cell receptor repertoire and clonal expansion in peripheral T-cell lymphoma using RNA-seq data. *Sci. Rep.* 2017;7(1):11301.
25. Linnemann T, Gellrich S, Lukowsky A, et al. Polyclonal expansion of T cells with the TCR V beta type of the tumour cell in lesions of cutaneous T-cell lymphoma: evidence for possible superantigen involvement. *Br. J. Dermatol.* 2004;150(5):1013–1017.
26. Bagot M, Echchakir H, Mami-Chouaib F, et al. Isolation of Tumor-Specific Cytotoxic CD4⁺ and CD4⁺CD8dim⁺ T-Cell Clones Infiltrating a Cutaneous T-Cell Lymphoma. *Blood.* 1998;91(11):4331 LP-4341.
27. Dobbeling U, Dummer R, Schmid MH, Burg G. Lack of expression of the recombination activating genes RAG-1 and RAG-2 in cutaneous T-cell lymphoma: pathogenic implications. *Clin. Exp. Dermatol.* 2006;22(5):230–233.
28. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. *Nat. Genet.* 2015;47(9):1011–1019.
29. Burg G, Kempf W, Haeffner A, et al. From inflammation to neoplasia: new concepts in the pathogenesis of cutaneous lymphomas. *Recent results cancer Res. Fortschritte der Krebsforsch. Prog. dans les Rech. sur le cancer.* 2002;160:271–280.
30. Jackow CM, Cather JC, Hearne V, et al. Association of erythrodermic cutaneous T-cell lymphoma, superantigen-positive *Staphylococcus aureus*, and oligoclonal T-cell receptor V beta gene expansion. *Blood.* 1997;89(1):32–40.
31. Keith R. Hudson, Helen Robinson, and John D. Two adjacent residues in staphylococcal enterotoxins A and E determine T cell receptor V beta specificity. *J. Exp. Med.* 1993;177(1):175–184.

32. Dean J, Emerson RO, Vignali M, et al. Annotation of pseudogenic gene segments by massively parallel sequencing of rearranged lymphocyte receptor loci. *Genome Med.* 2015;7(1):123.
33. Brunner PM, Emerson RO, Tipton C, et al. Nonlesional atopic dermatitis skin shares similar T-cell clones with lesional tissues. *Allergy.* 2017;72(12):2017–2025.

Chapter 3: Skin colonization by circulating neoplastic clones in cutaneous T-cell lymphoma*

3.1 Abstract

Mycosis fungoides (MF) is a mature T-cell lymphoma currently thought to develop primarily in the skin by a clonal expansion of a transformed, resident memory T-cell. However, this concept does not explain the key characteristics of MF such as the debut with multiple, widespread skin lesions or inability of skin directed therapies to provide cure. The testable inference of the mature T-cell theory is the clonality of MF with respect to all rearranged T-cell receptor (TCR) genes. Here we have used whole exome sequencing approach to detect and quantify TCR α , - β and - γ clonotypes in tumor cell clusters microdissected from MF lesions. This method allows us to calculate the tumor cell fraction of the sample and therefore an unequivocal identification of the TCR clonotypes as neoplastic. Analysis of TCR sequences from 29 patients with MF stage I-IV proved existence of multiple T-cell clones within the tumor cell fraction, with a considerable variation between patients and between lesions from the same patient (median 11 clones, range 2-80 clones/sample). We have also detected multiple neoplastic clones in the peripheral blood in all examined patients. Based on these findings we propose that circulating neoplastic T-cell clones continuously replenish the lesions of MF thus increasing their heterogeneity by a mechanism analogous to the consecutive tumor seeding. We hypothesize that circulating neoplastic clones might be a promising target for therapy and could be exploited as a potential biomarker in MF.

* A version of this work has been published as A. Iyer, D. Hennessey, S. O'Keefe, J. Patterson, W. Wang, G. Ka-Shu Wong, R. Gniadecki, Skin colonization by circulating

neoplastic clones in cutaneous T-cell lymphoma, *Blood* (2019): blood.2019002516. Web. 13 Sept2019.

3.2 Background

Cutaneous T-cell lymphomas (CTCL) are mature T-cell neoplasms, among which mycosis fungoides (MF) is the most common disease entity.¹ MF presents initially as scaly, erythematous patches and plaques on skin, which may progress to tumors and disseminate to lymph nodes and other organs, such as the central nervous system.²⁻⁴ The pathogenesis of MF has been studied for decades as a model disease reflecting key characteristics of low-grade lymphomas such as progressive course and lack of curative treatments.

MF is believed to originate from the mature, memory, tissue-resident T-cells expressing skin homing markers CLA and CCR4.^{5,6} This straightforward hypothesis explains the affinity of MF to the skin and its low capacity to disseminate to extracutaneous sites. However, some clinical and molecular features of MF are incompatible with the model of the skin-resident memory T-cell as origin of MF. It is unexplainable why the disease usually starts multifocally in different areas of the skin rather than in a single site representing the location of the founding, transformed T-cell. Second, even profound depletion of lymphocytes in the skin (e.g. by electron beam radiation therapy or psoralen ultraviolet A therapy [PUVA]) almost never results in a cure but only in a short-term responses.⁷⁻¹⁰ Third, cells sharing molecular characteristics of malignant T-cells in MF have been found in the bone marrow of the patients years before the emergence of skin lesions of the disease¹¹ and CTCL can be transmitted via bone marrow transplant from asymptomatic donors.^{12,13} Fourth, MF may share the common precursor with other lymphomas (e.g. Hodgkin lymphoma) that do not originate in the skin but primarily occupy extracutaneous sites, such as lymph nodes.¹⁴

These observations are more compatible with a scenario in which CTCL originate by hematogenous spread of precursor neoplastic cells to the skin niche.¹⁵ However, this concept was met with skepticism and resistance because analysis of the clonality of the T-cell receptor (TCR) seemed to strongly indicate that this disease is monoclonal and originates in the skin.^{6,16}

TCR clonality assays have been the most powerful technique to dissect the pathogenesis of CTCL and other T-cell lymphomas.^{17,18} During T-cell development, the *V*, (*D*) and *J* gene segments of *TCRG*, *TCRB* and *TCRA* undergo sequential rearrangements producing unique CDR3 sequences which are retained in the mature T-cells. The diversity at CDR3 is further increased by insertions or deletions at *V(D)J* junctions and therefore those sequences constitute a unique signature of a given T-cell clone.¹⁹ Most research focused on *TCRG* because of its relatively small size and limited diversity and only recently on *TCRB* which is more diverse and has a unique property of allelic exclusion which simplifies data analysis.

We have recently shown that MF cells sampled from a plaque or a tumor may share the same TCR γ clonotype but exhibit different TCR β and TCR α clonotypes.^{20,21} Since TCR γ loci (*TCRG*) rearrange before TCR β (*TCRB*) and TCR α (*TCRA*) and the unique *TCRG* CDR3 sequences are inherited by the T-cells derived from those early clones, these findings are incompatible with the current model of mature T-cell as precursor of MF where all malignant cells would share an identical clonotype for all rearranged TCR genes (*TCRG*, *TCRB* and *TCRA*).

Skin-resident T-cells do not recirculate but remain in the tissue where they can survive and proliferate without migration to the lymph nodes.^{5,22,23} Although atypical malignant T-cells are conspicuously absent in the blood in MF patients in the early stages and clonality assays are usually negative, it may be argued that standard TCR γ detection methods are rather insensitive for detecting rare clones in the background of highly diverse normal T-cells.²⁴ Indeed, using more careful experimental approaches such as tumor fraction enrichment by laser capture microdissection and analysis of purified lymphocytes or mononuclear cells from the blood, some authors were able to find clonal, circulating cells even in early stages of disease development.^{20,25–28} Unfortunately, due to the difficulties in PCR amplification of *TCRB* and *TCRA* from genomic DNA these findings rely heavily on the analyses of TCR γ , which may give false positive results because of the low diversity of *TCRG*. Likewise, relying on RT-PCR-based methods for TCR β may miss malignant clones with unproductive TCR rearrangements.

Here, we have applied the technique of TCR detection by whole exome sequencing²¹ to revisit the hypothesis of circulating neoplastic cells in MF. By comparing TCR clonotypes in the skin and blood in patients with MF we reveal a complex pattern of recirculation of tumor subclones. We propose that substantial clonotypic heterogeneity of skin lesions in MF is caused by the mechanism of consecutive seeding of the skin niche by multiple subclones of neoplastic cells.

3.3 Material and Methods

3.3.1 Patients, sample collection and storage

We included 29 patients with diagnosis of mycosis fungoides established by a certified dermatopathologist or a hematopathologist. Patient characteristics are summarized in Appendix **Table B1**. None of the patients received systemic therapy, radiotherapy or phototherapy at the time of tissue sampling. Most patients used medium and high potency topical steroids, but the biopsies were obtained from the lesions that have not been treated with steroids for at least 3 days. Ethical approval was obtained from the Health Research Ethics Board of Alberta, Cancer Committee HREBA.CC-16-0820-REN1. After informed consent, 4mm punch skin biopsies were collected from patients and embedded in optimal cutting temperature (OCT) medium stored at -80°C. 10 ml of blood was collected into EDTA tube, peripheral blood mononuclear cells (PBMC) were isolated by Ficoll centrifugation which were resuspended in 50% of Dulbecco's modified eagle medium (DMEM) (cat# 11965-084) (Thermo Fisher Scientific, Massachusetts, United States), 40% Fetal bovine serum (FBS) (cat# 16000044) (Thermo Fisher Scientific) and 10% Dimethyl Sulfoxide (DMSO) (cat# 20688) (Thermo Fisher Scientific) and frozen in liquid nitrogen until further use. Before DNA isolation the PBMC cells were thawed, resuspended in Roswell Park Memorial Institute (RPMI) 1640 medium with 10% FBS. DNA and RNA were isolated with Trizol reagent (cat# 15596026) (Invitrogen, Carlsbad, California, United States).

3.3.2 Cryosectioning, laser capture microdissection (LCM) and sample preparation for whole exome sequencing (WES)

Skin biopsies were cryosectioned and prepared for whole exome sequencing according to the previously reported protocol.²¹ NEBNext® Ultra™ II DNA library prep kit for illumina (cat# E7645S) (New England Biolabs, Massachusetts, United States) was used for preparing the

samples for sequencing and SSELXT Human All exon V6 +UTR probes (Agilent Technologies, California, United State) were used for the exome capture. The DNA libraries were sequenced on an Illumina HiSeq 1500 sequencer using paired-end (PE) 150 kit (cat# PE-402-4002) (Hiseq PE rapid cluster kit V2) or NovaSeq 6000 S4 reagent kit 300 cycles (cat# 20012866).

3.3.3 Data analysis

The fastq files were analyzed using MiXCR²⁹ (version 2.10.0) to identify the TCR clonotypes. For WES data, partial reads were filtered out as these might be the captures of only V or J sequences. The reads were processed using the GATK4³⁰(version 4.0.10) generic data-preprocessing workflow, then analyzed with Titan³¹(version 1.20.1) to determine copy number aberration and tumor cell fraction (TCF) using the hg38 Human reference genome. The tcR³² package for R was used to calculate the overlapping clones.

The number of neoplastic TCR β clonotypes (n_β) is calculated to satisfy the following formula:

$\sum_{i=1}^{n_\beta} TCRB_i \approx TCF$, where $TCRB_i$ is the percentage of the TCR β clonotype of i -rank (the rank $i=1$ being the most abundant, dominant clonotype) and TCF is the tumor cell fraction in the sample calculated from WES. We assumed that the proportion of malignant T-cells cells with >1 rearranged TCRB is negligible (allelic exclusion) and therefore the number of neoplastic TCR β clonotypes n_β is equal to the number of malignant T-cell clones. Although we have also computed the number of neoplastic clonotypes for TCR α (n_α) and TCR γ (n_γ), those values cannot directly be used for estimating the number of malignant clones because $TCRG$ and $TCRA$ often, but not always, rearrange on both chromosomes and $TCRA$ may rearrange producing >2 clonotypes for each TCR β clonotype.

3.4 Results

3.3.1 *TCRB* sequencing of MF shows lesional and topological heterogeneity of malignant clonotypes

We have previously provided evidence that whole exome sequencing (WES) can be used for identifying neoplastic TCR clonotypes, i.e. the TCR CDR3 sequences specific to malignant T-cells in MF.²¹ Our method relies on the sequencing of CDR3 regions and quantification of the fraction of TCR α , - β and - γ clonotypes corresponding to the tumor cell percentage, thus filtering out the TCR sequences from the reactive, tumor-infiltrating T-cells. Using the same approach here, we have identified neoplastic clonotypes in 29 patients with MF using microdissected samples of 46 biopsies from cutaneous lesions (**Fig 3.1**, Appendix **Table B1**). To quantify clonogenic heterogeneity, we focused on *TCRB* which is sufficiently diverse to avoid the risk identical rearrangements of unrelated T-cell clones and which is rearranged on a single chromosome in >98% of all T-cells (allelic exclusion) thus unequivocally defining a T-cell clone.^{33–36} In a purely monoclonal disease, one can expect that the frequency of the single, dominant TCR β clonotype matches the tumor cell fraction of the sample. However, in our 46 skin biopsies the tumor cell fraction significantly exceeded the frequency of the most abundant TCR β clonotype which unequivocally proved clonotypic heterogeneity of MF. We identified a range of 2-80 TCR β clonotypes per sample, which corresponded to the tumor cell fraction of the sample (i.e. neoplastic clonotypes) (**Fig 3.2A**, Appendix **Fig B1A**). On average, the most frequent (dominant) TCR β clonotype comprised only 19.32% of the tumor fraction, which was similar to the values obtained in other studies using the PCR-based approach and NGS.^{37,38} The number of neoplastic TCR β clonotypes correlated with the tumor cell fraction

but not with the stage of the lesion (T1 plaque vs T2 tumor) which further supported the notion that those clonotypes represented true tumor clones and were not derived from infiltrating, reactive T-cells (**Fig 3.2B**).

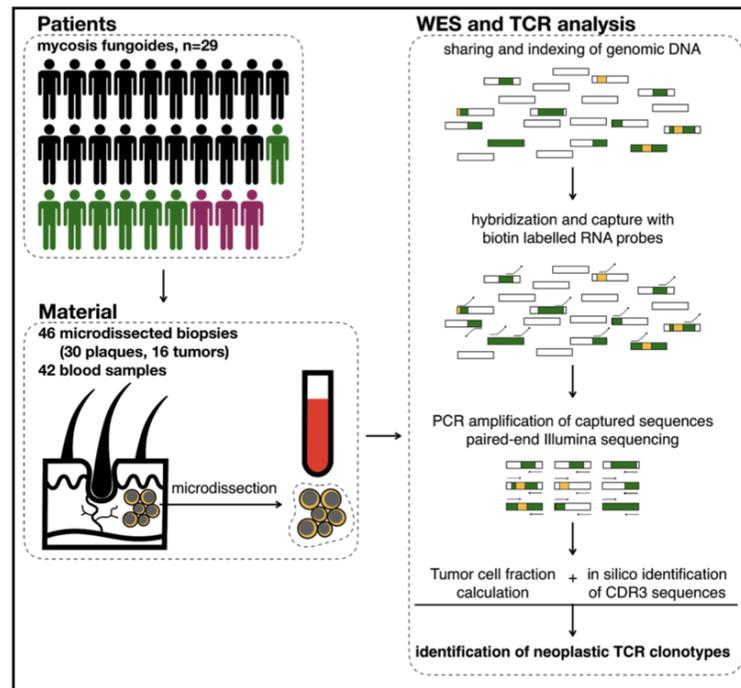


Figure 3.1: Schematic representation of sample collection and processing.

4mm single or multiple punch biopsies from and 10 ml of total blood was collected from 29 patients. In 7 patients we collected more than one biopsy (green silhouettes) whereas three patients were followed longitudinally with several biopsies and/or blood samples. The skin biopsies were cryosectioned and used for laser microdissection of clusters of tumor cells, which along with blood PBMC were processed for WES. Sequenced data was analyzed for identifying rearranged CDR3 sequences of *TCRA*, *TCRB* and *TCRG* and to determine tumor cell fraction. Rectangles represent DNA fragments, green areas are exons, yellow areas are rearranged TCR genes.

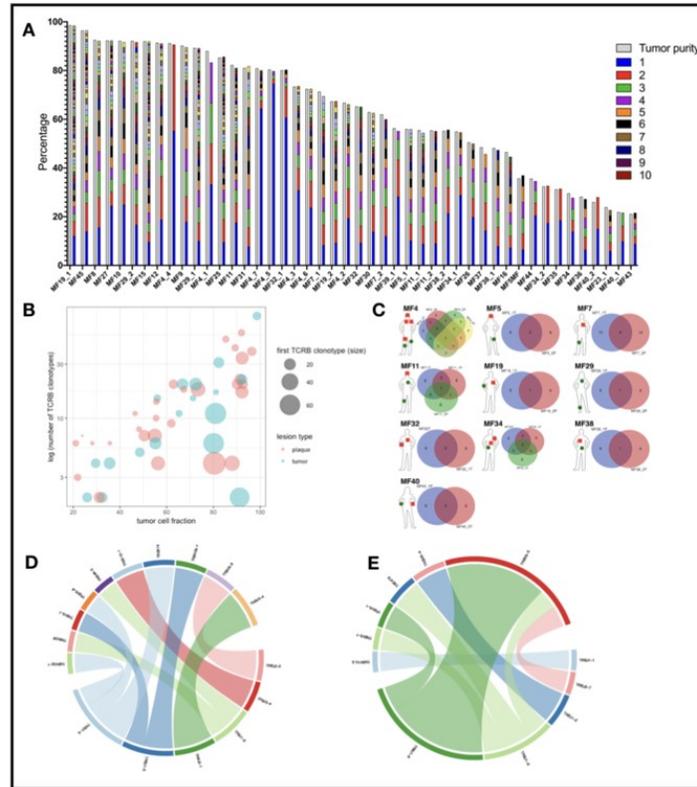


Figure 3.2: Clonotypic heterogeneity of skin lesions in MF.

(A) The tumor purity for the skin samples was estimated by the copy number aberration (CNA) data from the WES. The tumor cell fraction (grey bars) is plotted versus the cumulative frequency of the most abundant TCR β clonotypes. The frequencies of the clonotypes are represented with stacked bars representing clonotypes from the most abundant (rank #1) to the least frequent. The ranks of clonotypes are color-coded as in the legend. (B) Correlation between tumor cell fraction and the number of neoplastic clonotypes. Note that the clonotypic heterogeneity is not dependent on the stage of the lesion (tumor vs plaque). The size of the circle is proportional to the percentage of the most dominant (rank #1) TCR β clonotype. (C) Topological heterogeneity in MF. Venn diagrams illustrating the number of overlapping TCR β clonotypes across different skin lesions. The location and type of the lesion is plotted for each patient (green circle - plaque, red square - tumor). (D-E) VJ combination diversity of TCR β clonotypes in MF. The combinations of *VJ* genes of the neoplastic clonotypes is presented in (D). MF32 and (E) MF16.

In 10 patients we obtained biopsies from multiple skin lesions which enabled us to study clonotypic heterogeneity between different areas of the skin (Fig 3.2C). In 8 patients we compared the biopsies from the plaque and the tumor (late lesion) (patients MF4, MF5, MF7,

MF11, MF19, MF34, MF38 and MF40) and in 3 patients we compared lesions in the same stage (two plaques from MF29, stage T2 and two tumors from MF4 and MF32, stage T3). The heterogeneity of any single lesion was lower than the combined heterogeneity of all biopsies from the same patient, measured by the clonotype richness (number of different neoplastic TCR β clonotypes) and Simpson index (the probability that two clonotypes, randomly drawn from the sample are different) (Appendix **Fig B1**). Surprisingly, the degree of overlapping clonotypes between different lesions was very low (1-2 clonotypes) and in one case (MF7 tumor and plaque) we did not detect any shared clonotypes. Importantly, the shared clonotypes were not always the most frequent ones. Thus, extensive clonotypic heterogeneity was not detected only on the level of a single lesion, but also between different lesions (topological heterogeneity).

3.3.2 Neoplastic clonotypes are frequently detected in the peripheral blood in MF

Having established that MF shows high clonotypic heterogeneity, both with regard to the composition of a single lesion and between different lesions, we realized that this heterogeneity could not be generated in the skin *in situ* because cells in the MF infiltrate do not express RAG1/2 and TdT, the key enzymes needed for TCR gene recombination (ref. 39 and our unpublished data on RNAseq of MF). We considered the possibility that there is a pool of clonotypically heterogeneous neoplastic cells in the circulation which are able to seed the skin and undergo clonal expansion in the skin niche. Therefore, we investigated whether malignant T-cell clones could be found in the peripheral blood. For the consistency across samples, we assumed conservatively that the top 10 frequent TCR clonotypes from skin represent the true, tumor-related clonotypes. This assumption is based on the observation that

the 10 most abundant TCR β clonotypes contributed to up to 85% (95% upper confidence interval value) of the tumor-related clonotypes whereas the remaining clonotypes (ranked >10) had a low abundance (95% CI: 1.3%-1.5%,) and only contributed to 6%-15% (95% CI) of the total number of clonotypes. To provide a second layer of validation, we have also compared the *TCRA* and *TCRG* CDR3 sequences in the blood and the skin.

In 79% (15/19) of patients we have detected at least one shared TCR β clonotype between the skin and the blood. The same number of patients had one or more common TCR γ shared clonotype in the skin and blood, whereas 16 patients had circulating neoplastic TCR α clonotypes (**Fig 3.3A**, Appendix **Fig B2A and B3A**). Thus, all patients had at least one shared clonotype TCR α , - β or - γ between the skin and the blood at the time of sampling. The number of identical clonotypes in the skin and the blood was highest for TCR α (1-7 clonotypes, **Fig S2A**), probably due to the fact that a single clone of T-cells defined by a common TCR β clonotype may comprise 1-3 different TCR α clonotypes (see Appendix **Fig B4**). Importantly, the neoplastic clonotypes were composed of multiple *V-J* gene combinations that indicated they originated at the stage of T-cell development (**Fig 3.2D, E**).

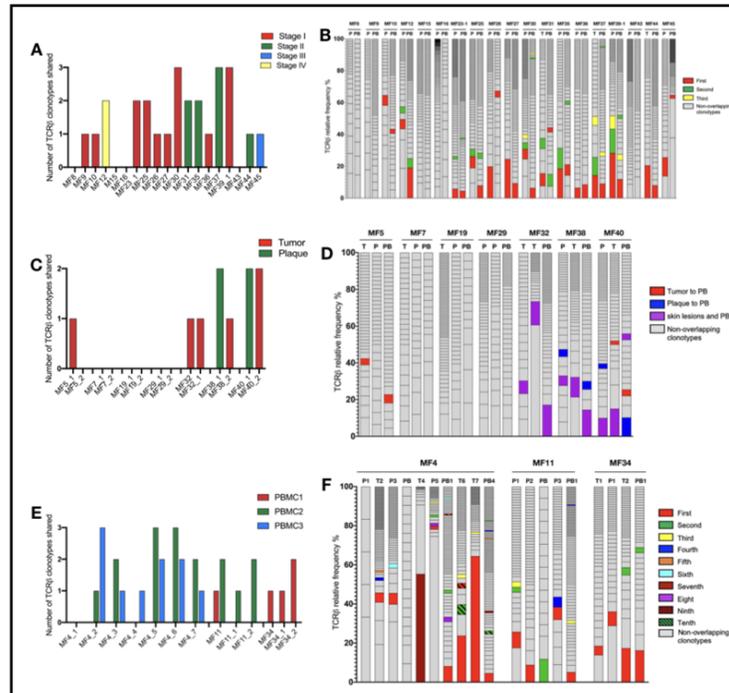


Figure 3.3: Detection of the neoplastic TCR β clonotypes in the peripheral blood.

The sequences of the TCR β clonotypes in the blood that matched any of the top 10 neoplastic TCR β clonotypes identified in the corresponding skin sample to detect the neoplastic clonotypes in the circulation. The number and frequency of those shared neoplastic clonotypes are shown separately for the three groups of patients as defined in **Fig 1**: **(A, B)**: 19 patients with a single biopsy, **C-F**: patients with multiple skin biopsies, of whom in 7 patients the biopsies were obtained at a single time point (**C, D**) whereas 3 patients were sampled longitudinally (**E, F**). In **B, D, F** the first ranking shared clonotype in the skin is indicated in red and the subsequent shared clonotypes are color-coded as indicated in the legend. The non-overlapping clonotypes are indicated in gray.

We subsequently analysed whether the frequency of the clonotypes in the skin correlated with the frequency of those in the blood. The dominant (most abundant) TCR β clonotype from the skin was identified in the blood in 8 patients and 6 of those clonotypes were also dominant in the blood (**Fig 3.3B**). The pattern was more complicated for TCR γ and TCR α , where dominant clonotypes could be detected in the blood and the skin in only 2 patients (MF36 and MF37) (**Fig S2B and S3B**).

Similar occurrence of tumor-derived clonotypes was detected in 10 patients in whom we analyzed multiple skin biopsies. 7 patients were analyzed with paired biopsies (tumor and plaque) and for 3 patients multiple biopsies (between 3 and 7). All patients shared at least one clonotype (TCR β , - α or - γ) between the blood and one or both skin biopsies (**Fig 3.3C**, Appendix **B2C**, **Fig B3C**) and the degree of clonotypic sharing was higher between the blood and the skin than between two skin biopsies (**Fig 3.3D**). No correlation was found between the number of shared clonotypes and the stage of the disease and the progression-free survival (Appendix **Table B1**).

It has not escaped our attention that some clonotypes were shared between different samples (both skin and blood) from different patients, such as the CDR3 sequence CDNNNDMRF (TRAV16/TRAJ43) which was found among malignant clonotypes of patients 8 of 29 patients and CAASRGC_AKNIQYF (TRBV18/TRBJ2-4) that was found in 20 of 29 patients (see Appendix **Table B2**). We have previously noticed frequent clonotypic sharing between patients with MF and excluded laboratory error as a possible cause.²¹ We have also excluded the possibility that those sequences represented sequences parts of unrelated captured exomes, with the secondary verification using blastn and blastp that indicated the sequences to be TCR (data not shown).

3.3.3 Temporal dynamics of malignant TCR clonotypes in the skin and blood

Clonotypic heterogeneity could be achieved by seeding the skin with malignant clones, a mechanism which is responsible for the formation of metastases in solid tumors.⁴⁰⁻⁴³

Metastatic seeding may occur by single cancer cells (in which case the metastasis represents a single subclone) or via continuous seeding when clusters of cancer cells transfer the entire

heterogeneity of the primary tumor to the metastases.^{41–43} The third seeding mechanism, referred to as the consecutive seeding, relies on sequential recruitment of neoplastic cells to the metastases and result in the metastatic lesions that only represent a fraction of the heterogeneity of the primary tumor (Appendix **Fig B5**).⁴³ Clonotypic heterogeneity of skin lesions excluded the single-cell seeding events whereas the fact that we detected single neoplastic clonotypes in the blood rather than combinations of different clonotypes argued against continuous seeding. To further elucidate the mechanism of tumor seeding, we followed the malignant clonotypes in the skin and the blood in 3 patients (MF4, MF11, MF34) over a period of 9 to 22 months (**Fig 3.3E, F and Fig 3.4**). In each case we found neoplastic T-cell clonotypes in the blood, defined as those *TCRB* CDR3 sequences that were found in at least one skin biopsy. The number of circulating neoplastic clonotypes varied from 2 clonotypes in MF34 to 10 clonotypes in MF4. Circulating neoplastic clonotypes were not detected constantly in all blood samples and certain CDR3 sequences could be detected in the blood before occurrence in skin biopsies, e.g. GPGTRLLVLGERGLLGRGRGR_ WVFRLRGVPGLCSGANVLTF or CASCYPH_VSCRRP that were found in the blood of patient MF4 months before their detection in skin biopsies (**Fig 3.4A**). Together with the finding that a single skin lesion contained only a fraction of all possible neoplastic clonotypes, our data strongly supported the model of the development of the skin lesions in MF by consecutive seeding (**Fig 3.5 and Appendix Fig B5**).

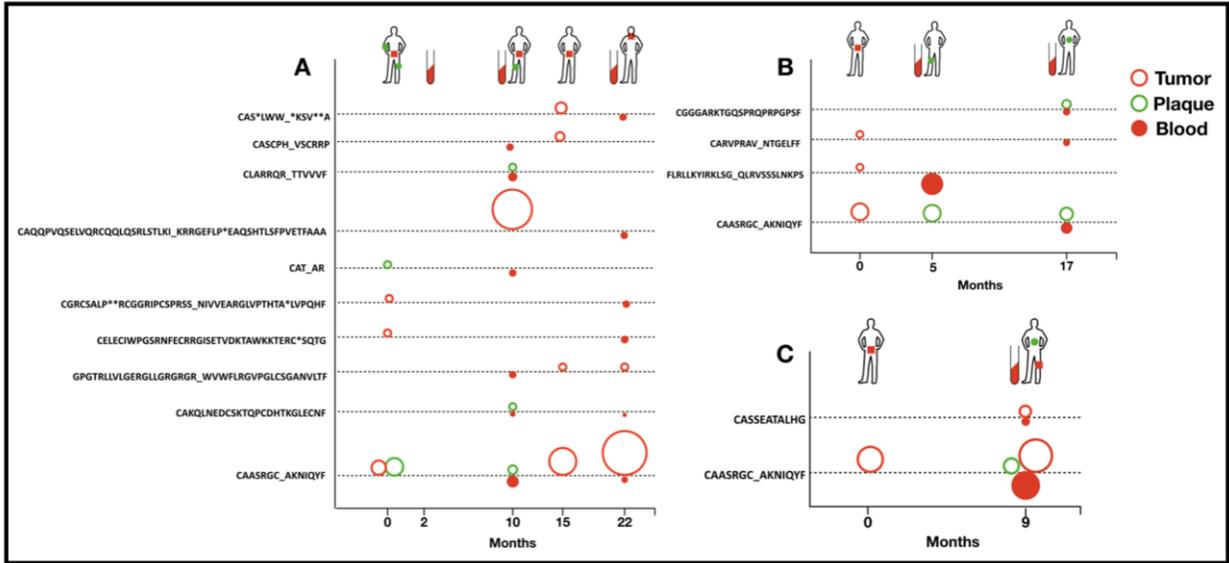


Figure 3.4: Dynamics of neoplastic clonotypes in the skin and the blood.

Three patients were followed longitudinally with multiple skin biopsies and/or blood sampling. All shared neoplastic clonotypes are plotted on the time axis for individual patients (A) MF4, (B) MF11 and (C) MF34. The location and type of the lesion is indicated for each patient on a silhouette (green circle - plaque, red square - tumor). Each dotted line corresponds to a single, shared clonotype of the indicated amino acid sequence of CDR3 β . Circles above the line are skin clonotypes (open red- tumor, open green- plaque) whereas the solid red circles below the line are the clonotypes detected in the blood. The size of the dot is proportional to the frequency of the shared clonotype in the sample.

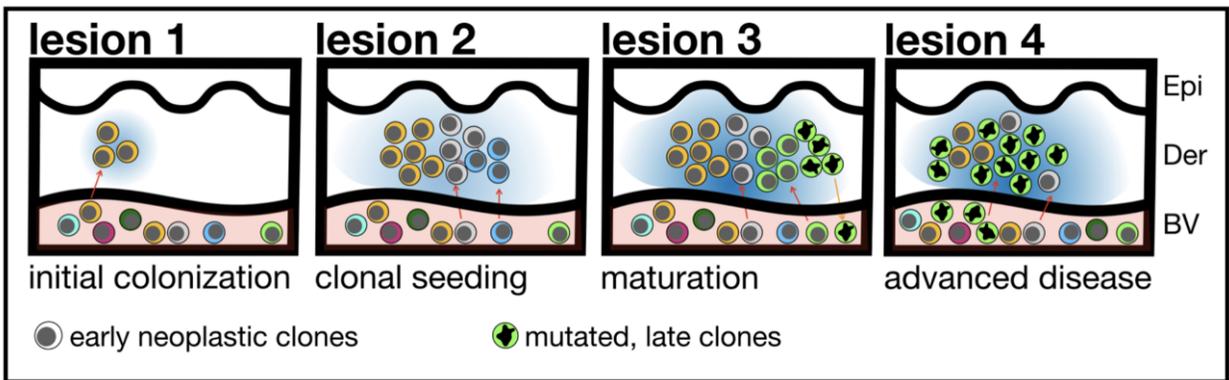


Figure 3.5: The hypothesis of tumor seeding in the pathogenesis of mycosis fungoides.

Even in the early stages of the disease the patients have circulating neoplastic T-cell clones in peripheral blood. Early lesions are initiated by the pioneer clones and create a niche (blue-shaded area) that facilitates seeding of this area of the skin (“lesion 1”) with subsequent clones (“lesion 2”) (consecutive seeding model⁴³ and Fig S6). The clonal composition of different lesions may differ (“lesion 3”) due to the stochastic nature of cancer seeding. Some clones

may have a higher proliferation capacity in the skin and may overgrow other clones (green clone in “lesion 3”, further mutate, re-enter the circulation (orange arrow) and re-seed another area of the skin (lesion 4). The figures represent symbolically the structure of the skin with the epidermis (Epi), dermis (Der) and a pink-shaded blood vessel (BV). Different clones of neoplastic T-cells are marked with different colors.

3.5 Discussion

The major finding of this study was that lesions of mycosis fungoides comprise a highly diverse collection of malignant T-cell clones which percolate between the skin and the blood. We propose that consecutive seeding of circulating neoplastic T-cell subclones in the skin is a likely mechanism of growth and evolution of the lesions in this cutaneous lymphoma.

Difficulties in distinguishing between the CDR3 sequences in malignant T-cells from those of normal, reactive T-cells has been the major limitation of TCR analyses. Several indirect approaches have been used to mitigate this problem, such as setting arbitrary thresholds for clonotypic frequencies or surrogate measures of tumor cell fraction as a ratio of the frequency of the most abundant TCR β clonotype to the sum of all TCR β clonotypes.³⁷ Obviously, both approaches assume that all tumor cells are clonal (i.e. share a single TCR β clonotype). This assumption is a cornerstone of the theory of CTCL as a neoplasm of the mature T-cell but has not rigorously been tested experimentally. Experimental findings suggesting clonotypic heterogeneity in CTCL⁴⁴ have either been ignored or attributed to contamination by clonal inflammatory T-cells⁴⁵, activation of skin resident T-cells by superantigens⁴⁴ or age-dependent clonal expansion.⁴⁶⁻⁴⁹ We were able to solve this methodological problem by calculating the tumor cell fraction in the sample by WES and thus directly enumerate the clonotypes derived from neoplastic T-cells²¹. Our method relies on DNA sequencing which eliminates a potential

error due to aberrant TCR expression in cancer cells. With this approach we confirmed our previous findings²⁰ demonstrating clonal heterogeneity in MF. Our findings contribute to the increasing body of evidence showing that CTCL as well as other T-cell leukemias and lymphomas exhibit genomic and transcriptomic heterogeneity, similar to what is seen in solid tumors.^{50–53}

We have found a large variation in the clonotypic richness (number of different TCR β clonotypes) in MF lesions, ranging from 2 up to 80 distinct clonotypes. In most samples, the frequency of the most abundant TCR β clonotype has been relatively high (19.4%) which is well above the usual 15% threshold accepted by many authors as a hallmark of monoclonality. This is reflected by a relatively low Simpson index (a probability that a random draw from the pool of clonotypes yields two different clonotypes; median 7% 1st and 3rd quartile 3.2% - 10.7%) (Appendix **Fig B1**). Thus, even consecutive draws from the pool of clonotypes of a given lesion are most likely to yield identical clonotypes, which can be misinterpreted as monoclonality (e.g. 10 consecutive draws yield identical results in ~50% of cases). In 33% of biopsies, the frequency of the first TCR β clone was lower than 10%, which corresponds well to the observed proportion of cases of MF in which monoclonality cannot be detected by standard assays.⁵⁴

There is compelling evidence that staphylococcal enterotoxins facilitate progression of MF.⁵⁵ Eradication of *S. aureus* with antibiotics leads to spectacular responses in some cases of MF.⁵⁶ It is controversial whether superantigens stimulate the neoplastic cells directly or rather indirectly, via activation of IL-2 by bystander T-cells.^{55,57} Since binding of the superantigen to

TCR is efficient for only a small subset of TCR V β ⁵⁸ we asked whether those TRVB segments are found in neoplastic cells. In contrast to previous reports⁵⁹⁻⁶¹ we did not see any obvious evidence of overrepresentation of superantigen-reactive TRVB segments, although the presence of TRVB5.3, TRVB5.1, TRVB19 was noted in 5/29 patients (Appendix **Table B2**). Larger studies are needed to determine whether patients expressing those TCR β variants are more responsive to antibiotic therapy and *S.aureus* elimination.

We have considered potential mechanisms which could explain the finding of clonotypic heterogeneity in absence of a true clonal diversity. We have rejected the idea that clonotypic heterogeneity is attributed to secondary somatic mutations within the already assembled CDR3 region because we detected that the rearranged *TCRB* were composed by numerous combinations of various VJ segments. Moreover, extensive secondary mutations within TCRB comes from the quantification of the ratio between TCR α and TCR β clonotypes. In a population of normal memory T-cells the *TCRA/TCRB* is between 2 and 3 because *TCRA* rearrangements are usually biallelic and there are occasional secondary rearrangements of this locus.⁶² Mutations in *TCRB* would result in an increase in the number of unique TCR β sequences and thus in a decrease in the *TCRA/TCRB* ratio, which was not the case in our material (*TCRA/TCRB* ratio was 2.76). Third, we have detected a higher than expected overlap between clonotypes between patients (Appendix **Table S2**) which argues against the influence of random mutations. Clonotypic richness of TCR β would also be increased if malignant T-cells were not subject to allelic exclusion of *TCRB* locus. However, this scenario is unlikely because allelic exclusion is a stochastic process, which is very resistant to perturbations^{63,64} and is not violated in other T-cell malignancies investigated to date.^{52,65} Lack of allelic

exclusion would also decrease the *TCRA/TCRB* ratio, in a similar way as described above for secondary mutations. Finally, recombination of *TCRB* in post-thymic lymphocytes is impossible because neither normal nor neoplastic T-cells in MF express the essential RAG1/2 recombinases.³⁹

Conceptually, the clonotypic heterogeneity described here is different from the mutational subclonal heterogeneity, because it cannot be generated continuously in the tumor but only in the time span when RAG1/2 recombinases are active (i.e. at the level of immature T-cell precursor). It has been suggested though, that multiple malignant clones can be generated from the pool of normal, reactive lymphocytes in the skin undergoing malignant transformation.⁴⁴ However, such a mechanism would have to operate at an unprecedented efficiency to generate hundreds of cancer clones simultaneously in different areas of the skin and is therefore unlikely. The most likely mechanism is by accrual of malignant T-cells clones from the circulation to the skin. Several lines of indirect evidence support such a hypothesis. First, we were able to detect malignant TCR clonotypes in the blood. It has long been known that TCR γ clonotypes identical to those in lesional skin could be detected in the blood in MF in 5%-10% of patients in early stage disease without any prognostic impact.^{49,66-68} We have found that the presence of circulating malignant clones is a rule rather than an exception; at least one malignant TCR β or TCR α clonotype was present in the peripheral blood in all examined patients. Although only a small fraction of the entire pool of neoplastic clonotypes could be detected in the blood, the chances of detecting TCR β clonotypes increased by repeated sampling, probably because of their very low frequency compared to the background of normal T-cells. Second, the neoplastic clonotypes were found in the blood even in the very

early stages of the disease, they were not correlated with the stage of the disease nor they always represented the dominant clonotype in the skin. Therefore, those circulating clones could not solely represent subclinical leukemization due to disease progression. Third, clonotypic overlap (the number of shared TCR α , - β or - γ clonotypes) was higher between the skin and the peripheral blood than between discrete skin lesions.

Taken together, these observations are compatible with the presence of a pool of very diverse neoplastic T-cell clones in the peripheral blood that may seed at different frequency to the skin where they develop further into the lesions of MF (**Fig 5 and Appendix B6**). Our findings are compatible with the model of consecutive seeding in which only a fraction of all neoplastic clones transfers the diversity to the developing lesions of lymphoma. Previous modeling of consecutive metastatic seeding events revealed that each lesion is likely to be funded by 10-150 cancer clones⁴³, a number which is in the same range as the 6-20 clonotypes (1st-3rd quartiles) detected by us in MF. Although not investigated here directly, we hypothesize that neoplastic clones do not migrate unidirectionally from the blood to the skin. Recent findings that downregulation of CD69 enables skin resident memory cells to exit the tissue and to recirculate⁶⁹ indicates that neoplastic cells in MF should also be able to re-enter the circulation and contribute to the pool of circulating neoplastic clones. This mechanism is reminiscent of the phenomenon of tumor self-seeding^{40,70} where circulating metastatic cells colonize the primary tumor. Tumor self-seeding is enhanced by changes in the tissue niche occupied by cancer, which makes the niche more receptive for the subsequent entry of the waves of circulating tumor cells.⁷¹ In MF, the source of those dermatropic neoplastic T-cell clones remains unknown, although there is some evidence that bone marrow may play such a role.¹¹⁻

¹³ Research showed that the bone marrow niche provides shelter for different types of malignant cells (e.g. breast or prostate cancers) and that the bone marrow pool of cancer might be responsible for relapses after therapy.^{71,72}

We wish to highlight several limitations of this study. Sensitivity of WES for TCR analyses is limited and our method does not detect rare clonotypes. It is therefore possible that we have not captured the entire repertoire of malignant clonotypes. It is also evident that microdissected cells are not representative for the heterogeneity of the entire lesion. It is likely that different tumor clones are not distributed homogeneously throughout the MF lesion and our sampling method underestimates the true clonal diversity. Finally, our evidence of clonal heterogeneity relies on the statistical analyses of the frequency of different clonotypes in relation to the tumor cell fraction rather than analyses of individual cells. With the emerging technologies allowing for mutation calling and TCR sequencing in single cells^{52,73} it should be possible to confirm our findings directly and by independent methods.

The proposed mechanism of tumor self-seeding in the pathogenesis of MF may have several practical implications. Circulating neoplastic T-cell clones may constitute an interesting target for therapy. It is tempting to speculate that mogamulizumab⁷⁴ which blocks CCR4, an essential cutaneous homing receptor⁷⁵, exerts its therapeutic efficacy via inhibition of skin seeding by circulating cancer clones. Identification of TCR β clonotypes could also be used diagnostically in CTCL. Interestingly, we have found significant interindividual overlap in TCR β clonotypic sequences, dramatically exceeding the frequency of shared, public clonotypes in healthy individual. Although the mechanism of clonotypic sharing remains

elusive, the common *TCRB* CDR3 sequences could easily, quantitatively and cost-effectively be detected with PCR-based techniques and used for the diagnosis and monitoring of therapy in CTCL.

3.6 References

1. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375–2390.
2. Olsen E, Vonderheid E, Pimpinelli N, et al. Revisions to the staging and classification of mycosis fungoides and Sezary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the cutaneous lymphoma task force of the European Organization of Research and Treatment of Cancer (EORTC). *Blood*. 2007;110(6):1713–1722.
3. Kim YH, Willemze R, Pimpinelli N, et al. TNM classification system for primary cutaneous lymphomas other than mycosis fungoides and Sezary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the Cutaneous Lymphoma Task Force of the European Organization of Research and Treatment of Cancer (EORTC). *Blood*. 2007;110(2):479–484.
4. Willemze R. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768–3785.
5. Clark RA, Watanabe R, Teague JE, et al. Skin effector memory T cells do not recirculate and provide immune protection in alemtuzumab-treated CTCL patients. *Sci. Transl. Med.* 2012;4(117):117ra7.
6. Campbell JJ, Clark RA, Watanabe R, Kupper TS. Sezary syndrome and mycosis fungoides arise from distinct T-cell subsets: a biologic rationale for their distinct clinical behaviors. *Blood*. 2010;116(5):767–771.
7. Whittaker S, Ortiz P, Dummer R, et al. Efficacy and safety of bexarotene combined with psoralen-ultraviolet A (PUVA) compared with PUVA treatment alone in stage IB-IIA

- mycosis fungoides: final results from the EORTC Cutaneous Lymphoma Task Force phase III randomized clinical trial 21011. *Br. J. Dermatol.* 2012;167(3):678–687.
8. Kamstrup MR, Gniadecki R, Iversen L, et al. Low-dose (10-Gy) total skin electron beam therapy for cutaneous T-cell lymphoma: an open clinical study and pooled data analysis. *Int. J. Radiat. Oncol. Biol. Phys.* 2015;92(1):138–143.
 9. Chowdhary M, Chhabra AM, Kharod S, Marwaha G. Total Skin Electron Beam Therapy in the Treatment of Mycosis Fungoides: A Review of Conventional and Low-Dose Regimens. *Clin. Lymphoma Myeloma Leuk.* 2016;16(12):662–671.
 10. Vieyra-Garcia P, Fink-Puches R, Porkert S, et al. Evaluation of Low-Dose, Low-Frequency Oral Psoralen-UV-A Treatment With or Without Maintenance on Early-Stage Mycosis Fungoides: A Randomized Clinical Trial. *JAMA Dermatol.* 2019;
 11. Gniadecki R, Lukowsky A, Rossen K, et al. Bone marrow precursor of extranodal T-cell lymphoma. *Blood.* 2003;102(10):3797–3799.
 12. Berg KD, Brinster NK, Huhn KM, et al. Transmission of a T-cell lymphoma by allogeneic bone marrow transplantation. *N. Engl. J. Med.* 2001;345(20):1458–1463.
 13. Fahy CMR, Fortune A, Quinn F, et al. Development of mycosis fungoides after bone marrow transplantation for chronic myeloid leukaemia: transmission from an allogeneic donor. *Br. J. Dermatol.* 2014;170(2):462–467.
 14. Davis TH, Morton CC, Miller-Cassman R, Balk SP, Kadin ME. Hodgkin's disease, lymphomatoid papulosis, and cutaneous T-cell lymphoma derived from a common T-cell clone. *N. Engl. J. Med.* 1992;326(17):1115–1122.
 15. Kamstrup MR, Gniadecki R, Skovgaard GL. Putative cancer stem cells in cutaneous malignancies. *Exp. Dermatol.* 2007;16(4):297–301.

16. Delfau-Larue M, Petrella T, Lahet C, et al. Value of clonality studies of cutaneous T lymphocytes in the diagnosis and follow-up of patients with mycosis fungoides. *J. Pathol.* 1998;184(2):185–190.
17. Jones D, Duvic M. The current state and future of clonality studies in mycosis fungoides. *J. Invest. Dermatol.* 2003;121(3):ix–x.
18. Mahe E, Pugh T, Kamel-Reid S. T cell clonality assessment: past, present and future. *J. Clin. Pathol.* 2018;71(3):195–200.
19. Krangel MS. Mechanics of T cell receptor gene rearrangement. *Curr. Opin. Immunol.* 2009;21(2):133–139.
20. Hamrouni A, Fogh H, Zak ZL, Odum N, Gniadecki R. Clonotypic diversity of the T-cell receptor corroborates the immature precursor origin of cutaneous T-cell lymphoma. *Clin. Cancer Res.* 2019;
21. Iyer A, Hennessey D, O’Keefe S, et al. Clonotypic heterogeneity in cutaneous T-cell lymphoma (mycosis fungoides) revealed by comprehensive whole-exome sequencing. *Blood Adv.* 2019;3(7):1175–1184.
22. Boyman O, Hefti HP, Conrad C, et al. Spontaneous Development of Psoriasis in a New Animal Model Shows an Essential Role for Resident T Cells and Tumor Necrosis Factor- α . *J. Exp. Med.* 2004;199(5):731–736.
23. Watanabe R, Gehad A, Yang C, et al. Human skin is protected by four functionally and phenotypically discrete populations of resident and recirculating memory T cells. *Sci. Transl. Med.* 2015;7(279):279ra39.
24. Langerak AW, Molina TJ, Lavender FL, et al. Polymerase chain reaction-based clonality testing in tissue samples with reactive lymphoproliferations: usefulness and pitfalls. A

- report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia*. 2007;21(2):222–229.
25. Marcus Mucbe J, Karenko L, Gellrich S, et al. Cellular coincidence of clonal T cell receptor rearrangements and complex clonal chromosomal aberrations—a hallmark of malignancy in cutaneous T cell lymphoma. *J. Invest. Dermatol.* 2004;122(3):574–578.
 26. Mucbe JM, Lukowsky A, Heim J, et al. Demonstration of frequent occurrence of clonal T cells in the peripheral blood but not in the skin of patients with small plaque parapsoriasis. *Blood*. 1999;94(4):1409–1417.
 27. Beylot-Barry M, Sibaud V, Thiebaut R, et al. Evidence that an identical T cell clone in skin and peripheral blood lymphocytes is an independent prognostic factor in primary cutaneous T cell lymphomas. *J. Invest. Dermatol.* 2001;117(4):920–926.
 28. Fraser-Andrews EA, Woolford AJ, Russell-Jones R, Whittaker SJ, Seed PT. Detection of a Peripheral Blood T Cell Clone is an Independent Prognostic Marker in Mycosis Fungoides. *J. Invest. Dermatol.* 2000;114(1):117–121.
 29. Bolotin DA, Poslavsky S, Mitrophanov I, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods*. 2015;12:380.
 30. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 2011;43(5):491–498.
 31. Ha G, Roth A, Khattra J, et al. TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. *Genome Res*. 2014;24(11):1881–1893.

32. Nazarov VI, Pogorelyy MV, Komech EA, et al. tcR: an R package for T cell receptor repertoire advanced data analysis. *BMC Bioinformatics*. 2015;16:175.
33. Pogorelyy MV, Elhanati Y, Marcou Q, et al. Persisting fetal clonotypes influence the structure and overlap of adult human T cell receptor repertoires. *PLoS Comput. Biol.* 2017;13(7):e1005572.
34. Dupic T, Marcou Q, Walczak AM, Mora T. Genesis of the $\alpha\beta$ T-cell receptor. *PLoS Comput. Biol.* 2019;15(3):e1006874.
35. Khor B, Sleckman BP. Allelic exclusion at the TCRbeta locus. *Curr. Opin. Immunol.* 2002;14(2):230–234.
36. Mora T, Walczak AM. Quantifying lymphocyte receptor diversity. *arXiv*:1604.00487
37. de Masson A, O'Malley JT, Elco CP, et al. High-throughput sequencing of the T cell receptor β gene identifies aggressive early-stage mycosis fungoides. *Sci. Transl. Med.* 2018;10(440.):
38. Kirsch IL, Sherwood A, Williamson D, et al. The utility of HTS TCR analyses in the management of patients with T-cell malignancies. *J. Clin. Oncol.* 2014;32(15_suppl):8565–8565.
39. Döbbeling U, Dummer R, Hess Schmid M, Burg G. Lack of expression of the recombination activating genes RAG-1 and RAG-2 in cutaneous T-cell lymphoma: pathogenic implications. *Clin. Exp. Dermatol.* 1997;22(5):230–233.
40. Kim M-Y, Oskarsson T, Acharyya S, et al. Tumor Self-Seeding by Circulating Cancer Cells. *Cell*. 2009;139(7):1315–1326.
41. Aceto N, Bardia A, Miyamoto DT, et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell*. 2014;158(5):1110–1122.

42. Cleary AS, Leonard TL, Gestl SA, Gunther EJ. Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature*. 2014;508(7494):113–117.
43. Heyde A, Reiter JG, Naxerova K, Nowak MA. Consecutive seeding and transfer of genetic diversity in metastasis. *Proc. Natl Acad. Sci.* 2019;116(28):14129–14137.
44. Vega F, Luthra R, Medeiros LJ, et al. Clonal heterogeneity in mycosis fungoides and its relationship to clinical course. *Blood*. 2002;100(9):3369–3373.
45. Kolowos W, Herrmann M, Ponner BB, et al. Detection of restricted junctional diversity of peripheral T cells in SLE patients by spectratyping. *Lupus*. 1997;6(9):701–707.
46. Gniadecki R, Lukowsky A. Monoclonal T-cell dyscrasia of undetermined significance associated with recalcitrant erythroderma. *Arch. Dermatol.* 2005;141(3):361–367.
47. Wack A, Cossarizza A, Heltai S, et al. Age-related modifications of the human alphabeta T cell repertoire due to different clonal expansions in the CD4+ and CD8+ subsets. *Int. Immunol.* 1998;10(9):1281–1288.
48. Posnett DN, Sinha R, Kabak S, Russo C. Clonal populations of T cells in normal elderly humans: the T cell equivalent to “benign monoclonal gammopathy.” *J. Exp. Med.* 1994;179(2):609–618.
49. Muche JM, Sterry W, Gellrich S, et al. Peripheral blood T-cell clonality in mycosis fungoides and nonlymphoma controls. *Diagn. Mol. Pathol.* 2003;12(3):142–150.
50. Buus TB, Willerslev-Olsen A, Fredholm S, et al. Single-cell heterogeneity in Sézary syndrome. *Blood Adv.* 2018;2(16):2115–2126.

51. Gaydosik AM, Tabib T, Geskin LJ, et al. Single-Cell Lymphocyte Heterogeneity in Advanced Cutaneous T-cell Lymphoma Skin Tumors. *Clin. Cancer Res.* 2019;25(14):4443–4454.
52. Farmanbar A, Kneller R, Firouzi S. RNA sequencing identifies clonal structure of T-cell repertoires in patients with adult T-cell leukemia/lymphoma. *NPJ Genom Med.* 2019;4:10.
53. Litvinov IV, Tetzlaff MT, Thibault P, et al. Gene expression analysis in Cutaneous T-Cell Lymphomas (CTCL) highlights disease heterogeneity and potential diagnostic and prognostic indicators. *Oncoimmunology.* 2017;6(5):e1306618.
54. Dereure O, Levi E, Vonderheid EC, Kadin ME. Improved sensitivity of T-cell clonality detection in mycosis fungoides by hand microdissection and heteroduplex analysis. *Arch. Dermatol.* 2003;139(12):1571–1575.
55. Woetmann A, Lovato P, Eriksen KW, et al. Nonmalignant T cells stimulate growth of T-cell lymphoma cells in the presence of bacterial toxins. *Blood.* 2007;109(8):3325–3332.
56. Lindahl LM, Willerslev-Olsen A, Gjerdrum LMR, et al. Antibiotics inhibit tumor and disease activity in cutaneous T cell lymphoma. *Blood.* 2019;
57. Willerslev-Olsen A, Krejsgaard T, Lindahl LM, et al. Staphylococcal enterotoxin A (SEA) stimulates STAT3 activation and IL-17 expression in cutaneous T-cell lymphoma. *Blood.* 2016;127(10):1287–1296.
58. Fraser JD, Proft T. The bacterial superantigen and superantigen-like proteins. *Immunol. Rev.* 2008;225:226–243.

59. Vonderheid EC, Bigler RD, Hou JS. On the possible relationship between staphylococcal superantigens and increased Vbeta5.1 usage in cutaneous T-cell lymphoma. *Br. J. Dermatol.* 2005;152(4):825–6; author reply 827.
60. Vonderheid EC, Boselli CM, Conroy M, et al. Evidence for restricted Vbeta usage in the leukemic phase of cutaneous T cell lymphoma. *J. Invest. Dermatol.* 2005;124(3):651–661.
61. Jackow CM, Cather JC, Hearne V, et al. Association of erythrodermic cutaneous T-cell lymphoma, superantigen-positive *Staphylococcus aureus*, and oligoclonal T-cell receptor V beta gene expansion. *Blood.* 1997;89(1):32–40.
62. Hamrouni A, Aublin A, Guillaume P, Maryanski JL. T cell receptor gene rearrangement lineage analysis reveals clues for the origin of highly restricted antigen-specific repertoires. *J. Exp. Med.* 2003;197(5):601–614.
63. Farcot E, Bonnet M, Jaeger S, et al. TCR β Allelic Exclusion in Dynamical Models of V(D)J Recombination Based on Allele Independence. *J. Immunol.* 2010;185(3):1622–1632.
64. Jaeger S, Lima R, Meyroneinc A, et al. A dynamical model of TCR β gene recombination: Coupling the initiation of D β -J β rearrangement to TCR β allelic exclusion. *bioRxiv.* 2019; <https://doi.org/10.1101/200444>
65. Brown SD, Hapgood G, Steidl C, et al. Defining the clonality of peripheral T cell lymphomas using RNA-seq. *Bioinformatics.* 2017;33(8):1111–1115.
66. Muche JM, Marcus Muche J, Lukowsky A, et al. Peripheral Blood T Cell Clonality in Mycosis Fungoides –An Independent Prognostic Marker? *J. Invest. Dermatol.* 2000;115(3):504–505.

67. Scarisbrick JJ, Hodak E, Bagot M, et al. Developments in the understanding of blood involvement and stage in mycosis fungoides/Sezary syndrome. *Eur. J. Cancer*. 2018;101:278–280.
68. Talpur R, Singh L, Daulat S, et al. Long-term outcomes of 1,263 patients with mycosis fungoides and Sézary syndrome from 1982 to 2009. *Clin. Cancer Res*. 2012;18(18):5051–5060.
69. Klicznik MM, Morawski PA, Höllbacher B, et al. Human CD4CD103 cutaneous resident memory T cells are found in the circulation of healthy individuals. *Sci Immunol*. 2019;4(37.):
70. Massagué J, Obenauf AC. Metastatic colonization by circulating tumour cells. *Nature*. 2016;529(7586):298–306.
71. Shiozawa Y, Pedersen EA, Havens AM, et al. Human prostate cancer metastases target the hematopoietic stem cell niche to establish footholds in mouse bone marrow. *J. Clin. Invest*. 2011;121(4):1298–1312.
72. Zhang XH-F, Jin X, Malladi S, et al. Selection of bone metastasis seeds by mesenchymal signals in the primary tumor stroma. *Cell*. 2013;154(5):1060–1073.
73. Singh M, Al-Eryani G, Carswell S, et al. High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes. *Nat. Commun*. 2019;10(1):3120.
74. Kim YH, Bagot M, Pinter-Brown L, et al. Mogamulizumab versus vorinostat in previously treated cutaneous T-cell lymphoma (MAVORIC): an international, open-label, randomised, controlled phase 3 trial. *Lancet Oncol*. 2018;19(9):1192–1204.

75. Soler D, Humphreys TL, Spinola SM, Campbell JJ. CCR4 versus CCR10 in human cutaneous TH lymphocyte trafficking. *Blood*. 2003;101(5):1677–1682.

Chapter 4: Branched evolution and genomic intratumor heterogeneity in the pathogenesis of cutaneous T-cell lymphoma*

4.1 Abstract

Mycosis fungoides (MF) is a slowly progressive cutaneous T-cell lymphoma (CTCL) for which there is no cure. In the early plaque stage, the disease is indolent, but development of tumors heralds an increased risk of metastasis and death. Previous research into the genomic landscape of CTCL revealed a complex pattern of >50 driver mutations implicated in more than a dozen of signaling pathways. However, the genomic mechanisms governing disease progression and treatment resistance remain unknown. Building on our previous discovery of the clonotypic heterogeneity of MF, we hypothesized that this lymphoma does not progress in a linear fashion as currently thought but comprises heterogeneous mutational subclones. We sequenced exomes of 49 cases of MF and identified 28 previously unreported putative driver genes. MF exhibited extensive intratumoral heterogeneity (ITH) of a median of six subclones showing branched pattern of phylogenetic relationships. Stage progression was correlated with an increase in ITH and redistribution of mutations from the stem to the clades. The pattern of clonal driver mutations was highly variable with no consistent mutations between patients. A similar intratumoral heterogeneity was detected in leukemic CTCL (Sezary syndrome). Based on these findings we propose a model of the pathogenesis of MF comprising neutral, divergent evolution of cancer subclones and discuss how ITH impacts the efficacy of targeted drug therapies and immunotherapies of CTCL.

* A version of this thesis is published online as a preprint in the format of A. Iyer, D. Hennessey, S. O’Keefe, J. Patterson, W. Wang, G. Ka-Shu Wong, R. Gniadecki, Branched evolution and genomic intratumor heterogeneity in the pathogenesis of cutaneous T-cell lymphoma. *bioRxiv*. 2019;804351. This chapter will subsequently be submitted for peer review publication.

4.2 Background

Cutaneous T-cell lymphoma (CTCL) is the most common form of T-cell lymphoma representing 5-10% of the total non-Hodgkin lymphomas.^{1,2} Mycosis fungoides (MF) is the prevalent form of CTCL that initially presents as erythematous, scaly patches and plaques on the skin (stage T1-T2, early lesions) which eventually progresses to advanced lesions, tumors (T3) and erythroderma (T4). Progression to stage T3 is a threshold event during the clinical evolution of MF, associated with a rapid drop in 5-year overall survival from >80% to 44%.³ The tumors may appear de novo on the clinically normal skin or may arise within the pre-existing plaque. Therefore, stage T3 patients most often present with a combination of early and late lesions. Thus, MF presents a unique opportunity to study the genetic mechanism of the progression of a T-cell lymphoma and to analyse phyletic relationships between cancer clones in early and advanced stages.

The genomic hallmarks of progressing MF have not been investigated in detail. The majority (84%) of the currently available sequencing data are not derived from MF but from the Sézary syndrome, a leukemic form of CTCL which albeit related, is an entity different from MF.⁴⁻¹³ The constellation of mutations in MF is very complex comprising at least 55 potential driver genes, with a considerable variability between patients.⁸ In advanced disease, mutations in the p53 and NF-kB pathways occur frequently and are mutually exclusive, but the survival seems to be similar whichever pathway is affected.¹⁴ It is also unknown how the pattern of genomic mutations correlate with the phenotype of the lesion.¹⁴

We hypothesized that lack of a repetitive constellation of mutations in CTCL is a result of the heterogeneity in the mutational processes. It is well documented that in heterogeneous cancers, the bulk sequencing approach will not provide insight into the essential driver genes

because most of the detected mutations will be subclonal and only relevant for a subpopulation of malignant cells.¹⁵ Intratumor heterogeneity (ITH), i.e. existence of genetically different subclones of neoplastic cells within the tumor, has been extensively studied and well documented in different types of solid cancers.¹⁵⁻¹⁷ ITH turned out to be important in tumor evolution which happens via selection of the fittest subclones. Thus tumors with a high degree of ITH tend to be more aggressive and notoriously difficult to cure because presence of subclones increases the risk of metastasis, facilitates immunological escape and development of resistance to chemotherapy and immunotherapy.¹⁵⁻¹⁷

The question whether ITH plays a role in the pathogenesis of CTCL has never been addressed before. CTCL is considered to represent a relatively homogenous malignancy in which all neoplastic cells descend from a transformed, mature T-cell in the skin.¹⁸ We have recently provided evidence supporting an alternative scenario. By analysing the repertoire of T-cell receptor (TCR) sequences in malignant cells in MF, we found that even early lesions, such as patches and plaques, show a considerable level of clonotypic diversity which occurs via seeding of early precursors in the skin.¹⁹⁻²¹ It is conceivable that those malignant clonotypes differ with respect to their mutational history and provide material for ITH in MF.

In this study we were able to confirm the existence of extensive subclonal heterogeneity in MF. We describe differences between early and advanced lesions with respect to the distribution of driver mutations and copy number aberrations and show patterns of branched evolution of cancer genomes during the clinical progression of the disease.

4.3 Material and Methods

4.3.1. Samples and sequencing

Ethical approval HREBA.CC-16-0820-REN1 was obtained from the Health Research Ethics Board of Alberta, Cancer Committee. Material (4mm punch skin biopsies from lesional skin and 10 ml blood) was collected from 31 consented patients with the diagnosis of mycosis fungoides in stages IA-IVA2 (**Fig 4.1** and Appendix **Table C1**). The biopsies and blood were processed for storage as explained in previous methods.²⁰ Frozen biopsies were sectioned at 10 μ and microdissected to isolate clusters of malignant cells, as previously described in detail.²⁰ Peripheral blood mononuclear cells were used as a source of control DNA except for samples MF2 and MF18 for which we did not have matching blood and therefore used microdissected epithelial cells from epidermis as the control. NEBNext® Ultra™ II DNA library prep kit for illumina (cat# E7645S) (New England Biolabs, Massachusetts, United States) was used for DNA library preparation and SSELXT Human All exon V6 +UTR probes (Agilent Technologies, California, United State) were used for the exome and UTR sequence capture. The DNA libraries were sequenced on an Illumina HiSeq 1500 sequencer using paired-end (PE) 150 kit (cat# PE-402-4002) (Hiseq PE rapid cluster kit V2) or NovaSeq 6000 S4 reagent kit 300 cycles (cat# 20012866).

4.3.2 Bioinformatic analysis

The raw fastq files generated from WES were processed through the GATK (version 4.0.10) best practices workflow²² and aligned to the hg38 reference genome. Somatic Variants (SV), that include Single Somatic Mutations (SSMs) and indels were identified by two different variant callers, MuTect2 (version 2.1)^{22,23} and Strelka2 (version 2.9.10).²⁴ The variants filtering as “Passed” by both variant callers were used for downstream analysis. The

functional effects of SVs were identified by the Variant Effect Predictor (VEP) (version 95.2).²⁵ The Copy Number Aberrations (CNA) and Tumor Cell Fraction (TCF) were identified using TitanCNA (version 1.20.1).^{25,26} PhyloWGS (version 1.0-rc2) was used for phylogenetic analysis to identify the clones and subclones (**Fig 4.1**).²⁵⁻²⁷

Sequencing data from previous studies (Appendix **Table C2**) in CTCL were obtained from public databases and were subjected to the same bioinformatics analysis as described above, with the exception that only MuTect2 was used for variant calling.

4.4 Results

4.4.1. Genomic landscape of driver genes in MF

We have decided to revisit the genomics of MF because previous studies have largely focused on Sezary syndrome, a rare leukemia-lymphoma syndrome which although related to MF, is considered to be a separate entity. Only 11% of all previously sequenced CTCL cases were MF and the material for sequencing was the entire skin biopsy which might have introduced an error due to the contribution of mutations from other cells than the lymphoma.²⁸ To enrich the material in neoplastic cells, we microdissected clusters of lymphoma cells from 49 skin biopsies from 31 MF patients in various stages (I-IV) (Appendix **Table C1**) for whole exome sequencing (WES) (**Fig 4.1**). For the analysis, we decided to group the samples not only by the clinical stage, but also by the morphological features of the biopsied lesion. In stages I-IIA the lesions as per definition were either patches or plaques (abbreviated for the purpose of this paper as ESP, early-stage plaques) but in stages \geq IIB we have distinguished between the biopsies from tumors (TMR) and the plaques (referred to as the late-stage plaques, LSP).

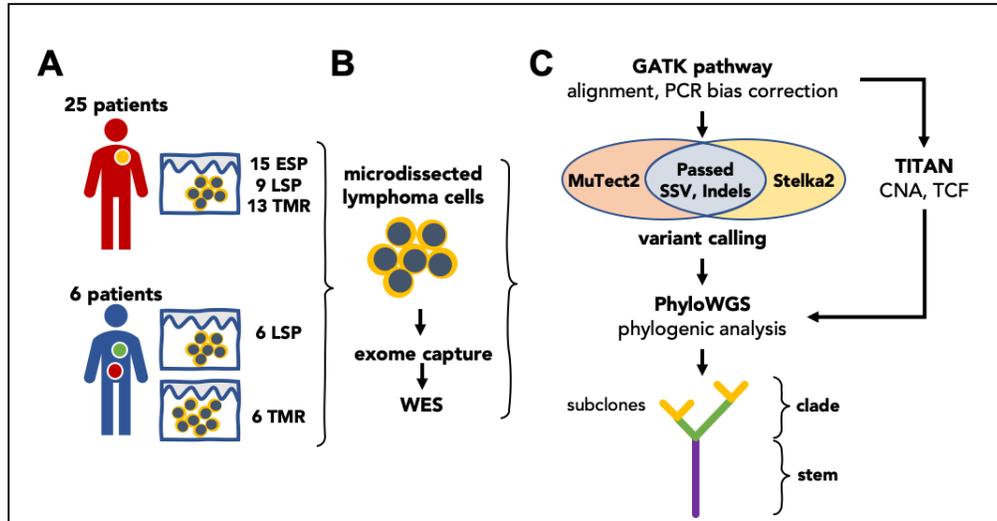


Figure 4.1: Summary of experimental methods and data analysis.

(A) 49 biopsies were obtained from 31 patients with MF. In 6 patients with tumor stage disease paired biopsies from TMR and LSP were obtained. (B) Tumor cell clusters microdissected from the lesional skin and matching control tissue (peripheral blood or the epidermis, not shown) were sequenced by WES. (C) The genetic aberration data (SNV and CNA) was used for the reconstruction of phylogenetic trees of MF. Abbreviations: ESP, early stage plaque, LSP, late stage plaque, TMR, tumor, WES, whole exome sequencing, SV, somatic variants, CNA, copy number aberration, TCF, tumor cell fraction.

We identified a median of 765 non-synonymous SV in ESP, 1269 in LSP and 2133 in TMR (Fig 4.2), documenting that mutations accumulate during disease progression. These numbers are an order of magnitude higher than was previously found in CTCL (median of 42 non-synonymous mutations⁸). The high number of mutations was likely a result of the high tumor cell fraction (TCF) in our material (Appendix Fig C1) as well as the deeper exome sequencing compared with the previously published data. Among the mutations detected in our material, 265 genes were previously adjudicated as driver genes in cancer²⁹ and 56 genes were reported in CTCL.^{4-6,8-13} Current analysis adds 28 additional genes which fulfill the criteria of cancer drivers²⁹ that we found to be mutated in >20% of the patients (Fig 4.2, see Appendix Table

C3 for the complete list of the mutated genes). Among those 28 genes, five genes (*ZFHX3*, *CIC*, *EP300*, *PIK3CB* and *HUWE1*) were found in 33-45% of the samples (Fig 4.2). Most mutated genes previously described as drivers, mapped to the already known pathways such as transcription factors (34 genes) followed by chromatin modification (22 genes). The new mutated pathways found here were the Wnt/B Catenin (4 mutated genes), microRNA processing (*DICER1*), protein homeostasis (*HUWE1*) and genome integrity (*POLE*, *PDS5B*). Lastly, we found that mutations in the tumor suppressor genes dominated over the known oncogenes (Appendix Fig C2), analogously to what have previously been reported in other heterogeneous solid tumors.^{30,31}

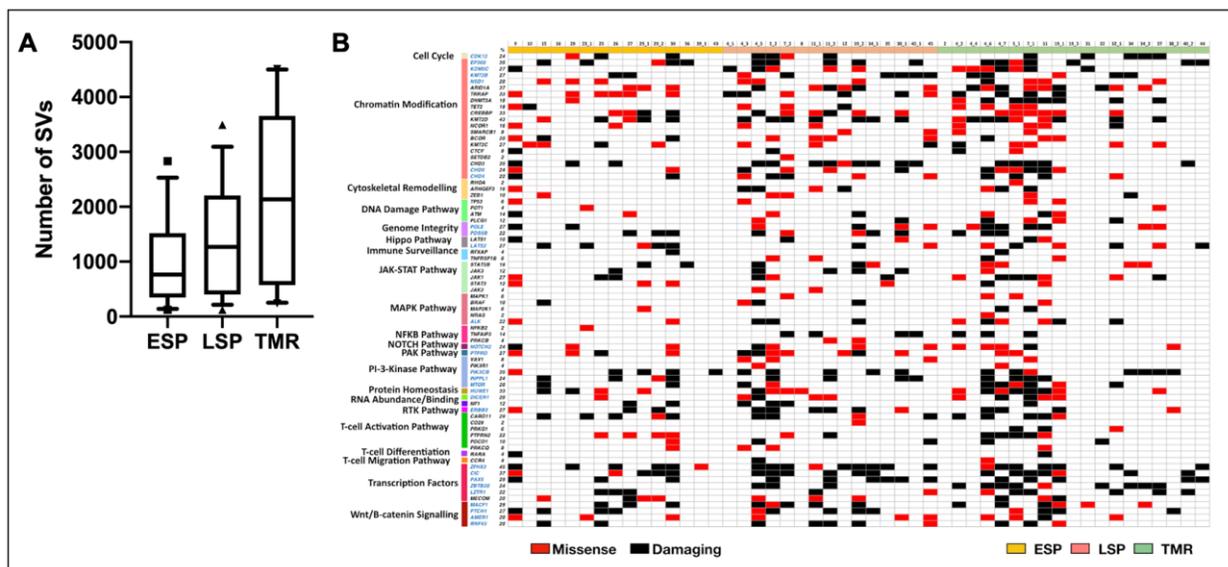


Figure 4.2: Mutational landscape of putative driver genes in MF.

(A) Number of non-synonymous SVs in samples from ESP, LSP and TMR. Box and whisker plot showing 90th percentile respectively. (B) Identification of amino acid altering mutations in 75 putative driver genes across 21 different pathways. Black gene symbols annotate the previously reported 47 driver genes in CTCL; the previously unreported 28 potential drivers identified in this study are highlighted in blue. Damaging mutations indicates frameshift mutations, short read insertion and deletion (<6bp), stop gain or stop lost.

We also investigated the CNA profiles for all our samples (**Fig 4.3A**) and were able to confirm the previously noted amplifications of chromosome 1 and 7 and deletions in chromosome 9 in MF.³² The patterns of changes in the CNA profiles were similar for samples TMR and LSP but different from ESP that surprisingly was characterised by larger CNV fragments and increased number of copy number gains across all chromosomes except 6, 9, 13, 19 and 21 (**Fig 4.3A**). We have also analyzed CNA changes in the putative driver genes. We reproduced the previous finding of Choi et al.^{13,20} of the deletion of *TNFAIP3* and found additional genes that were affected in all investigated samples: deletions in *RHOA* (tumor suppressors) and amplifications in the oncogene *BRAF* (**Fig 4.3B**). In summary, we have found that progression from early stage I to advanced (IIB or higher) was associated with an increased number of non-synonymous SNV, both in the tumors and the plaques. Many of those aberrations affected the potential driver genes and are reported here for the first time.

MF is very rich in mutations, both with respect to SNV and CNA. We have therefore asked whether those mutations are clonal or rather a manifestation of the subclonal architecture of MF.

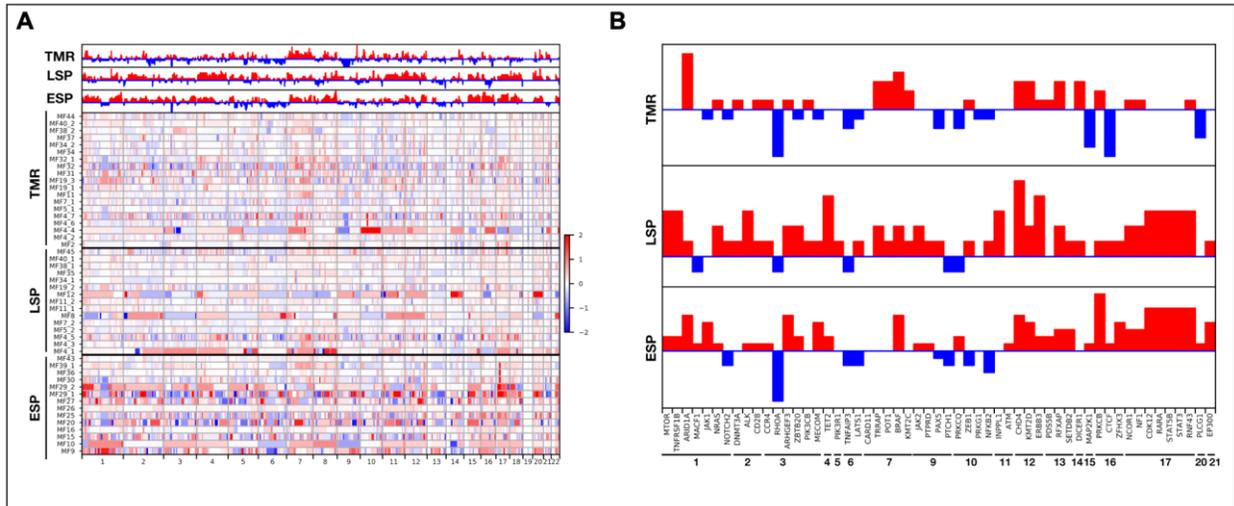


Figure 4.3: Genomic copy number changes in MF.

(A) Integrative genomic viewer (IGV) graph showing the changes in copy number for the 49 MF samples separated by the type of the lesion (ESP, LSP, TMR). The red bar indicates an amplification and the blue bar indicates the deletion. The number of amplifications or deletions are presented as the log₂ scale. The sparklines show combined frequencies of amplifications or deletions at each chromosome. (B) Frequency of deletion or amplification for putative driver genes in each subgroup of MF samples.

4.4.2. Intratumor heterogeneity in MF

Genetic aberrations (SVs and CNAs) in solid tumors have often clonal (present in all cells) or subclonal distribution. The subclones may be present in the common stem (“trunk”) of the phylogenetic tree or may occur as a result of branched evolution. In the latter situation the subclonal mutations present in only a subset of cancer cells, often referred to as the clades, or the branches of the tree (Fig 4.1).^{16,33}

To investigate whether MF is characterized by a subclonal structure, we have used a bioinformatic approach where the combined information from SVs and CNA for each of our samples was used to reconstruct a phylogenetic tree. None of the 49 MF samples analyzed here was clonal. We found the median of 6 subclones with a maximum of 9 clones (Fig 4.4A and Appendix Fig C3). The TMR samples tended to have more subclones than ESP or LSP

and the phylogenetic trees of TMR were also more branched than those of the plaques. The highly branched trees (more than one branching node) were found in 12/19 of TMR (95% CI 40.9%-81.8%) and in 14/27 (52%; 95% CI 33.6-69.8) of the plaques. Only in a minority of cases the phylogenetic tracing showed a linear pattern of subclones (one case of ESP and 2 cases of LSP and TMR) (**Fig 4.4B**, Appendix **Fig C3**).

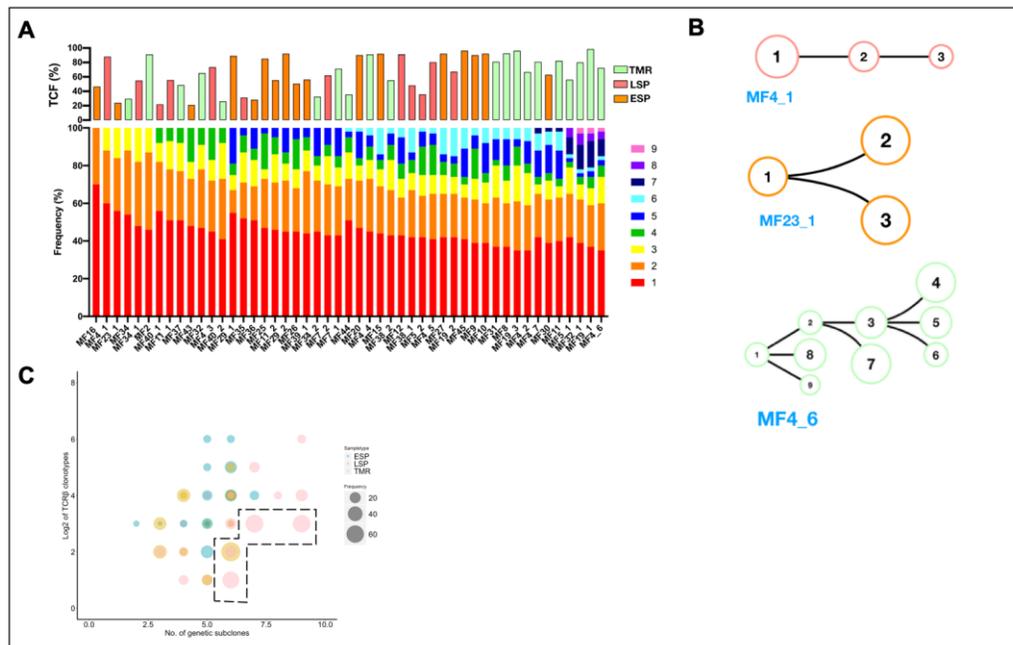


Figure 4.4: Intratumoral heterogeneity in MF.

Combined data from SVs and CNA for each sample was subjected to phylogenetic analysis to identify genetic subclones, as in Fig 1. (A) Rainbow graph representing the number and frequency of the subclones identified in each sample. The samples are arranged by increasing number of subclones. The top bar graph shows TCF for each sample; the colour of the bars indicates the type of the lesion (ESP, LSP, TMR). (B) Examples of three major categories of phylogenetic trees: non-branched linear sequence of subclones (upper), simple branched structure with one generation of subclones (middle) and complex structure with several generations of subclones (lower). All phylogenetic trees are shown in supplementary Figure S3. (C) Bubble plot showing correlation between the number of neoplastic clonotypes and the number of subclones in the samples. The size of the bubble is proportional to the frequency of the first-ranked (the most abundant) clonotype. Dashed line highlights the samples where the first-ranked clonotype had a relative frequency of 60% or higher.

The phylogenetic trees reconstructed from the analysis of the distribution of mutations do not necessarily reflect presence of actual cellular clones, defined as a group of identical cells that share a common ancestry. However, being derived from mature T-cells, MF provides an additional opportunity to analyze clonal composition-by counting the clonotypes, the unique CDR3 sequences of rearranged *TCRB* genes.²¹ Because *TCRB* locus is rearranged on only one allele and not re-rearranged in mature T-cells, it is possible to calculate the richness and diversity of the repertoire of T-cells. To avoid confusion between different definitions of clones, we will refer to the TCR heterogeneity as the “clonotypic”.

There was a weak, but significant correlation between ITH and clonotypic richness (**Fig 4.4C**) and between the respective Simpson diversity indices (Appendix **Fig C4**). This suggests that MF is characterized by a divergent evolution, in which the individual T-cell clones accumulate mutations independently of one another. However, we also noticed that the samples in which the most abundant TCR β clonotype outnumbered the remaining clonotypes (relative frequency $\geq 60\%$) also had multiple mutational subclones (>5) (**Fig 4.4C**). This represented expansion of some neoplastic clones and further branching into multiple mutational subclones (examples of such highly branched phylogenetic trees are MF5_1, MF32_1, MF19_1 and MF4_6 that present with 8-9 subclones, Appendix **Fig C3**).

To further examine whether ITH is present in other types of CTCL, we re-analyzed the sequencing data from 56 samples Sezary syndrome, 13 MF and 8 CTCL not specified, available through data sharing platforms (Appendix **Table C2**). 7 samples of SS did not present any ITH, whereas the remaining cases demonstrated different degree of ITH ranging from 2 to 9 subclones (median of 4 subclones) (Appendix **Fig C5A**). All MF samples showed ITH similar to that found in our material.

4.4.3. Subclonal distribution of mutations in MF

We compared the mutational burden in the stem and clades of phylogenetic trees. The distribution of mutational burden changed with disease progression. In most ESPs (11 of 15, 73%), the majority (>50%) of the mutations were concentrated in stem (**Fig 4.5A**). The situation was reversed in advanced stages. In 73% of LSP and 68% of TMR had >50% mutation in the clades rather than in the stem. The gradual enrichment in the mutations in the clades has been found to be characteristic for divergent evolution³⁴, which underscores our conclusion that this is the dominant evolutionary pattern of MF. In only 5 of 49 samples all mutations were concentrated in the stem and these were also the samples with the lowest number of subclones (2-3 subclones) (**Fig 4.5A**). A slightly higher number of cases with clonal (stem) mutations were found in Sezary syndrome (23%) (Appendix **Fig C5B**), but a direct comparison with our data may be affected by different values of TCF (Appendix **Fig C1**).

distribution between the stem and the clade. *PIK3CB* and *ERBB3* were only found in stem whereas *KMT2C*, *RFXAP* and *TNFAIP3* (essential for chromatin modification and immune surveillance) were only in clades, which as will be discussed below, may have functional importance for the pathogenesis of MF.

4.4.4. Topologic subclonal heterogeneity in MF

We have previously shown that lesions of MF in the same patient have different clonotypic compositions, a phenomenon which we named “topologic heterogeneity”.²⁰ To investigate whether the topologic (interlesional) heterogeneity is also detectable in relation to tumor subclones, we determined the phylogenetic relationships between the subclones in different skin lesions of a single patient. We collected pairs of tumor and plaque biopsies from 6 MF patients and used combined data of SSMs and CNAs from both lesions to map the phylogenetic trees (**Fig 4.6**). Two patients (MF5, MF40) had no common ancestral clone shared between the lesions and 3 other patients (MF4, MF7, MF38) had only a single clone shared between the LSP and TMR. Each of the lesion presented an independent phylogenetic branch with multiple subclones (**Fig 4.6**). We interpret these findings as an additional evidence for divergent evolution of MF witnessing that each lesion evolves in relative isolation from other lesions.

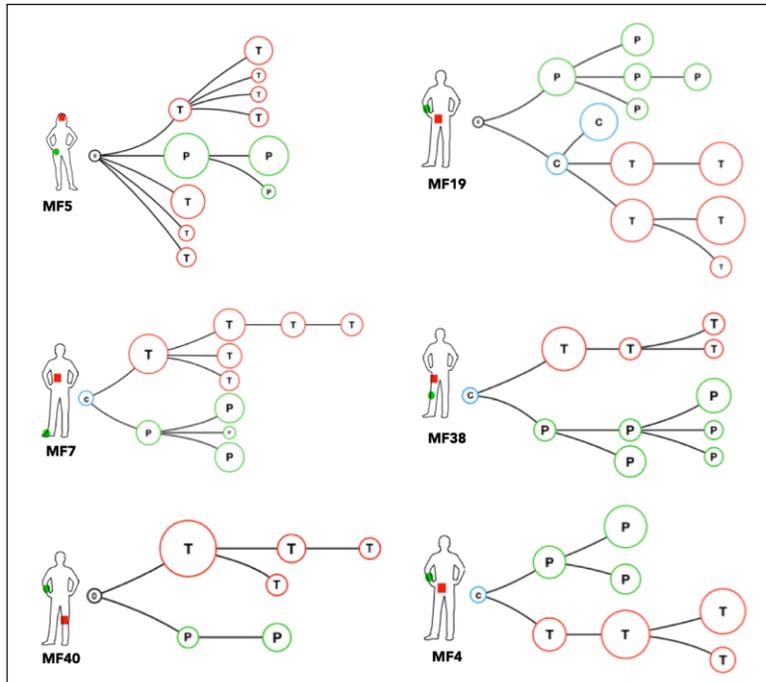


Figure 4.6: Phylogenetic relationship between different lesions in the same patient: topologic heterogeneity in MF.

Pairs of LSP and TMR biopsies were collected from six patients and analyzed as in Fig 4. Red branches represent TMR (T), green branches symbolize the evolution of the LSP (P). The blue circles represent common clones shared by TMR and LSP. The black circle represents a phylogenetic tree without identifiable ancestor clone identifiable.

4.5 Discussion

Intratumor heterogeneity (ITH) refers to the recently described phenomenon of the distribution of somatic mutations in subsets of malignant cells (subclones) rather than being found in all malignant cells (“clonal” mutations). ITH has been documented in solid tumors^{16,35} and in non-Hodgkin lymphomas^{36–38} and is considered to be a genomic manifestation of tumor evolution. ITH arises due to differences in the proliferation and survival between cells bearing different mutations and enhances the evolution by providing

material for natural selection where the fittest subclones determine prognosis and resistance to therapy.¹⁵

Here we demonstrate that MF, previously considered to be a mutationally homogenous lymphoma, is a highly heterogeneous neoplasm composed by multiple subclones. We have found evidence of branched evolution in the majority (92%) of the analyzed MF cases. In addition, subclonal structure was found for the Sezary syndrome, which indicates that ITH is a general feature of CTCL.

Although extensive ITH has been found even in the early stages of MF (T1 in stage IA), disease progression was associated with further accumulation of mutations (SVs and CNA) and increase in ITH. Thus, in contrast to the widely held presumption that progression of MF is caused by a selection and expansion of a single aggressive clone³⁹, our data indicate that it is caused by an evolutionary branching leading to enrichment in neoplastic subclones. Several lines of evidence indicate that progression of MF may occur in absence of strong natural selection but happens via divergent evolution which is neutral or only mildly affected by selective pressure on the subclones. The divergent evolution in ITH suggest that each genetic subclone exhibits diversity from the parent clone due to gain or loss of mutations. In addition to the already mentioned hallmarks of neutral tumor evolution such as branched phylogenetic structure and progressive increase in the number of subclones³⁴, we have also observed a predicted increase in the number of clade mutations versus stem mutations in advanced stages. Neutral divergent evolution also explains our previous finding of why the number of malignant T-cell clonotypes is comparable in early and late stage disease. Malignant T-cell expand and branch into subclones that cohabit the skin niche without evidence of competition and clonal elimination.

We hypothesized that neutral, divergent evolution could explain the resistance of advanced MF to therapy. It has been shown in several types of cancer that ITH is negatively correlated with the sensitivity to chemotherapy or immunotherapy^{15,17} because multiple, genetically diverse malignant subclones provide material for selection of resistant cells. We have explored this question further by analyzing the clonality of putative driver mutations in MF. We added 28 new potential driver mutations to the list of known mutated driver genes in CTCL now including important targetable genes such as *JAK1*, *JAK3*, *BRAF*, *ALK*, *MTOR*, or *PTCH1*.

Unfortunately, we did not find any consistent pattern in the clonal driver mutations which were very heterogeneous and varied vastly from patient to patient. However, we found that certain pathways, in particular chromatin modification and transcription factors were very frequently mutated in the phylogenetic stem in at least one constituent gene in most samples. Based on this observation pathway targeting could be a more promising therapeutic strategy in MF as compared to targeting of specific mutations. Already the drugs that affect chromatin modification mechanisms, such as histone deacetylase blockers, have proved efficacy in MF.^{40,41}

Most potential driver mutations in advanced disease are confined to the subclones (clades) which is likely to limit the efficacy of targeted treatment. However, it is worthwhile to mention here that even subclonal mutations may present with drug targeting opportunities if the given subclone is important for the progression of the entire tumor. Cooperativity between distinct subclones was described for some metastasizing tumors⁴²⁻⁴⁴, and it was proposed that disruption of clonal cooperation might be an interesting therapeutic approach. Especially the targeting of Wnt and Hedgehog seems to be promising⁴², the pathway which we show here to be frequently mutated in MF (genes *MACF1*, *PTCH1*, *RNF43*).

Perhaps a better understanding of ITH and the impact of therapy on the subclonal evolution of MF could be gained using multi-omics single cell sequencing of the samples collected before and after the treatment. The subclones detected by bioinformatic reconstruction of phylogenetic trees do not necessarily identify true cellular clones, defined as a group of mutationally identical cells derived from a common ancestor. It has recently been discovered by us and other groups that CTCL is heterogenous at transcriptomics and protein level.⁴⁵⁻⁴⁸ Single cell approach would allow for precise mapping of malignant T-cell clones identified by a unique pair of rearranged *TCRA* and *TCRB* sequences to the mutational pattern and gene expression profiles. However, our results identified one potential problem in studies on ITH in multifocal tumor such as MF. By comparing ITH between samples taken from different lesions we found very little phylogenetic relationship between the subclones. In 2/6 cases, the distant lesions did not share any parental clone. Lack of relatedness between discrete lesions could be a result of sampling and missing out the lesion that provides a phylogenetic link. This may indicate that ITH of the totality of MF exceeds significantly the ITH found in single biopsies. A similar conclusion was reached for ITH of lung cancer by mutational analysis of multiple biopsies.³⁵ Driver mutations that were clonal in some biopsies were found to be subclonal in other areas of the tumor. We recognize lack of data from repeated biopsies of multiple skin lesions as a limitation of this study.

Another limitation of our study is the lack of longitudinal follow-up on ITH. We have noted that some phylogenetic trees did not have an identifiable common clone. It is likely that such ancestor clones may become extinct during the progression of the tumor and are no longer detectable, a phenomenon which was described during the evolution of another highly heterogeneous tumor, the glioblastoma.⁴⁴

Finally, we would like to propose a model for the pathogenesis of MF that accounts for the previously found clonotypic heterogeneity^{19–21} and the ITH described here (**Fig 4.7**). ITH is readily detectable in ESP, which testifies to the history of mutational tumor evolution before seeding of neoplastic cells in the skin. As the disease progresses, the seeded T-cell clones undergo further mutations and branching into subsequent generations of subclones. Analysis of the branching structure could also confirm our hypothesis that malignant clones from one lesion can re-enter the circulation and seed another lesion. Exchange of malignant T-cell clones between lesions could explain why there is more resemblance between LSP and TMR than between LSP and ESP. A more direct evidence of cancer self-seeding was found in patient MF19 (**Fig 4.6**) where a subclone shared between the plaque and the tumor was interjected among the branches of the phylogenetic tree.

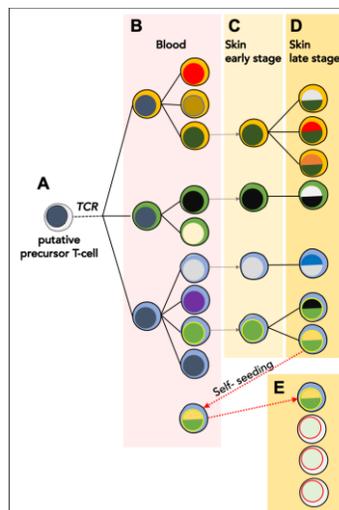


Figure 4.7: Proposed model of the evolution of MF.

The skin lesions of MF are formed by seeding with the circulating malignant T-cell clones which undergo further mutational evolution. It is likely malignant clones originate from an immature T-cell transformed before TCRB rearrangement (A) and therefore show clonotypic heterogeneity (highlighted by different colors of the “cytoplasm”). These circulating neoplastic T-cells undergo expansion and accumulate mutations leading to emergence of

genetically different malignant subclones (different colors of the “nucleus”) **(B)**. Some of the circulating malignant cells seed into the skin (stippled grey arrows) **(C)** where they proliferate, accumulate further mutations and develop additional subclones as the disease progresses **(D)**. Some subclones may re-enter the circulation and seed other skin lesions (red stippled arrow) further increasing the heterogeneity of the lesions and causing disease progression **(E)**. Solid lines symbolize the phylogenetic relationship between the generations of malignant cells that follow the pattern of divergent, neutral evolution. Based on data in this paper and our previous work.^{19–21,49}

4.6 References

1. Rizvi MA. T-cell non-Hodgkin lymphoma. *Blood*. 2006;107(4):1255–1264.
2. Ansell SM. Non-Hodgkin Lymphoma: Diagnosis and Treatment. *Mayo Clinic Proceedings*. 2015;90(8):1152–1163.
3. Agar NS, Wedgworth E, Crichton S, et al. Survival outcomes and prognostic factors in mycosis fungoides/Sézary syndrome: validation of the revised International Society for Cutaneous Lymphomas/European Organisation for Research and Treatment of Cancer staging proposal. *J. Clin. Oncol*. 2010;28(31):4730–4739.
4. McGirt LY, Jia P, Baerenwald DA, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. *Blood*. 2015;126(4):508–519.
5. Ungewickell A, Bhaduri A, Rios E, et al. Genomic analysis of mycosis fungoides and Sézary syndrome identifies recurrent alterations in TNFR2. *Nat. Genet*. 2015;47(9):1056–1060.
6. Wang L, Ni X, Covington KR, et al. Genomic profiling of Sézary syndrome identifies alterations of key T cell signaling and differentiation genes. *Nat. Genet*. 2015;47(12):1426–1434.
7. Moffitt AB, Dave SS. Clinical Applications of the Genomic Landscape of Aggressive Non-Hodgkin Lymphoma. *J. Clin. Oncol*. 2017;35(9):955–962.
8. Park J, Yang J, Wenzel AT, et al. Genomic analysis of 220 CTCLs identifies a novel recurrent gain-of-function alteration in RLTPR (p.Q575E). *Blood*. 2017;130(12):1430 LP–1440.
9. da Silva Almeida AC, Abate F, Khiabani H, et al. The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome. *Nat. Genet*. 2015;47(12):1465–1470.

10. Tensen CP. PLCG1 Gene Mutations in Cutaneous T-Cell Lymphomas Revisited. *J. Invest. Dermatol.* 2015;135(9):2153–2154.
11. Prasad A, Rabionet R, Espinet B, et al. Identification of Gene Mutations and Fusion Genes in Patients with Sézary Syndrome. *J. Invest. Dermatol.* 2016;136(7):1490–1499.
12. Kiel MJ, Sahasrabudde AA, Rolland DCM, et al. Genomic analyses reveal recurrent mutations in epigenetic modifiers and the JAK–STAT pathway in Sézary syndrome. *Nature Communications.* 2015;6(1.):
13. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. *Nature Genetics.* 2015;47(9):1011–1019.
14. Chang L-W, Patrone CC, Yang W, et al. An Integrated Data Resource for Genomic Analysis of Cutaneous T-Cell Lymphoma. *J. Invest. Dermatol.* 2018;138(12):2681–2683.
15. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature.* 2013;499(7457):214–218.
16. McGranahan N, Rosenthal R, Hiley CT, et al. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell.* 2017;171(6):1259–1271.e11.
17. Keenan TE, Burke KP, Van Allen EM. Genomic correlates of response to immune checkpoint blockade. *Nat. Med.* 2019;25(3):389–402.
18. Campbell JJ, Clark RA, Watanabe R, Kupper TS. Sezary syndrome and mycosis fungoides arise from distinct T-cell subsets: a biologic rationale for their distinct clinical behaviors. *Blood.* 2010;116(5):767–771.
19. Hamrouni A, Fogh H, Zak Z, Ødum N, Gniadecki R. Clonotypic Diversity of the T-cell Receptor Corroborates the Immature Precursor Origin of Cutaneous T-cell Lymphoma. *Clin. Cancer Res.* 2019;25(10):3104–3114.

20. Iyer A, Hennessey D, O’Keefe S, et al. Skin colonization by circulating neoplastic clones in cutaneous T-cell lymphoma. *Blood*. 2019;
21. Iyer A, Hennessey D, O’Keefe S, et al. Clonotypic heterogeneity in cutaneous T-cell lymphoma (mycosis fungoides) revealed by comprehensive whole-exome sequencing. *Blood Adv*. 2019;3(7):1175–1184.
22. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics*. 2013;43:11.10.1–33.
23. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol*. 2013;31:213.
24. Kim S, Scheffler K, Halpern AL, et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat. Methods*. 2018;15(8):591–594.
25. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122.
26. Ha G, Roth A, Khattra J, et al. TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. *Genome Res*. 2014;24(11):1881–1893.
27. Deshwar AG, Vembu S, Yung CK, et al. PhyloWGS: Reconstructing subclonal composition and evolution from whole-genome sequencing of tumors. *Genome Biol*. 2015;16(1):35.
28. Martincorena I, Roshan A, Gerstung M, et al. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science*. 2015;348(6237):880–886.

29. Bailey MH, Tokheim C, Porta-Pardo E, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell*. 2018;174(4):1034–1035.
30. Zhu K, Liu Q, Zhou Y, et al. Oncogenes and tumor suppressor genes: comparative genomics and network perspectives. *BMC Genomics*. 2015;16 Suppl 7:S8.
31. Osborne C, Wilson P, Tripathy D. Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications. *Oncologist*. 2004;9(4):361–377.
32. Gug G, Huang Q, Chiticariu E, Solovan C, Baudis M. DNA copy number imbalances in primary cutaneous lymphomas. *Journal of the European Academy of Dermatology and Venereology*. 2019;33(6):1062–1075.
33. De Mattos-Arruda L, Sammut S-J, Ross EM, et al. The Genomic and Immune Landscapes of Lethal Metastatic Breast Cancer. *Cell Rep*. 2019;27(9):2690–2708.e10.
34. Williams MJ, Werner B, Barnes CP, Graham TA, Sottoriva A. Identification of neutral tumor evolution across cancer types. *Nat. Genet*. 2016;48(3):238–244.
35. Jamal-Hanjani M, Wilson GA, McGranahan N, et al. Tracking the Evolution of Non-Small-Cell Lung Cancer. *N. Engl. J. Med*. 2017;376(22):2109–2121.
36. Araf S, Wang J, Korfi K, et al. Genomic profiling reveals spatial intra-tumor heterogeneity in follicular lymphoma. *Leukemia*. 2018;32(5):1261–1265.
37. Magnoli F, Tibiletti MG, Uccella S. Unraveling Tumor Heterogeneity in an Apparently Monolithic Disease: BCL2 and Other Players in the Genetic Landscape of Nodal Follicular Lymphoma. *Front. Med*. 2019;6:44.
38. Beà S, Valdés-Mas R, Navarro A, et al. Landscape of somatic mutations and clonal evolution in mantle cell lymphoma. *Proc. Natl. Acad. Sci. U. S. A*. 2013;110(45):18250–18255.

39. de Masson A, O'Malley JT, Elco CP, et al. High-throughput sequencing of the T cell receptor β gene identifies aggressive early-stage mycosis fungoides. *Sci. Transl. Med.* 2018;10(440.):
40. Oka T, Miyagaki T. Novel and Future Therapeutic Drugs for Advanced Mycosis Fungoides and Sézary Syndrome. *Front. Med.* 2019;6:116.
41. Lopez AT, Bates S, Geskin L. Current Status of HDAC Inhibitors in Cutaneous T-cell Lymphoma. *Am. J. Clin. Dermatol.* 2018;19(6):805–819.
42. Zhou H, Neelakantan D, Ford HL. Clonal cooperativity in heterogenous cancers. *Semin. Cell Dev. Biol.* 2017;64:79–89.
43. Caswell DR, Swanton C. The role of tumour heterogeneity and clonal cooperativity in metastasis, immune evasion and clinical outcome. *BMC Med.* 2017;15(1):133.
44. Vinci M, Burford A, Molinari V, et al. Functional diversity and cooperativity between subclonal populations of pediatric glioblastoma and diffuse intrinsic pontine glioma cells. *Nat. Med.* 2018;24(8):1204–1215.
45. Buus TB, Willerslev-Olsen A, Fredholm S, et al. Single-cell heterogeneity in Sézary syndrome. *Blood Adv.* 2018;2(16):2115–2126.
46. Borcherding N, Voigt AP, Liu V, et al. Single-Cell Profiling of Cutaneous T-Cell Lymphoma Reveals Underlying Heterogeneity Associated with Disease Progression. *Clin. Cancer Res.* 2019;25(10):2996–3005.
47. Litvinov IV, Tetzlaff MT, Thibault P, et al. Gene expression analysis in Cutaneous T-Cell Lymphomas (CTCL) highlights disease heterogeneity and potential diagnostic and prognostic indicators. *Oncoimmunology.* 2017;6(5):e1306618.

48. Gaydosik AM, Tabib T, Geskin LJ, et al. Single-Cell Lymphocyte Heterogeneity in Advanced Cutaneous T-cell Lymphoma Skin Tumors. *Clin. Cancer Res.* 2019;25(14):4443–4454.
49. Gniadecki R, Lukowsky A, Rossen K, et al. Bone marrow precursor of extranodal T-cell lymphoma. *Blood.* 2003;102(10):3797–3799.

Chapter 5: Independent evolution of cutaneous lymphoma subclones in different microenvironments of the skin

5.1 Abstract

Mycosis fungoides (MF) is the most common cutaneous T-cell lymphoma. Lesions of MF are formed by hematogenous seeding the skin with polyclonal (clonotypically diverse) neoplastic T-cells which accumulate numerous mutations and display a high degree of mutational, intratumoral heterogeneity (ITH). A characteristic but poorly studied feature of MF is epidermotropism, the tendency to infiltrate skin epithelial layer (epidermis) in addition to the vascularized dermis. By sequencing the exomes of the microdissected clusters of lymphoma cells from the epidermis and the dermis, we found that those microenvironments comprised different malignant clonotypes. Subclonal structure witnessed to the independent mutational evolution in the epidermis and dermis. Malignant cells in epidermal and dermal niches exhibited different patterns of driver gene mutations. Thus, the epidermal involvement in MF could not be explained by gradual infiltration from the dermis but was caused by a separate seeding process. Independent invasion of different microenvironments in the skin allowed for a higher degree of ITH because cancer subclones developed in each niche via a neutral, branched evolution. In conclusion, tissue microenvironments shape the subclonal architecture in MF leading to “ecological heterogeneity” which contributes to the total ITH. Since ITH adversely affects cancer prognosis, targeting the microenvironment may present therapeutic opportunities in MF and other cancers.

5.2 Background

Mycosis fungoides (MF) is one of the most common diseases in the realm of extranodal T-cell lymphomas.¹ It is a skin-tropic lymphoid neoplasm that initially presents as scaly, erythematous patches and plaques, which may progress to tumours and disseminate to lymph nodes and other organs.²⁻⁴

ITH has recently emerged as an important characteristic of solid and hematopoietic malignancies.⁵ Although mutations in few driver genes may be sufficient to initiate tumorigenesis, it is now evident that the progression depends on the accumulation of multiple mutations to promote expansion and invasion of the primary niche and surrounding tissues.^{6,7} Mutations occur randomly in malignant cells within the tumour, leading to the emergence of multiple subclones. ITH allows cancer to withstand selection pressure from the microenvironment and therapies by promoting the expansion of subclones harboring mutations advantageous to these cells.⁶

The generation of ITH is usually viewed as an evolutionary process with a single transformed cell as a starting point. This cell proliferates and branches into phylogenetically related subclones (Appendix **Fig D1**) that infiltrate the tissue. Although this model may be applicable to cancers that grow expansively as single tumours, this is not necessarily true for all malignancies. Many cancers comprise the entire ecosystem of primary and metastatic lesions that are physically separated from each other. It has been shown that in such situations, tumour heterogeneity may be augmented by cross-seeding by circulating, genetically diverse

cancer subclones, for example, cancer self-seeding by the cells from the metastatic lesion re-entering the primary tumour.⁸ We hypothesized that a similar mechanism may operate at the microscopic scale for primary cancers, where different compartments within an organ can be colonized by different cancer subclones. Independent seeding of different microscopic compartments within the same organ would increase the heterogeneity of the entire lesion beyond what would have been possible by a continuous evolution from only one ancestral clone (Appendix **Fig D1**).

MF provides with a convenient model to test this hypothesis. The skin has a simple layered structure comprising ectodermal derived epidermis and the mesodermal dermis. Both layers can be occupied by cancer cells in MF. The histopathology of MF reveals disconnected areas of malignant cell clusters in the dermis and the epidermis. Dermal infiltrate is usually perivascular or diffuse whereas lymphoma foci in the epidermis form well-demarcated clusters of cells known as Pautrier abscesses (Appendix **Fig D2**). Pautrier microabscesses are a characteristic feature of MF and are present in approximately 20% of all biopsies.^{9,10} Unlike the dermal perivascular infiltrates that comprise a significant proportion of reactive cells, Pautrier abscesses are believed to contain predominantly cancer cells with a minor admixture of apoptotic Langerhans cells and eosinophils.¹¹ The initial points of entry of malignant cells are the capillaries in the upper (papillary) dermis. Therefore, Pautrier microabscesses are a manifestation of the infiltrative growth in MF by which “epidermotropic” subclones migrate from the dermis to the epidermis.

We and others have recently studied the heterogeneity of MF on the genomic, transcriptomic

and cellular levels.¹²⁻¹⁷ In contrast to previous views considering MF as a relatively simple, monoclonal lymphoproliferation derived from a mature T-cell, we showed that MF comprises multiple mature T-cell clones which undergo branched evolution producing generations of cancer subclones.^{16,17} Interestingly, there seems to be very little competition between different subclones and the disease progression is associated with an increase in subclonal diversity rather than a selection of the fittest subclones. We have therefore asked whether different microcompartments in the skin (epidermis vs dermis) play a role in the generation of ITH in MF. We found that Pautrier microabscesses do not comprise a subpopulation of the dermal malignant cells emigrating to the epidermis, but that they originate independently from distinct seeding event and undergo autonomous evolution, reminiscent of the peripatric type of ecological speciation of the organisms.

5.3 Material and Methods

Sample collection, cryosectioning, laser capture microdissection (LCM) and sample preparation for whole-exome sequencing (WES)

Samples (4mm punch biopsy and 10ml of blood) were obtained from 7 patients under Ethics approval number HREBA.CC-16-0820-REN1 approved by Health Research Ethics Board of Alberta, Cancer Committee. Peripheral blood mononuclear cells (PBMC) were used as normal control except in sample MF18 where the epidermal cells were used as normal control for data analysis. Frozen biopsies were sectioned at 10 μ m, transferred on 2 μ m PEN membrane slides and stained with hematoxylin and eosin. Clusters of atypical cells representing malignant lymphocytes were microdissected from the dermis and the epidermis under 20x or 40x magnification in Leica DM6000B microscope (Wetzlar, Germany). The

microdissected epidermal lymphocytes represented Pautrier microabscesses which could readily be identified based on their enlarged hyperchromatic nuclei, lighter cytoplasm and a cleavage separating them from the surrounding epidermis (Appendix **Fig D2**). Sequencing libraries were prepared with NEBNext® Ultra™ II kit for Illumina (cat# E7645S) (New England Biolabs, MA) and exomes were captured with SSELXT Human All exon V6 +UTR probes (Agilent Technologies, CA). Samples were sequenced on Illumina HiSeq 1500 sequencer or NovaSeq 6000 platform. Detailed protocol for samples processing for storage and sequencing explained in previous methods.¹⁵

Data analysis

To identify the TCR sequences, the fastq files were analyzed using MiXCR (version 2.10.0).¹⁸ To identify the genomic subclones, the sequenced reads were processed using the GATK4 (version 4.0.10). Somatic variants (SVs) were identified by MuTect2 (version 2.1)^{19,20} and Strelka2 (version 2.9.10).²¹ Variants filtered as “Pass” from both variant callers were used for downstream analysis. Variant effect predictor (VEP, version 95.2) was used to assign functional significance to the predicted SVs.²² Titan-CNA (version 1.20.1) was used to identify copy number aberrations (CNA) and predict the tumour cell fraction (TCF).^{22,23} PhyloWGS (version 1.0-rc2) was used for phylogenetic analysis of the genetic subclones.^{22–24}

5.4 Results

Clonotypic diversity of malignant T-cells in epidermal and dermal niches in the skin

Since the epidermis is not vascularized, the intraepidermal neoplastic cells of Pautrier microabscesses must necessarily originate from the cells that initially enter the papillary

dermis. Therefore, Pautrier microabscesses are assumed to represent a fraction of the dermal cells that acquired an ability to survive and proliferate in the epidermis. To examine this hypothesis, we microdissected atypical cells from both layers (epidermal and dermal) of the skin in 7 MF patients (Appendix **Table D1**) and analysed their clonotypic composition by comparing their T-cell receptor β (TCR β) repertoires. We used the previously described methodology where the CDR3 sequences of the rearranged *TCRB* genes are detected by bioinformatic analysis of WES data.¹⁵ Since *TCRB* locus is rearranged only on one chromosome (allelic exclusion) at the stage of the double-positive thymocyte, the unique CDR3 sequences constitute a molecular barcode identifying a single clone of the T-cell.²⁵ Unlike the mutational heterogeneity which is constantly changing during tumour evolution by hypermutation of cancer genomes, the clonotypic heterogeneity can only be generated at the level of pre-malignant precursor T-cell, because the essential recombinases RAG1/2 that mediate V, D, J rearrangements are neither expressed in mature T-cells nor in tumour lymphocytes of MF.

We identified malignant TCR β clonotypes by matching their frequency to TCF of the samples and noticed that epidermal samples had a higher TCR β clonotype diversity in comparison to the dermis (median of 25 clonotypes (range 1-70) versus 11 clonotypes (range 7-49), respectively) (**Fig 1A, B**). However, the number of shared clonotypes between epidermis and dermis was very low, from no shared clonotypes (sample MF17), 1 shared clonotype (MF41) to a maximum of 2-5 clonotypes (MF18, MF22, MF23, MF28 and MF42) (**Fig 1C**). These results indicated that the pools of malignant T-cells in the epidermis and dermis are largely clonotypically unrelated that suggested that they originate from separate seeding events by

different T-cell clones. Indeed, we observed that in 4 of 6 samples analyzed, T-cells from the epidermal and dermal compartments individually shared between 1-8 TCR β clonotypes with those in the circulating blood (**Fig 1D**), which represented a higher degree of overlap than seen for the epidermal and dermal compartments. Taken together, the epidermal and dermal compartments of the skin are likely to be seeded by different circulating malignant clones.

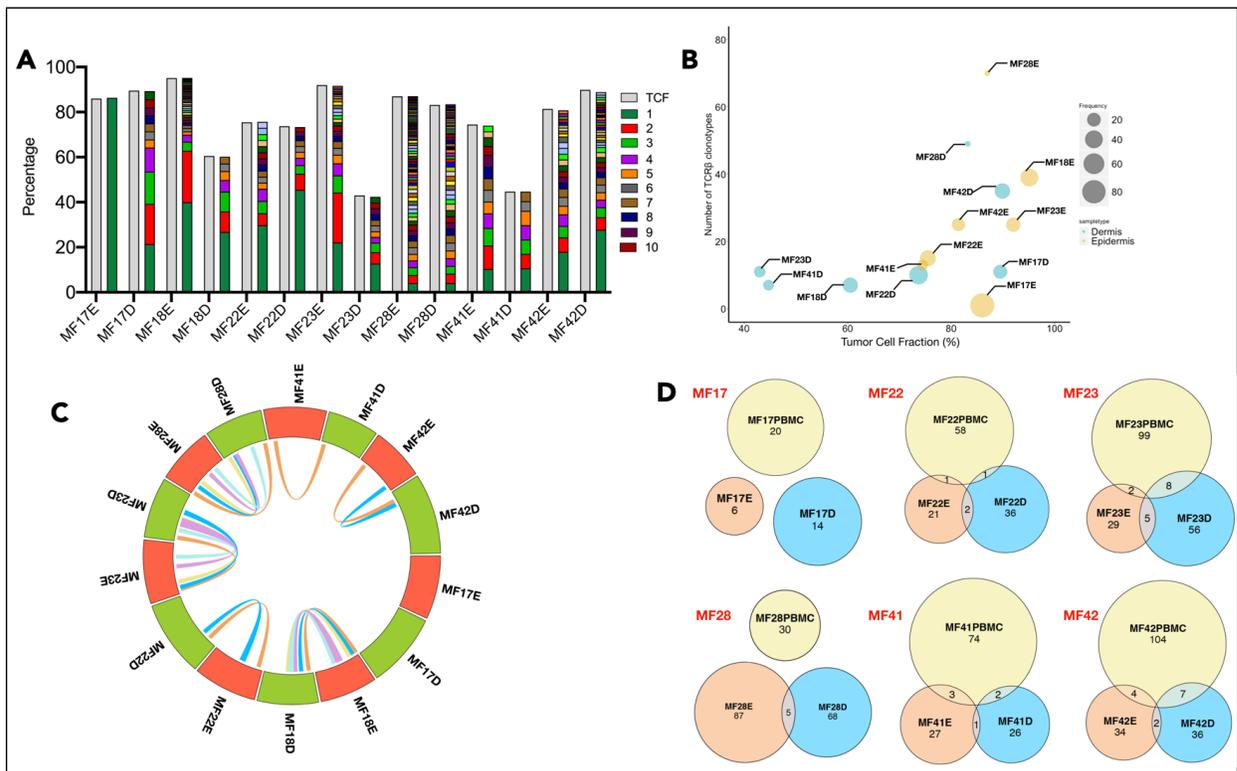


Figure 5.1: Clonotypic heterogeneity and tumour cell seeding of the skin microenvironment in MF.

Percentage of Tumor cell fraction (TCF) and relative frequency of TCR β clonotype sequences for cells isolated from different skin layer (epidermis and dermis) was calculated and plotted as a bar graph (**A**). The green and brown colour indicate the first and the 10th most frequent TCR β clonotype in the sample. Gray colour indicates the tumour cell fraction (TCF). (**B**) Bubble plot presenting the correlation between TCF and the number of neoplastic TCR β clonotypes in cells from epidermis and dermis of each sample. The size of the bubble is equivalent to the relative frequency of the most frequent TCR β clonotype in the sample. (**C**) Circos plot indicates the frequency of TCR β clonotype for cells isolated from epidermis and dermis of each sample. The connecting lines inside indicate the number of overlapping TCR β

clonotype between the two regions of the same sample. E- Epidermis; D-Dermis. **(D)** Venn diagram indicating the number of identical TCR β clonotype between the epidermis, dermis and the circulating blood in samples MF17, MF22, MF23, MF28, MF41 and MF42.

Mutational diversity in neoplastic T-cells in epidermis and dermis

The substantial clonotypic discordance between the epidermal and dermal compartments prompted a question regarding differences and similarities in their mutational evolution. In our previous work, we characterized 75 putative driver mutations involved in the pathogenesis and progression of MF.¹⁷ Similarities in the patterns of driver mutations between the epidermal and the dermal infiltrate would suggest parallel evolution in both compartments whereas lack of substantial overlap would indicate a neutral evolution.

We identified a median of 856 non-synonymous mutations in cells from the dermal region and 1431 non-synonymous mutations in cells from the epidermal region (**Fig 2A**). The majority of the mutations (48-93%) were in the Pautrier microabscess fraction and the overlap between the compartments was less than 7% across all 7 samples (**Fig 2B**). When driver genes were considered, 37 drivers were mutated in both epidermis and dermis, 13 genes (*NCOR1*, *ARHGEF3*, *ZEB1*, *TP53*, *PLCG1*, *RFXAP*, *CD58*, *TNFRSF1B*, *JAK3*, *MAPK1*, *PRKCB*, *MTOR* and *NF1*) were exclusively mutated in epidermis and 9 genes (*DNMT3A*, *TET2*, *SMARCB1*, *KDM6A*, *SETDB2*, *STAT3*, *NFKB2*, *NOTCH2* and *CARD11*) were mutated only in dermis (**Fig 2C**). The mutations present only in the malignant T-cells of epidermis were in the genes involved in cytoskeletal remodelling, DNA damage and immune surveillance. However, the driver mutation profile in the epidermal and dermal fractions showed non-overlapping patterns arguing against parallel evolution. Thus, the data supported the model of independent mutational evolution of neoplastic cells in different skin microenvironments.

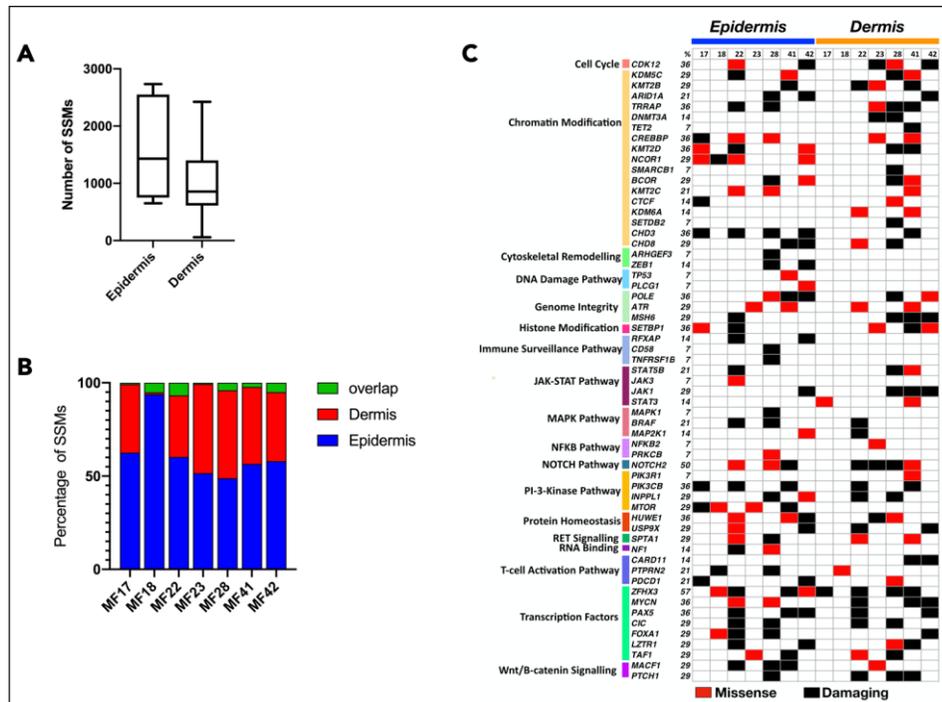


Figure 5.2: Mutational landscape of putative driver genes in anatomical layers of skin.

Neoplastic T-cells isolated from epidermis and dermis were analyzed for somatic variants (SVs) in putative driver genes. (A) Number of non-synonymous SVs in neoplastic cells isolated from epidermis and dermis. Box and whisker plot showing 90th percentile respectively. (B) Bar graph represents the number of SSMs identified and the percent overlapping mutations between epidermis and dermis. (C) Mutations in 59 genes across 18 different pathways were identified. The mutations were classified as missense or damaging. Frameshift, insertion or deletion (<6bp), stop gain or lost are classified as damaging as these mutations are likely to be deleterious.

Phylogenetic development of tumor T-cells in skin microenvironment

To further examine the phylogenetic relationships between the subclones in the epidermal and dermal compartments we adopted the previously described bioinformatic approach based on the analysis of the mutational pattern between cancer cells.¹⁷ We found evidence of subclonal heterogeneity in all samples confirming previous findings of ITH in MF.¹⁷ A slightly higher number of subclones were found in epidermal (5-8 subclones) versus the dermal layers (4-5

subclones) (**Fig 3A**) reflecting the differences in clonotypic richness between those compartments (**Fig 3B**). In the epidermal fraction, the mutational burden was mostly in the clades whereas the dermal fraction tended to have a higher proportion of clonal (stem) mutations (**Fig 3C**). Thus, the number of subclones correlated with the proportion of subclonal mutations, as predicted for the neutral, branched evolution pattern.²⁶ We also analyzed driver gene mutations in the stem and clade population for the dermis and epidermis and found that mutations in *STAT5B* and *CDK12* were only present in clades in Pautrier microabscesses whereas *NOTCH2* and *PRKCB* mutations were only in the dermal fraction, either in the stem or clades (**Fig 3D**).

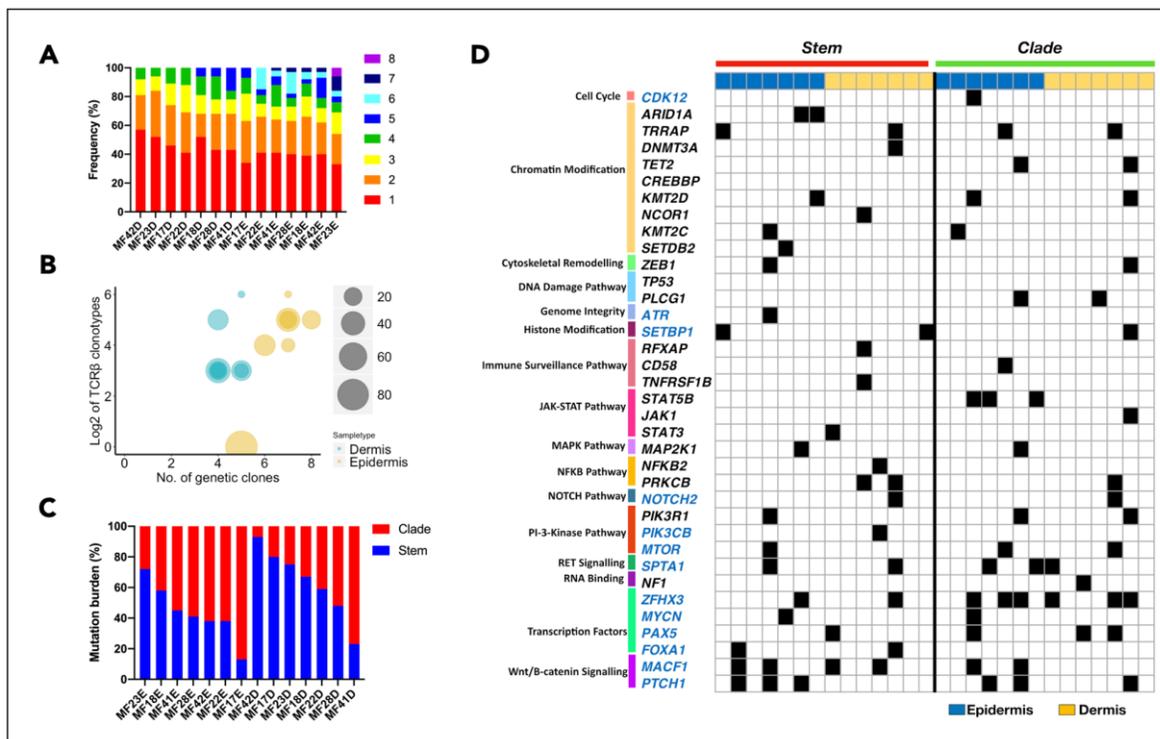


Figure 5.3: Evolutionary facets of the genetic clones in skin microenvironment.

Combined data from SVs and CNA for each sample was subjected to phylogenetic analysis to identify genetic subclones. (**A**) Rainbow graph representing the number and proportion of the subclones identified in each sample. (**B**) Bubble plot representing the correlation between the TCRβ clonotypes and the genetic subclones. The number of TCRβ clonotypes are represented

as Log2 scale. (C) Phylogenetic trees are composed of stem and clades (also recognized as branches). Bar graph represents the percentage of all mutations in each section (stem and clade) of the phylogenetic tree. The blue and red colour represents the mutations in stem and clades respectively (D) Mutational landscape of the putative driver genes in the different sections of the phylogenetic tree for two layers of skin (epidermis and dermis). Function significance of the mutations include missense, frameshift, insertions, deletions, stop gain or loss and variant in 3' and 5' UTR. No colour indicates absence of mutation in the sample.

To visualize how different subclones in the clades are related to each other, we reconstructed the phylogenetic trees. In one case (MF41) there was no common ancestor clone linking epidermal and dermal subclones. In other cases, we detected 1-2 subclones forming the stem of the tree. All samples showed branched evolution of the subclones, with the epidermal and dermal clades clearly separated from each other (Fig 4).

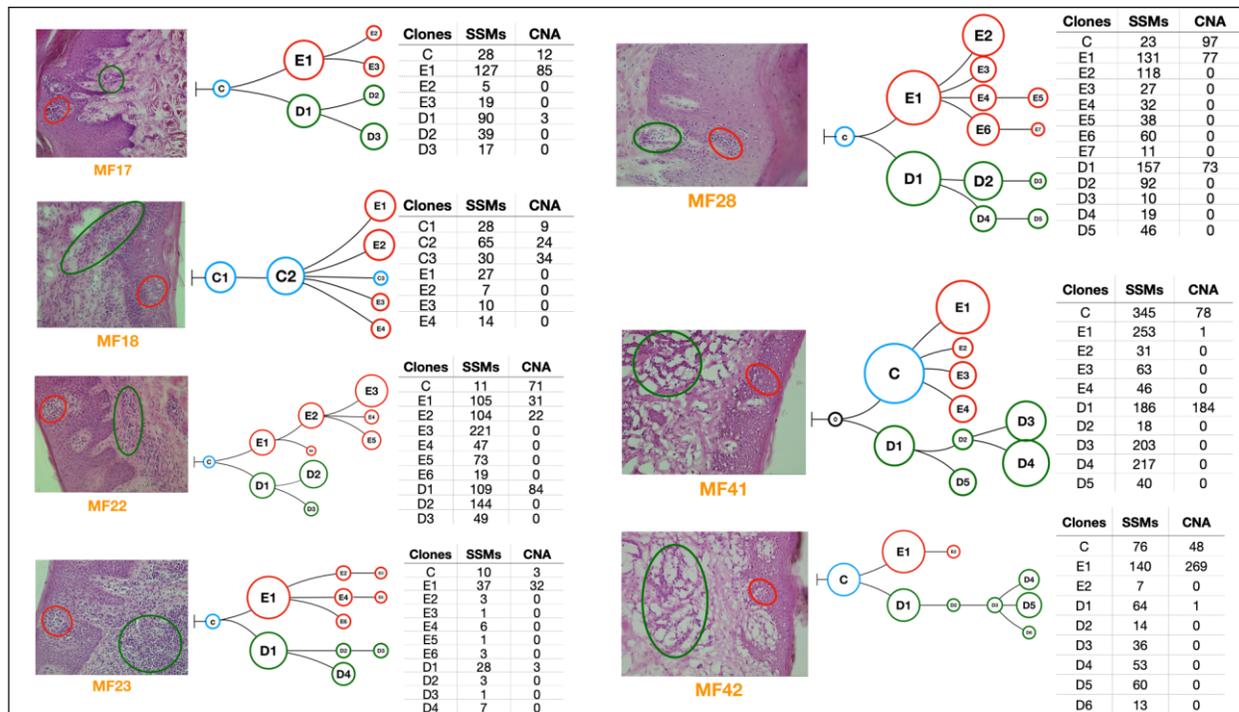


Figure 5.4: Phylogenetic analysis of the neoplastic T-cells in skin microenvironment. Genetic abnormalities (SVs and CNA) for neoplastic cells microdissected from epidermis and dermis were subjected to phylogenetic analysis. Each phylogenetic tree represents an individual patient sample. The blue circles indicate the common clone between the two skin layers. Red and green indicate the subclones in epidermis and dermis respectively. Black

circles indicate absence of common ancestral clone. The tables adjacent to each figure provides the number of SVs and CNA identified in each of the subclones in the phylogenetic tree.

5.5 Discussion

Many normal tissues comprise a system of morphologically and functionally distinctive niches that differ by their cellular composition, extracellular matrix, metabolic conditions, and accessibility to the immune system. Although tissue microenvironment has been recognized as a major factor that influences tumour cell morphology and function,^{27,28} the impact of the niche on ITH and mutational evolution is poorly understood and largely limited to metastasis.^{29,30}

Our results help to understand how the distinct microenvironments of the skin influence the evolution of MF, a primary cutaneous, extranodal T-cell lymphoma. Our previous research showed that MF is clonotypically and genetically diverse exhibiting a high degree of ITH. The main mechanism responsible for the heterogeneity is the hematogenous seeding of skin lesions by clonotypically diverse neoplastic T-cells. The finding that the epidermal and dermal layers of the skin comprise distinct malignant clonotypes allowed us to conclude here that those compartments had been colonized by different clones of cancer cells (in this context, we define the clone as a population of malignant T-cells that are derived from the common precursor cell and exhibit the same TCR β clonotype). Thus, the lesion of MF does not develop via a gradual infiltration of the tissue by the expanding tumour, but by independent microinvasion events in which different niches in the skin are colonized independently by various T-cell clones (**Fig 5** and appendix **Fig D1**). Our conclusion was further confirmed by the finding that clonotypic diversity of the intraepidermal malignant cells exceeded the

diversity found in the dermis. If the infiltration of the epidermis had been caused by some clones in the dermal infiltrate, the opposite phenomenon would have been found, i.e. higher number of malignant clonotypes in the dermis and a smaller number of epidermal clonotypes overlapping with the dermal clones. These findings reinforce and broaden the concept of epidermotropism in MF, which originally described the morphological impression of movement of malignant T-cells from the epidermis to the dermis. It seems that epidermotropism is a feature of early malignant T-cell clones which seed the epidermis more readily than the dermis.

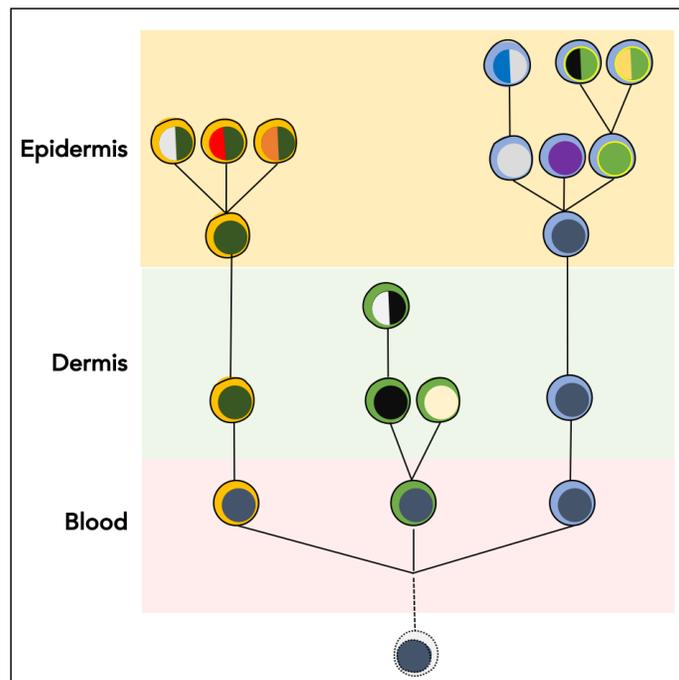


Figure 5.5: Generation of ecological heterogeneity in MF.

Skin lesions of MF are initiated by circulating, clonotypically heterogeneous malignant T-cell clones (various clonotypes are highlighted by different colours of the “cytoplasm”). Upon entering the skin some clones remain in the dermis where they proliferate whereas others pass directly to the epidermis. Expanding clones accumulate mutations leading to emergence of genetically different malignant subclones (different colours of the “nucleus”). Solid lines symbolize the phylogenetic relationship between the generations of malignant cells and illustrate divergent, neutral evolution of the subclones. Based on data in this paper and our previous work.^{15-17,36}

We were also able to conclude that the existence of different skin niches colonized by cancer facilitates the development of mutational subclones and augments ITH. By analyzing the phylogenetic trees of MF, we found that cancer subclones seem to develop independently in each compartment via a neutral, branched evolutionary process. Similar patterns of neutral evolution have previously been found in other solid neoplasms such as the lung or colorectal cancers.^{26,31} The symmetric shapes of the trees suggested lack of perceptible competition between epidermal and dermal subclones. We postulate therefore that evolving subclones in different compartments do not directly compete for the niche or the nutrients and develop independently adding to the overall ITH of the lesion. Of note, lack of competition does not imply lack of interactions between different subclones. It is likely (although at this stage still hypothetical) that different MF subclones cooperate to achieve optimal tumour growth in a similar manner to what has already been shown for solid cancers.³²

The described independent evolution of cancer subclones in different tissue microenvironments is reminiscent of the well-known phenomenon of ecological speciation during the evolution of organisms, in particular, the so-called peripatric speciation.^{33,34} It occurs when a small fraction of the population becomes separated into a new environment. The major difference, however, is the nature of the isolation. Rather than reproductive isolation essential for speciation, different microenvironments of the tissue separate evolving subclones protecting them against direct competition. The result is a more rapid increase in ITH than what could be achieved in a homogenous environment. We would like to propose the term “ecological heterogeneity” to describe the difference in cancer subclones in various microenvironments of the tissue.

High ITH of the tumors has been correlated to unfavourable prognosis over a large range of cancers.⁵ Currently, data are too limited to be able to investigate the prognostic role of ITH in MF. Some indirect evidence, such as the correlation between the presence of Pautrier microabscesses with the risk of progression³⁵ and higher ITH in MF tumors as compared to early plaques¹⁷ suggest that this indeed may be the case. One of the indications that ecological heterogeneity in MF may have functional significance is the non-overlapping pattern of driver mutations in epidermal and dermal neoplastic cells, which potentially would limit the efficacy of targeted therapies. However, a more detailed understanding of the differences between functionally significant signalling pathways on the level of transcriptome and the protein is needed.

5.6 References

1. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375–2390.
2. Olsen E, Vonderheid E, Pimpinelli N, et al. Revisions to the staging and classification of mycosis fungoides and Sezary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the cutaneous lymphoma task force of the European Organization of Research and Treatment of Cancer (EORTC). *Blood*. 2007;110(6):1713–1722.
3. Kim YH, Willemze R, Pimpinelli N, et al. TNM classification system for primary cutaneous lymphomas other than mycosis fungoides and Sezary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the Cutaneous Lymphoma Task Force of the European Organization of Research and Treatment of Cancer (EORTC). *Blood*. 2007;110(2):479–484.
4. Willemze R. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768–3785.
5. McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell*. 2017;168(4):613–628.
6. Vogelstein B, Kinzler KW. The Path to Cancer --Three Strikes and You're Out. *N. Engl. J. Med*. 2015;373(20):1895–1898.
7. Ellsworth DL, Blackburn HL, Shriver CD, et al. Single-cell sequencing and tumorigenesis: improved understanding of tumor evolution and metastasis. *Clin. Transl. Med*. 2017;6(1):15.

8. Heyde A, Reiter JG, Naxerova K, Nowak MA. Consecutive seeding and transfer of genetic diversity in metastasis. *Proc. Natl. Acad. Sci. U. S. A.* 2019;116(28):14129–14137.
9. Klemke CD, Booken N, Weiss C, et al. Histopathological and immunophenotypical criteria for the diagnosis of Sézary syndrome in differentiation from other erythrodermic skin diseases: a European Organisation for Research and Treatment of Cancer (EORTC) Cutaneous Lymphoma Task Force Study of 97 cases. *Br. J. Dermatol.* 2015;173(1):93–105.
10. Massone C, Kodama K, Kerl H, Cerroni L. Histopathologic features of early (patch) lesions of mycosis fungoides: a morphologic study on 745 biopsy specimens from 427 patients. *Am. J. Surg. Pathol.* 2005;29(4):550–560.
11. Chu AC, Morgan EW, MacDonald DM. Ultrastructural identification of T lymphocytes in tissue sections of mycosis fungoides. *J. Invest. Dermatol.* 1980;74(1):17–20.
12. Buus TB, Willerslev-Olsen A, Fredholm S, et al. Single-cell heterogeneity in Sézary syndrome. *Blood Adv.* 2018;2(16):2115–2126.
13. Gaydosik AM, Tabib T, Geskin LJ, et al. Single-Cell Lymphocyte Heterogeneity in Advanced Cutaneous T-cell Lymphoma Skin Tumors. *Clin. Cancer Res.* 2019;25(14):4443–4454.
14. Litvinov IV, Tetzlaff MT, Thibault P, et al. Gene expression analysis in Cutaneous T-Cell Lymphomas (CTCL) highlights disease heterogeneity and potential diagnostic and prognostic indicators. *Oncoimmunology.* 2017;6(5):e1306618.

15. Iyer A, Hennessey D, O’Keefe S, et al. Clonotypic heterogeneity in cutaneous T-cell lymphoma (mycosis fungoides) revealed by comprehensive whole-exome sequencing. *Blood Adv.* 2019;3(7):1175–1184.
16. Iyer A, Hennessey D, O’Keefe S, et al. Skin colonization by circulating neoplastic clones in cutaneous T-cell lymphoma. *Blood.* 2019;blood.2019002516.
17. Iyer A, Hennessey D, O’Keefe S, et al. Branched evolution and genomic intratumor heterogeneity in the pathogenesis of cutaneous T-cell lymphoma.
18. Bolotin DA, Poslavsky S, Mitrophanov I, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods.* 2015;12:380.
19. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics.* 2013;43:11.10.1–33.
20. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* 2013;31:213.
21. Kim S, Scheffler K, Halpern AL, et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat. Methods.* 2018;15(8):591–594.
22. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol.* 2016;17(1):122.
23. Ha G, Roth A, Khattra J, et al. TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. *Genome Res.* 2014;24(11):1881–1893.

24. Deshwar AG, Vembu S, Yung CK, et al. PhyloWGS: Reconstructing subclonal composition and evolution from whole-genome sequencing of tumors. *Genome Biol.* 2015;16(1):35.
25. Krangel MS. Mechanics of T cell receptor gene rearrangement. *Curr. Opin. Immunol.* 2009;21(2):133–139.
26. Williams MJ, Werner B, Barnes CP, Graham TA, Sottoriva A. Identification of neutral tumor evolution across cancer types. *Nat. Genet.* 2016;48(3):238–244.
27. Bissell MJ, Labarge MA. Context, tissue plasticity, and cancer: are tumor stem cells also regulated by the microenvironment? *Cancer Cell.* 2005;7(1):17–23.
28. Anderson ARA, Weaver AM, Cummings PT, Quaranta V. Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell.* 2006;127(5):905–915.
29. Polyak K, Haviv I, Campbell IG. Co-evolution of tumor cells and their microenvironment. *Trends Genet.* 2009;25(1):30–38.
30. Zhang L, Zhang S, Yao J, et al. Microenvironment-induced PTEN loss by exosomal microRNA primes brain metastasis outgrowth. *Nature.* 2015;527(7576):100–104.
31. Sottoriva A, Kang H, Ma Z, et al. A Big Bang model of human colorectal tumor growth. *Nat. Genet.* 2015;47(3):209–216.
32. Zhou H, Neelakantan D, Ford HL. Clonal cooperativity in heterogenous cancers. *Semin. Cell Dev. Biol.* 2017;64:79–89.
33. Templeton AR. The reality and importance of founder speciation in evolution. *Bioessays.* 2008;30(5):470–479.

34. Schluter D. Evidence for ecological speciation and its alternative. *Science*. 2009;323(5915):737–741.
35. Vonderheid EC, Pavlov I, Delgado JC, et al. Prognostic factors and risk stratification in early mycosis fungoides. *Leuk. Lymphoma*. 2014;55(1):44–50.
36. Gniadecki R, Lukowsky A, Rossen K, et al. Bone marrow precursor of extranodal T-cell lymphoma. *Blood*. 2003;102(10):3797–3799.
37. Hamrouni A, Fogh H, Zak Z, Ødum N, Gniadecki R. Clonotypic Diversity of the T-cell Receptor Corroborates the Immature Precursor Origin of Cutaneous T-cell Lymphoma. *Clin. Cancer Res*. 2019;25(10):3104–3114.

Chapter 6: Discussion and Conclusion

6.1 General Discussion

The work in this thesis challenges the long-standing concept of MF as a lymphoma developing from the skin resident memory T-cells. We show that MF is heterogenous on multiple levels: it comprises multiple T-cell clones and shows genetic heterogeneity being composed of multiple genetic subclones. We further provide evidence that malignant T-cells in MF are not solely skin resident but are also present in the circulating blood irrespective of the stage of the disease.

The major contribution of the thesis is as following.

1. By using WES to calculate the TCF we can determine the frequency of malignant T-cells and eliminate the need for arbitrary thresholds for “clonality”.
2. Identification of previously unreported mutations in putative driver genes in MF.
3. To provide evidence that MF skin lesions presents genetic and clonotypic heterogeneity.
4. To propose the model of MF in which skin lesions develop from seeding by the circulating precursors independently in different lesions and different niches in the skin which further develop by neutral branched evolution.

We are aware of certain limitations of this work. First, the sensitivity of WES for identifying the TCR repertoire was lower as compared with the PCR based amplification. Even with increased sequencing depth of ~800x the number of TCR clonotypes identified by WES cannot identify the low frequency clonotypes with the same sensitivity as target amplification method. Second, we used LCM to isolate the malignant T-cells to improve the TCF for the samples. Though we achieved this goal, we understand that the LCM cells may not represent

the complete diversity of the skin lesion. Third, we did not corroborate our TCR and genetic heterogeneity results at transcriptome level.

6.2 Conclusion and perspectives

MF is composed of a clonally and genetically heterogeneous mixture of mutation-rich, malignant T-cells. The disease develops and progresses via hematogenous spread of malignant cells that seed new areas of the skin as well as the already existing sites.

Our model of MF pathogenesis provides a plausible explanation of disease resistance and relapse. Skin directed therapies alone are not likely to provide cure as they fail to target the malignant T-cells in circulating blood. Drugs that target circulating cells and prevent their entry to the skin (such as mogamalizumab, the CCR4 inhibitor) are likely to be effective in MF. Finally, due to the heterogeneous nature of MF, combination therapies would be required to efficiently target multiple subclones. Future studies with single cell sequencing and additional normal control would be beneficial in identifying treatment targets and the characteristics of the immature T-cell precursors lying at the root of pathogenesis of MF.

Bibliography

References for chapter 1:

1. Rizvi MA. T-cell non-Hodgkin lymphoma. *Blood*. 2006;107(4):1255–1264.
2. Ansell SM. Non-Hodgkin Lymphoma: Diagnosis and Treatment. *Mayo Clinic Proceedings*. 2015;90(8):1152–1163.
3. Vd C, Ma W. Incidence of cutaneous t-cell lymphoma in the united states, 1973-2002. *Arch. Dermatol*. 2007;143(7):854–859.
4. Rodd AL, Ververis K, Karagiannis TC. Current and Emerging Therapeutics for Cutaneous T-Cell Lymphoma: Histone Deacetylase Inhibitors. *Lymphoma*. 2012;2012:1–10.
5. Bagherani N, Smoller BR. An overview of cutaneous T cell lymphomas. *F1000Research*. 2016;5:1882.
6. Wilcox RA. Cutaneous T-cell lymphoma: 2016 update on diagnosis, risk-stratification, and management. *Am. J. Hematol*. 2016;91(1):151–165.
7. Willemze R, Jaffe ES, Burg G, et al. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768 LP–3785.
8. Olsen E, Vonderheid E, Pimpinelli N, et al. Revisions to the staging and classification of mycosis fungoides and Sézary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the cutaneous lymphoma task force of the European Organization of Research and Treatment of Ca. *Blood*. 2007;110(6):1713 LP–1722.
9. Huber MA, Kunzi-Rapp K, Staib G, Scharffetter-Kochanek K. Management of refractory early-stage cutaneous T-cell lymphoma (mycosis fungoides) with a combination of oral

- bexarotene and psoralen plus ultraviolet bath therapy. *Journal of the American Academy of Dermatology*. 2004;50(3):475–476.
10. Foo SH, Shah F, Chaganti S, Stevens A, Scarisbrick JJ. Unmasking mycosis fungoides/Sézary syndrome from preceding or co-existing benign inflammatory dermatoses requiring systemic therapies: patients frequently present with advanced disease and have an aggressive clinical course. *Br. J. Dermatol*. 2016;174(4):901–904.
 11. Eklund Y, Aronsson A, Schmidtchen A, Relander T. Mycosis Fungoides: A Retrospective Study of 44 Swedish Cases. *Acta Derm. Venereol*. 2016;96(5):669–673.
 12. Song SX, Willemze R, Swerdlow SH, Kinney MC, Said JW. Mycosis fungoides: report of the 2011 Society for Hematopathology/European Association for Haematopathology workshop. *Am. J. Clin. Pathol*. 2013;139(4):466–490.
 13. Harvey NT, Spagnolo DV, Wood BA. “Could it be mycosis fungoides?”: an approach to diagnosing patch stage mycosis fungoides. *Journal of Hematopathology*. 2015;8(4):209–223.
 14. Trautinger F, Knobler R, Willemze R, et al. EORTC consensus recommendations for the treatment of mycosis fungoides/Sézary syndrome. *Eur. J. Cancer*. 2006;42(8):1014–1030.
 15. Hwang ST, Janik JE, Jaffe ES, Wilson WH. Mycosis fungoides and Sézary syndrome. *Lancet*. 2008;371(9616):945–957.
 16. Campbell JJ, Clark RA, Watanabe R, Kupper TS. Sezary syndrome and mycosis fungoides arise from distinct T-cell subsets: a biologic rationale for their distinct clinical behaviors. *Blood*. 2010;116(5):767–771.

17. Compton CC, Byrd DR, Garcia-Aguilar J, et al. AJCC Cancer Staging Atlas: A Companion to the Seventh Editions of the AJCC Cancer Staging Manual and Handbook. Springer Science & Business Media; 2012.
18. Vonderheid EC, Bernengo MG, Burg G, et al. Update on erythrodermic cutaneous T-cell lymphoma: report of the International Society for Cutaneous Lymphomas. *J. Am. Acad. Dermatol.* 2002;46(1):95–106.
19. Shin J, Monti S, Aires DJ, et al. Lesional gene expression profiling in cutaneous T-cell lymphoma reveals natural clusters associated with disease outcome. *Blood.* 2007;110(8):3015–3027.
20. Kim YH, Liu HL, Mraz-Gernhard S, Varghese A, Hoppe RT. Long-term outcome of 525 patients with mycosis fungoides and Sezary syndrome: clinical prognostic factors and risk for disease progression. *Arch. Dermatol.* 2003;139(7):857–866.
21. van Doorn R, van Kester MS, Dijkman R, et al. Oncogenomic analysis of mycosis fungoides reveals major differences with Sezary syndrome. *Blood.* 2009;113(1):127–136.
22. Bassing CH, Swat W, Alt FW. The Mechanism and Regulation of Chromosomal V(D)J Recombination. *Cell.* 2018;109(2):S45–S55.
23. Hayday AC, Pennington DJ. Key factors in the organized chaos of early T cell development. *Nature Immunology.* 2007;8(2):137–144.
24. Krangel MS. Gene segment selection in V(D)J recombination: accessibility and beyond. *Nat. Immunol.* 2003;4(7):624–630.
25. Xiong N, Kang C, Raulet DH. Redundant and Unique Roles of Two Enhancer Elements in the TCR γ Locus in Gene Regulation and $\gamma\delta$ T Cell Development. *Immunity.* 2002;16(3):453–463.

26. Cobb RM, Oestreich KJ, Osipovich OA, Oltz EMBT-. A in I. Accessibility Control of V(D)J Recombination. 2006;91:45–109.
27. Dupic T, Marcou Q, Walczak AM, Mora T. Genesis of the $\alpha\beta$ T-cell receptor. *PLoS Comput. Biol.* 2019;15(3):e1006874.
28. Han A, Glanville J, Hansmann L, Davis MM. Corrigendum: Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat. Biotechnol.* 2015;33(2):210.
29. van Dongen JJM, Langerak AW, Brüggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: Report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia.* 2003;17:2257.
30. Thurber SE, Zhang B, Kim YH, et al. T-cell clonality analysis in biopsy specimens from two different skin sites shows high specificity in the diagnosis of patients with suggested mycosis fungoides. *J. Am. Acad. Dermatol.* 2007;57(5):782–790.
31. Barba G, Matteucci C, Girolomoni G, et al. Comparative genomic hybridization identifies 17q11.2 approximately q12 duplication as an early event in cutaneous T-cell lymphomas. *Cancer Genet. Cytogenet.* 2008;184(1):48–51.
32. da Silva Almeida AC, Abate F, Khiabani H, et al. The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome. *Nat. Genet.* 2015;47(12):1465–1470.
33. Prochazkova M, Chevret E, Mainhaguet G, et al. Common chromosomal abnormalities in mycosis fungoides transformation. *Genes Chromosomes Cancer.* 2007;46(9):828–838.
34. Fischer TC, Gellrich S, Muche JM, et al. Genomic aberrations and survival in cutaneous T cell lymphomas. *J. Invest. Dermatol.* 2004;122(3):579–586.

35. Carbone A, Bernardini L, Valenzano F, et al. Array-based comparative genomic hybridization in early-stage mycosis fungoides: recurrent deletion of tumor suppressor genes BCL7A, SMAC/DIABLO, and RHOV. *Genes Chromosomes Cancer*. 2008;47(12):1067–1075.
36. Salgado R, Servitje O, Gallardo F, et al. Oligonucleotide array-CGH identifies genomic subgroups and prognostic markers for tumor stage mycosis fungoides. *J. Invest. Dermatol*. 2010;130(4):1126–1135.
37. Karenko L, Kähkönen M, Hyytinen ER, Lindlof M, Ranki A. Notable losses at specific regions of chromosomes 10q and 13q in the Sézary syndrome detected by comparative genomic hybridization. *J. Invest. Dermatol*. 1999;112(3):392–395.
38. So CC, Wong KF, Siu LL, Kwong YL. Large cell transformation of Sézary syndrome. A conventional and molecular cytogenetic study. *Am. J. Clin. Pathol*. 2000;113(6):792–797.
39. Mao X, Lillington D, Scarisbrick JJ, et al. Molecular cytogenetic analysis of cutaneous T-cell lymphomas: identification of common genetic alterations in Sézary syndrome and mycosis fungoides. *Br. J. Dermatol*. 2002;147(3):464–475.
40. Mao X, Lillington DM, Czepulkowski B, et al. Molecular cytogenetic characterization of Sézary syndrome. *Genes Chromosomes Cancer*. 2003;36(3):250–260.
41. Vermeer MH, van Doorn R, Dijkman R, et al. Novel and highly recurrent chromosomal alterations in Sézary syndrome. *Cancer Res*. 2008;68(8):2689–2698.
42. Caprini E, Cristofolletti C, Arcelli D, et al. Identification of key regions and genes important in the pathogenesis of sezary syndrome by combining genomic and expression microarrays. *Cancer Res*. 2009;69(21):8438–8446.

43. Boudjarane J, Essaydi A, Farnault L, et al. Characterization of the novel Sezary lymphoma cell line BKP1. *Exp. Dermatol.* 2015;24(1):60–62.
44. Laharanne E, Oumouhou N, Bonnet F, et al. Genome-wide analysis of cutaneous T-cell lymphomas identifies three clinically relevant classes. *J. Invest. Dermatol.* 2010;130(6):1707–1718.
45. Gug G, Huang Q, Chiticariu E, Solovan C, Baudis M. DNA copy number imbalances in primary cutaneous lymphomas. *Journal of the European Academy of Dermatology and Venereology.* 2019;33(6):1062–1075.
46. McGirt LY, Jia P, Baerenwald DA, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. *Blood.* 2015;126(4):508–519.
47. Ungewickell A, Bhaduri A, Rios E, et al. Genomic analysis of mycosis fungoides and Sézary syndrome identifies recurrent alterations in TNFR2. *Nat. Genet.* 2015;47(9):1056–1060.
48. Wang L, Ni X, Covington KR, et al. Genomic profiling of Sézary syndrome identifies alterations of key T cell signaling and differentiation genes. *Nat. Genet.* 2015;47(12):1426–1434.
49. Moffitt AB, Dave SS. Clinical Applications of the Genomic Landscape of Aggressive Non-Hodgkin Lymphoma. *J. Clin. Oncol.* 2017;35(9):955–962.
50. Park J, Yang J, Wenzel AT, et al. Genomic analysis of 220 CTCLs identifies a novel recurrent gain-of-function alteration in RLTPR (p.Q575E). *Blood.* 2017;130(12):1430 LP–1440.
51. Tensen CP. PLCG1 Gene Mutations in Cutaneous T-Cell Lymphomas Revisited. *J. Invest. Dermatol.* 2015;135(9):2153–2154.

52. Prasad A, Rabionet R, Espinet B, et al. Identification of Gene Mutations and Fusion Genes in Patients with Sézary Syndrome. *J. Invest. Dermatol.* 2016;136(7):1490–1499.
53. Kiel MJ, Sahasrabudde AA, Rolland DCM, et al. Genomic analyses reveal recurrent mutations in epigenetic modifiers and the JAK–STAT pathway in Sézary syndrome. *Nature Communications.* 2015;6(1.):
54. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. *Nature Genetics.* 2015;47(9):1011–1019.
55. McGranahan N, Rosenthal R, Hiley CT, et al. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell.* 2017;171(6):1259–1271.e11.
56. Mardis E. Faculty of 1000 evaluation for Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. *F1000 - Post-publication peer review of the biomedical literature.* 2015;
57. Zhang J, Fujimoto J, Zhang J, et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. *Science.* 2014;346(6206):256–259.
58. Harbst K, Lauss M, Cirenajwis H, et al. Multiregion Whole-Exome Sequencing Uncovers the Genetic Evolution and Mutational Heterogeneity of Early-Stage Metastatic Melanoma. *Cancer Res.* 2016;76(16):4765–4774.
59. Gundem G, Van Loo P, Kremeyer B, et al. The evolutionary history of lethal metastatic prostate cancer. *Nature.* 2015;520(7547):353–357.
60. Buus TB, Willerslev-Olsen A, Fredholm S, et al. Single-cell heterogeneity in Sézary syndrome. *Blood Adv.* 2018;2(16):2115–2126.

61. Borchering N, Voigt AP, Liu V, et al. Single-Cell Profiling of Cutaneous T-Cell Lymphoma Reveals Underlying Heterogeneity Associated with Disease Progression. *Clin. Cancer Res.* 2019;25(10):2996–3005.
62. Litvinov IV, Tetzlaff MT, Thibault P, et al. Gene expression analysis in Cutaneous T-Cell Lymphomas (CTCL) highlights disease heterogeneity and potential diagnostic and prognostic indicators. *Oncoimmunology.* 2017;6(5):e1306618.
63. Gaydosik AM, Tabib T, Geskin LJ, et al. Single-Cell Lymphocyte Heterogeneity in Advanced Cutaneous T-cell Lymphoma Skin Tumors. *Clin. Cancer Res.* 2019;25(14):4443–4454.
64. Bradford PT, Devesa SS, Anderson WF, Toro JR. Cutaneous lymphoma incidence patterns in the United States: a population-based study of 3884 cases. *Blood.* 2009;113(21):5064 LP–5073.
65. Kirsch IR, Watanabe R, O'Malley JT, et al. TCR sequencing facilitates diagnosis and identifies mature T cells as the cell of origin in CTCL. *Sci. Transl. Med.* 2015;7(308):308ra158.
66. Masson A de, de Masson A, O'Malley JT, et al. High-throughput sequencing of the T cell receptor β gene identifies aggressive early-stage mycosis fungoides. *Science Translational Medicine.* 2018;10(440):eaar5894.

References for chapter 2:

1. Olsen E, Vonderheid E, Pimpinelli N, et al. Revisions to the staging and classification of mycosis fungoides and Sézary syndrome: a proposal of the International Society for

- Cutaneous Lymphomas (ISCL) and the cutaneous lymphoma task force of the European Organization of Research and Treatment of Ca. *Blood*. 2007;110(6):1713 LP-1722.
2. Kim YH, Willemze R, Pimpinelli N, et al. TNM classification system for primary cutaneous lymphomas other than mycosis fungoides and Sezary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the Cutaneous Lymphoma Task Force of the European Organization of Resear. *Blood*. 2007;110(2):479–484.
 3. Willemze R, Jaffe ES, Burg G, et al. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768 LP-3785.
 4. Scarisbrick JJ, Prince HM, Vermeer MH, et al. Cutaneous Lymphoma International Consortium Study of Outcome in Advanced Stages of Mycosis Fungoides and Sezary Syndrome: Effect of Specific Prognostic Markers on Survival and Development of a Prognostic Model. *J. Clin. Oncol*. 2015;33(32):3766–3773.
 5. Kim EJ, Hess S, Richardson SK, et al. Immunopathogenesis and therapy of cutaneous T cell lymphoma. *J. Clin. Invest*. 2005;115(4):798–812.
 6. Davis MM, Bjorkman PJ. T-cell antigen receptor genes and T-cell recognition. *Nature*. 1988;334(6181):395–402.
 7. Manfras BJ, Terjung D, Boehm BO. Non-productive human TCR beta chain genes represent V-D-J diversity before selection upon function: insight into biased usage of TCRBD and TCRBJ genes and diversity of CDR3 region length. *Hum. Immunol*. 1999;60(11):1090–1100.
 8. Kaplinsky J, Arnaout R. Robust estimates of overall immune-repertoire diversity from high-throughput measurements on samples. *Nat. Commun*. 2016;7:11881.

9. de Masson A, O'Malley JT, Elco CP, et al. High-throughput sequencing of the T cell receptor beta gene identifies aggressive early-stage mycosis fungoides. *Sci. Transl. Med.* 2018;10(440).
10. Assaf C, Hummel M, Dippel E, et al. High detection rate of T-cell receptor beta chain rearrangements in T-cell lymphoproliferations by family specific polymerase chain reaction in combination with the GeneScan technique and DNA sequencing. *Blood.* 2000;96(2):640–646.
11. Hou X-L, Wang L, Ding Y-L, Xie Q, Diao H-Y. Current status and recent advances of next generation sequencing techniques in immunological repertoire. *Genes Immun.* 2016;17(3):153–164.
12. Kirsch IR, Watanabe R, O'Malley JT, et al. TCR sequencing facilitates diagnosis and identifies mature T cells as the cell of origin in CTCL. *Sci. Transl. Med.* 2015;7(308):308ra158.
13. Sufficool KE, Lockwood CM, Abel HJ, et al. T-cell clonality assessment by next-generation sequencing improves detection sensitivity in mycosis fungoides. *J. Am. Acad. Dermatol.* 2015;73(2):228–36.e2.
14. Levy E, Marty R, Gárate Calderón V, et al. Immune DNA signature of T-cell infiltration in breast tumor exomes. *Sci. Rep.* 2016;6:30064.
15. Ha G, Roth A, Khattra J, et al. TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. *Genome Res.* 2014;24(11):1881–1893.
16. Bolotin DA, Poslavsky S, Mitrophanov I, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods.* 2015;12:380.

17. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297–1303.
18. Nazarov VI, Pogorelyy M V, Komech EA, et al. tcR: an R package for T cell receptor repertoire advanced data analysis. *BMC Bioinformatics.* 2015;16(1):175.
19. Shugay M, Bagaev D V., Turchaninova MA, et al. VDJtools: Unifying Post-analysis of T Cell Receptor Repertoires. *PLOS Comput. Biol.* 2015;2015;11(11):e1004503
20. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 2011;43(5):491–498.
21. Hamrouni A, Aublin A, Guillaume P, Maryanski JL. T Cell Receptor Gene Rearrangement Lineage Analysis Reveals Clues for the Origin of Highly Restricted Antigen-specific Repertoires. *J. Exp. Med.* 2003;197(5):601 LP-614.
22. Linnemann C, Heemskerk B, Kvistborg P, et al. High-throughput identification of antigen-specific TCRs by TCR gene capture. *Nat. Med.* 2013;19(11):1534–1541.
23. Ruggiero E, Nicolay JP, Fronza R, et al. High-resolution analysis of the human T-cell receptor repertoire. *Nat. Commun.* 2015;6:8081.
24. Gong Q, Wang C, Zhang W, et al. Assessment of T-cell receptor repertoire and clonal expansion in peripheral T-cell lymphoma using RNA-seq data. *Sci. Rep.* 2017;7(1):11301.
25. Linnemann T, Gellrich S, Lukowsky A, et al. Polyclonal expansion of T cells with the TCR V beta type of the tumour cell in lesions of cutaneous T-cell lymphoma: evidence for possible superantigen involvement. *Br. J. Dermatol.* 2004;150(5):1013–1017.

26. Bagot M, Echchakir H, Mami-Chouaib F, et al. Isolation of Tumor-Specific Cytotoxic CD4⁺ and CD4⁺CD8dim⁺ T-Cell Clones Infiltrating a Cutaneous T-Cell Lymphoma. *Blood*. 1998;91(11):4331 LP-4341.
27. Dobbeling U, Dummer R, Schmid MH, Burg G. Lack of expression of the recombination activating genes RAG-1 and RAG-2 in cutaneous T-cell lymphoma: pathogenic implications. *Clin. Exp. Dermatol*. 2006;22(5):230–233.
28. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. *Nat. Genet*. 2015;47(9):1011–1019.
29. Burg G, Kempf W, Haeffner A, et al. From inflammation to neoplasia: new concepts in the pathogenesis of cutaneous lymphomas. *Recent results cancer Res. Fortschritte der Krebsforsch. Prog. dans les Rech. sur le cancer*. 2002;160:271–280.
30. Jackow CM, Cather JC, Hearne V, et al. Association of erythrodermic cutaneous T-cell lymphoma, superantigen-positive *Staphylococcus aureus*, and oligoclonal T-cell receptor V beta gene expansion. *Blood*. 1997;89(1):32–40.
31. Keith R. Hudson, Helen Robinson, and John D. Two adjacent residues in staphylococcal enterotoxins A and E determine T cell receptor V beta specificity. *J. Exp. Med*. 1993;177(1):175–184.
32. Dean J, Emerson RO, Vignali M, et al. Annotation of pseudogenic gene segments by massively parallel sequencing of rearranged lymphocyte receptor loci. *Genome Med*. 2015;7(1):123.
33. Brunner PM, Emerson RO, Tipton C, et al. Nonlesional atopic dermatitis skin shares similar T-cell clones with lesional tissues. *Allergy*. 2017;72(12):2017–2025.

References for chapter 3:

1. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375–2390.
2. Olsen E, Vonderheid E, Pimpinelli N, et al. Revisions to the staging and classification of mycosis fungoides and Sezary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the cutaneous lymphoma task force of the European Organization of Research and Treatment of Cancer (EORTC). *Blood*. 2007;110(6):1713–1722.
3. Kim YH, Willemze R, Pimpinelli N, et al. TNM classification system for primary cutaneous lymphomas other than mycosis fungoides and Sezary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the Cutaneous Lymphoma Task Force of the European Organization of Research and Treatment of Cancer (EORTC). *Blood*. 2007;110(2):479–484.
4. Willemze R. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768–3785.
5. Clark RA, Watanabe R, Teague JE, et al. Skin effector memory T cells do not recirculate and provide immune protection in alemtuzumab-treated CTCL patients. *Sci. Transl. Med*. 2012;4(117):117ra7.
6. Campbell JJ, Clark RA, Watanabe R, Kupper TS. Sezary syndrome and mycosis fungoides arise from distinct T-cell subsets: a biologic rationale for their distinct clinical behaviors. *Blood*. 2010;116(5):767–771.
7. Whittaker S, Ortiz P, Dummer R, et al. Efficacy and safety of bexarotene combined with psoralen-ultraviolet A (PUVA) compared with PUVA treatment alone in stage IB-IIA

- mycosis fungoides: final results from the EORTC Cutaneous Lymphoma Task Force phase III randomized clinical trial 21011. *Br. J. Dermatol.* 2012;167(3):678–687.
8. Kamstrup MR, Gniadecki R, Iversen L, et al. Low-dose (10-Gy) total skin electron beam therapy for cutaneous T-cell lymphoma: an open clinical study and pooled data analysis. *Int. J. Radiat. Oncol. Biol. Phys.* 2015;92(1):138–143.
 9. Chowdhary M, Chhabra AM, Kharod S, Marwaha G. Total Skin Electron Beam Therapy in the Treatment of Mycosis Fungoides: A Review of Conventional and Low-Dose Regimens. *Clin. Lymphoma Myeloma Leuk.* 2016;16(12):662–671.
 10. Vieyra-Garcia P, Fink-Puches R, Porkert S, et al. Evaluation of Low-Dose, Low-Frequency Oral Psoralen-UV-A Treatment With or Without Maintenance on Early-Stage Mycosis Fungoides: A Randomized Clinical Trial. *JAMA Dermatol.* 2019;
 11. Gniadecki R, Lukowsky A, Rossen K, et al. Bone marrow precursor of extranodal T-cell lymphoma. *Blood.* 2003;102(10):3797–3799.
 12. Berg KD, Brinster NK, Huhn KM, et al. Transmission of a T-cell lymphoma by allogeneic bone marrow transplantation. *N. Engl. J. Med.* 2001;345(20):1458–1463.
 13. Fahy CMR, Fortune A, Quinn F, et al. Development of mycosis fungoides after bone marrow transplantation for chronic myeloid leukaemia: transmission from an allogeneic donor. *Br. J. Dermatol.* 2014;170(2):462–467.
 14. Davis TH, Morton CC, Miller-Cassman R, Balk SP, Kadin ME. Hodgkin's disease, lymphomatoid papulosis, and cutaneous T-cell lymphoma derived from a common T-cell clone. *N. Engl. J. Med.* 1992;326(17):1115–1122.
 15. Kamstrup MR, Gniadecki R, Skovgaard GL. Putative cancer stem cells in cutaneous malignancies. *Exp. Dermatol.* 2007;16(4):297–301.

16. Delfau-Larue M, Petrella T, Lahet C, et al. Value of clonality studies of cutaneous T lymphocytes in the diagnosis and follow-up of patients with mycosis fungoides. *J. Pathol.* 1998;184(2):185–190.
17. Jones D, Duvic M. The current state and future of clonality studies in mycosis fungoides. *J. Invest. Dermatol.* 2003;121(3):ix–x.
18. Mahe E, Pugh T, Kamel-Reid S. T cell clonality assessment: past, present and future. *J. Clin. Pathol.* 2018;71(3):195–200.
19. Krangel MS. Mechanics of T cell receptor gene rearrangement. *Curr. Opin. Immunol.* 2009;21(2):133–139.
20. Hamrouni A, Fogh H, Zak ZL, Odum N, Gniadecki R. Clonotypic diversity of the T-cell receptor corroborates the immature precursor origin of cutaneous T-cell lymphoma. *Clin. Cancer Res.* 2019;
21. Iyer A, Hennessey D, O’Keefe S, et al. Clonotypic heterogeneity in cutaneous T-cell lymphoma (mycosis fungoides) revealed by comprehensive whole-exome sequencing. *Blood Adv.* 2019;3(7):1175–1184.
22. Boyman O, Hefti HP, Conrad C, et al. Spontaneous Development of Psoriasis in a New Animal Model Shows an Essential Role for Resident T Cells and Tumor Necrosis Factor- α . *J. Exp. Med.* 2004;199(5):731–736.
23. Watanabe R, Gehad A, Yang C, et al. Human skin is protected by four functionally and phenotypically discrete populations of resident and recirculating memory T cells. *Sci. Transl. Med.* 2015;7(279):279ra39.
24. Langerak AW, Molina TJ, Lavender FL, et al. Polymerase chain reaction-based clonality testing in tissue samples with reactive lymphoproliferations: usefulness and pitfalls. A

- report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia*. 2007;21(2):222–229.
25. Marcus Mucbe J, Karenko L, Gellrich S, et al. Cellular coincidence of clonal T cell receptor rearrangements and complex clonal chromosomal aberrations—a hallmark of malignancy in cutaneous T cell lymphoma. *J. Invest. Dermatol.* 2004;122(3):574–578.
 26. Mucbe JM, Lukowsky A, Heim J, et al. Demonstration of frequent occurrence of clonal T cells in the peripheral blood but not in the skin of patients with small plaque parapsoriasis. *Blood*. 1999;94(4):1409–1417.
 27. Beylot-Barry M, Sibaud V, Thiebaut R, et al. Evidence that an identical T cell clone in skin and peripheral blood lymphocytes is an independent prognostic factor in primary cutaneous T cell lymphomas. *J. Invest. Dermatol.* 2001;117(4):920–926.
 28. Fraser-Andrews EA, Woolford AJ, Russell-Jones R, Whittaker SJ, Seed PT. Detection of a Peripheral Blood T Cell Clone is an Independent Prognostic Marker in Mycosis Fungoides. *J. Invest. Dermatol.* 2000;114(1):117–121.
 29. Bolotin DA, Poslavsky S, Mitrophanov I, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods*. 2015;12:380.
 30. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* 2011;43(5):491–498.
 31. Ha G, Roth A, Khattra J, et al. TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. *Genome Res*. 2014;24(11):1881–1893.

32. Nazarov VI, Pogorelyy MV, Komech EA, et al. tcR: an R package for T cell receptor repertoire advanced data analysis. *BMC Bioinformatics*. 2015;16:175.
33. Pogorelyy MV, Elhanati Y, Marcou Q, et al. Persisting fetal clonotypes influence the structure and overlap of adult human T cell receptor repertoires. *PLoS Comput. Biol.* 2017;13(7):e1005572.
34. Dupic T, Marcou Q, Walczak AM, Mora T. Genesis of the $\alpha\beta$ T-cell receptor. *PLoS Comput. Biol.* 2019;15(3):e1006874.
35. Khor B, Sleckman BP. Allelic exclusion at the TCRbeta locus. *Curr. Opin. Immunol.* 2002;14(2):230–234.
36. Mora T, Walczak AM. Quantifying lymphocyte receptor diversity. *arXiv*:1604.00487
37. de Masson A, O'Malley JT, Elco CP, et al. High-throughput sequencing of the T cell receptor β gene identifies aggressive early-stage mycosis fungoides. *Sci. Transl. Med.* 2018;10(440.):
38. Kirsch IL, Sherwood A, Williamson D, et al. The utility of HTS TCR analyses in the management of patients with T-cell malignancies. *J. Clin. Oncol.* 2014;32(15_suppl):8565–8565.
39. Döbbeling U, Dummer R, Hess Schmid M, Burg G. Lack of expression of the recombination activating genes RAG-1 and RAG-2 in cutaneous T-cell lymphoma: pathogenic implications. *Clin. Exp. Dermatol.* 1997;22(5):230–233.
40. Kim M-Y, Oskarsson T, Acharyya S, et al. Tumor Self-Seeding by Circulating Cancer Cells. *Cell*. 2009;139(7):1315–1326.
41. Aceto N, Bardia A, Miyamoto DT, et al. Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell*. 2014;158(5):1110–1122.

42. Cleary AS, Leonard TL, Gestl SA, Gunther EJ. Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature*. 2014;508(7494):113–117.
43. Heyde A, Reiter JG, Naxerova K, Nowak MA. Consecutive seeding and transfer of genetic diversity in metastasis. *Proc. Natl Acad. Sci.* 2019;116(28):14129–14137.
44. Vega F, Luthra R, Medeiros LJ, et al. Clonal heterogeneity in mycosis fungoides and its relationship to clinical course. *Blood*. 2002;100(9):3369–3373.
45. Kolowos W, Herrmann M, Ponner BB, et al. Detection of restricted junctional diversity of peripheral T cells in SLE patients by spectratyping. *Lupus*. 1997;6(9):701–707.
46. Gniadecki R, Lukowsky A. Monoclonal T-cell dyscrasia of undetermined significance associated with recalcitrant erythroderma. *Arch. Dermatol.* 2005;141(3):361–367.
47. Wack A, Cossarizza A, Heltai S, et al. Age-related modifications of the human alphabeta T cell repertoire due to different clonal expansions in the CD4+ and CD8+ subsets. *Int. Immunol.* 1998;10(9):1281–1288.
48. Posnett DN, Sinha R, Kabak S, Russo C. Clonal populations of T cells in normal elderly humans: the T cell equivalent to “benign monoclonal gammopathy.” *J. Exp. Med.* 1994;179(2):609–618.
49. Muche JM, Sterry W, Gellrich S, et al. Peripheral blood T-cell clonality in mycosis fungoides and nonlymphoma controls. *Diagn. Mol. Pathol.* 2003;12(3):142–150.
50. Buus TB, Willerslev-Olsen A, Fredholm S, et al. Single-cell heterogeneity in Sézary syndrome. *Blood Adv.* 2018;2(16):2115–2126.

51. Gaydosik AM, Tabib T, Geskin LJ, et al. Single-Cell Lymphocyte Heterogeneity in Advanced Cutaneous T-cell Lymphoma Skin Tumors. *Clin. Cancer Res.* 2019;25(14):4443–4454.
52. Farmanbar A, Kneller R, Firouzi S. RNA sequencing identifies clonal structure of T-cell repertoires in patients with adult T-cell leukemia/lymphoma. *NPJ Genom Med.* 2019;4:10.
53. Litvinov IV, Tetzlaff MT, Thibault P, et al. Gene expression analysis in Cutaneous T-Cell Lymphomas (CTCL) highlights disease heterogeneity and potential diagnostic and prognostic indicators. *Oncoimmunology.* 2017;6(5):e1306618.
54. Dereure O, Levi E, Vonderheid EC, Kadin ME. Improved sensitivity of T-cell clonality detection in mycosis fungoides by hand microdissection and heteroduplex analysis. *Arch. Dermatol.* 2003;139(12):1571–1575.
55. Woetmann A, Lovato P, Eriksen KW, et al. Nonmalignant T cells stimulate growth of T-cell lymphoma cells in the presence of bacterial toxins. *Blood.* 2007;109(8):3325–3332.
56. Lindahl LM, Willerslev-Olsen A, Gjerdrum LMR, et al. Antibiotics inhibit tumor and disease activity in cutaneous T cell lymphoma. *Blood.* 2019;
57. Willerslev-Olsen A, Krejsgaard T, Lindahl LM, et al. Staphylococcal enterotoxin A (SEA) stimulates STAT3 activation and IL-17 expression in cutaneous T-cell lymphoma. *Blood.* 2016;127(10):1287–1296.
58. Fraser JD, Proft T. The bacterial superantigen and superantigen-like proteins. *Immunol. Rev.* 2008;225:226–243.

59. Vonderheid EC, Bigler RD, Hou JS. On the possible relationship between staphylococcal superantigens and increased Vbeta5.1 usage in cutaneous T-cell lymphoma. *Br. J. Dermatol.* 2005;152(4):825–6; author reply 827.
60. Vonderheid EC, Boselli CM, Conroy M, et al. Evidence for restricted Vbeta usage in the leukemic phase of cutaneous T cell lymphoma. *J. Invest. Dermatol.* 2005;124(3):651–661.
61. Jackow CM, Cather JC, Hearne V, et al. Association of erythrodermic cutaneous T-cell lymphoma, superantigen-positive *Staphylococcus aureus*, and oligoclonal T-cell receptor V beta gene expansion. *Blood.* 1997;89(1):32–40.
62. Hamrouni A, Aublin A, Guillaume P, Maryanski JL. T cell receptor gene rearrangement lineage analysis reveals clues for the origin of highly restricted antigen-specific repertoires. *J. Exp. Med.* 2003;197(5):601–614.
63. Farcot E, Bonnet M, Jaeger S, et al. TCR β Allelic Exclusion in Dynamical Models of V(D)J Recombination Based on Allele Independence. *J. Immunol.* 2010;185(3):1622–1632.
64. Jaeger S, Lima R, Meyroneinc A, et al. A dynamical model of TCR β gene recombination: Coupling the initiation of D β -J β rearrangement to TCR β allelic exclusion. *bioRxiv.* 2019; <https://doi.org/10.1101/200444>
65. Brown SD, Hapgood G, Steidl C, et al. Defining the clonality of peripheral T cell lymphomas using RNA-seq. *Bioinformatics.* 2017;33(8):1111–1115.
66. Muche JM, Marcus Muche J, Lukowsky A, et al. Peripheral Blood T Cell Clonality in Mycosis Fungoides –An Independent Prognostic Marker? *J. Invest. Dermatol.* 2000;115(3):504–505.

67. Scarisbrick JJ, Hodak E, Bagot M, et al. Developments in the understanding of blood involvement and stage in mycosis fungoides/Sezary syndrome. *Eur. J. Cancer*. 2018;101:278–280.
68. Talpur R, Singh L, Daulat S, et al. Long-term outcomes of 1,263 patients with mycosis fungoides and Sézary syndrome from 1982 to 2009. *Clin. Cancer Res*. 2012;18(18):5051–5060.
69. Klicznik MM, Morawski PA, Höllbacher B, et al. Human CD4CD103 cutaneous resident memory T cells are found in the circulation of healthy individuals. *Sci Immunol*. 2019;4(37.):
70. Massagué J, Obenauf AC. Metastatic colonization by circulating tumour cells. *Nature*. 2016;529(7586):298–306.
71. Shiozawa Y, Pedersen EA, Havens AM, et al. Human prostate cancer metastases target the hematopoietic stem cell niche to establish footholds in mouse bone marrow. *J. Clin. Invest*. 2011;121(4):1298–1312.
72. Zhang XH-F, Jin X, Malladi S, et al. Selection of bone metastasis seeds by mesenchymal signals in the primary tumor stroma. *Cell*. 2013;154(5):1060–1073.
73. Singh M, Al-Eryani G, Carswell S, et al. High-throughput targeted long-read single cell sequencing reveals the clonal and transcriptional landscape of lymphocytes. *Nat. Commun*. 2019;10(1):3120.
74. Kim YH, Bagot M, Pinter-Brown L, et al. Mogamulizumab versus vorinostat in previously treated cutaneous T-cell lymphoma (MAVORIC): an international, open-label, randomised, controlled phase 3 trial. *Lancet Oncol*. 2018;19(9):1192–1204.

75. Soler D, Humphreys TL, Spinola SM, Campbell JJ. CCR4 versus CCR10 in human cutaneous TH lymphocyte trafficking. *Blood*. 2003;101(5):1677–1682.

References for chapter 4:

1. Rizvi MA. T-cell non-Hodgkin lymphoma. *Blood*. 2006;107(4):1255–1264.
2. Ansell SM. Non-Hodgkin Lymphoma: Diagnosis and Treatment. *Mayo Clinic Proceedings*. 2015;90(8):1152–1163.
3. Agar NS, Wedgeworth E, Crichton S, et al. Survival outcomes and prognostic factors in mycosis fungoides/Sézary syndrome: validation of the revised International Society for Cutaneous Lymphomas/European Organisation for Research and Treatment of Cancer staging proposal. *J. Clin. Oncol*. 2010;28(31):4730–4739.
4. McGirt LY, Jia P, Baerenwald DA, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. *Blood*. 2015;126(4):508–519.
5. Ungewickell A, Bhaduri A, Rios E, et al. Genomic analysis of mycosis fungoides and Sézary syndrome identifies recurrent alterations in TNFR2. *Nat. Genet*. 2015;47(9):1056–1060.
6. Wang L, Ni X, Covington KR, et al. Genomic profiling of Sézary syndrome identifies alterations of key T cell signaling and differentiation genes. *Nat. Genet*. 2015;47(12):1426–1434.
7. Moffitt AB, Dave SS. Clinical Applications of the Genomic Landscape of Aggressive Non-Hodgkin Lymphoma. *J. Clin. Oncol*. 2017;35(9):955–962.

8. Park J, Yang J, Wenzel AT, et al. Genomic analysis of 220 CTCLs identifies a novel recurrent gain-of-function alteration in RLTPR (p.Q575E). *Blood*. 2017;130(12):1430 LP–1440.
9. da Silva Almeida AC, Abate F, Khiabani H, et al. The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome. *Nat. Genet.* 2015;47(12):1465–1470.
10. Tensen CP. PLCG1 Gene Mutations in Cutaneous T-Cell Lymphomas Revisited. *J. Invest. Dermatol.* 2015;135(9):2153–2154.
11. Prasad A, Rabionet R, Espinet B, et al. Identification of Gene Mutations and Fusion Genes in Patients with Sézary Syndrome. *J. Invest. Dermatol.* 2016;136(7):1490–1499.
12. Kiel MJ, Sahasrabudhe AA, Rolland DCM, et al. Genomic analyses reveal recurrent mutations in epigenetic modifiers and the JAK–STAT pathway in Sézary syndrome. *Nature Communications*. 2015;6(1.):
13. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. *Nature Genetics*. 2015;47(9):1011–1019.
14. Chang L-W, Patrone CC, Yang W, et al. An Integrated Data Resource for Genomic Analysis of Cutaneous T-Cell Lymphoma. *J. Invest. Dermatol.* 2018;138(12):2681–2683.
15. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013;499(7457):214–218.
16. McGranahan N, Rosenthal R, Hiley CT, et al. Allele-Specific HLA Loss and Immune Escape in Lung Cancer Evolution. *Cell*. 2017;171(6):1259–1271.e11.
17. Keenan TE, Burke KP, Van Allen EM. Genomic correlates of response to immune checkpoint blockade. *Nat. Med.* 2019;25(3):389–402.

18. Campbell JJ, Clark RA, Watanabe R, Kupper TS. Sezary syndrome and mycosis fungoides arise from distinct T-cell subsets: a biologic rationale for their distinct clinical behaviors. *Blood*. 2010;116(5):767–771.
19. Hamrouni A, Fogh H, Zak Z, Ødum N, Gniadecki R. Clonotypic Diversity of the T-cell Receptor Corroborates the Immature Precursor Origin of Cutaneous T-cell Lymphoma. *Clin. Cancer Res*. 2019;25(10):3104–3114.
20. Iyer A, Hennessey D, O’Keefe S, et al. Skin colonization by circulating neoplastic clones in cutaneous T-cell lymphoma. *Blood*. 2019;
21. Iyer A, Hennessey D, O’Keefe S, et al. Clonotypic heterogeneity in cutaneous T-cell lymphoma (mycosis fungoides) revealed by comprehensive whole-exome sequencing. *Blood Adv*. 2019;3(7):1175–1184.
22. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics*. 2013;43:11.10.1–33.
23. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol*. 2013;31:213.
24. Kim S, Scheffler K, Halpern AL, et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat. Methods*. 2018;15(8):591–594.
25. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122.
26. Ha G, Roth A, Khattra J, et al. TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. *Genome Res*. 2014;24(11):1881–1893.

27. Deshwar AG, Vembu S, Yung CK, et al. PhyloWGS: Reconstructing subclonal composition and evolution from whole-genome sequencing of tumors. *Genome Biol.* 2015;16(1):35.
28. Martincorena I, Roshan A, Gerstung M, et al. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science.* 2015;348(6237):880–886.
29. Bailey MH, Tokheim C, Porta-Pardo E, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell.* 2018;174(4):1034–1035.
30. Zhu K, Liu Q, Zhou Y, et al. Oncogenes and tumor suppressor genes: comparative genomics and network perspectives. *BMC Genomics.* 2015;16 Suppl 7:S8.
31. Osborne C, Wilson P, Tripathy D. Oncogenes and tumor suppressor genes in breast cancer: potential diagnostic and therapeutic applications. *Oncologist.* 2004;9(4):361–377.
32. Gug G, Huang Q, Chiticariu E, Solovan C, Baudis M. DNA copy number imbalances in primary cutaneous lymphomas. *Journal of the European Academy of Dermatology and Venereology.* 2019;33(6):1062–1075.
33. De Mattos-Arruda L, Sammut S-J, Ross EM, et al. The Genomic and Immune Landscapes of Lethal Metastatic Breast Cancer. *Cell Rep.* 2019;27(9):2690–2708.e10.
34. Williams MJ, Werner B, Barnes CP, Graham TA, Sottoriva A. Identification of neutral tumor evolution across cancer types. *Nat. Genet.* 2016;48(3):238–244.
35. Jamal-Hanjani M, Wilson GA, McGranahan N, et al. Tracking the Evolution of Non-Small-Cell Lung Cancer. *N. Engl. J. Med.* 2017;376(22):2109–2121.
36. Araf S, Wang J, Korfi K, et al. Genomic profiling reveals spatial intra-tumor heterogeneity in follicular lymphoma. *Leukemia.* 2018;32(5):1261–1265.

37. Magnoli F, Tibiletti MG, Uccella S. Unraveling Tumor Heterogeneity in an Apparently Monolithic Disease: BCL2 and Other Players in the Genetic Landscape of Nodal Follicular Lymphoma. *Front. Med.* 2019;6:44.
38. Beà S, Valdés-Mas R, Navarro A, et al. Landscape of somatic mutations and clonal evolution in mantle cell lymphoma. *Proc. Natl. Acad. Sci. U. S. A.* 2013;110(45):18250–18255.
39. de Masson A, O'Malley JT, Elco CP, et al. High-throughput sequencing of the T cell receptor β gene identifies aggressive early-stage mycosis fungoides. *Sci. Transl. Med.* 2018;10(440.):
40. Oka T, Miyagaki T. Novel and Future Therapeutic Drugs for Advanced Mycosis Fungoides and Sézary Syndrome. *Front. Med.* 2019;6:116.
41. Lopez AT, Bates S, Geskin L. Current Status of HDAC Inhibitors in Cutaneous T-cell Lymphoma. *Am. J. Clin. Dermatol.* 2018;19(6):805–819.
42. Zhou H, Neelakantan D, Ford HL. Clonal cooperativity in heterogenous cancers. *Semin. Cell Dev. Biol.* 2017;64:79–89.
43. Caswell DR, Swanton C. The role of tumour heterogeneity and clonal cooperativity in metastasis, immune evasion and clinical outcome. *BMC Med.* 2017;15(1):133.
44. Vinci M, Burford A, Molinari V, et al. Functional diversity and cooperativity between subclonal populations of pediatric glioblastoma and diffuse intrinsic pontine glioma cells. *Nat. Med.* 2018;24(8):1204–1215.
45. Buus TB, Willerslev-Olsen A, Fredholm S, et al. Single-cell heterogeneity in Sézary syndrome. *Blood Adv.* 2018;2(16):2115–2126.

46. Borcherdig N, Voigt AP, Liu V, et al. Single-Cell Profiling of Cutaneous T-Cell Lymphoma Reveals Underlying Heterogeneity Associated with Disease Progression. *Clin. Cancer Res.* 2019;25(10):2996–3005.
47. Litvinov IV, Tetzlaff MT, Thibault P, et al. Gene expression analysis in Cutaneous T-Cell Lymphomas (CTCL) highlights disease heterogeneity and potential diagnostic and prognostic indicators. *Oncoimmunology.* 2017;6(5):e1306618.
48. Gaydosik AM, Tabib T, Geskin LJ, et al. Single-Cell Lymphocyte Heterogeneity in Advanced Cutaneous T-cell Lymphoma Skin Tumors. *Clin. Cancer Res.* 2019;25(14):4443–4454.
49. Gniadecki R, Lukowsky A, Rossen K, et al. Bone marrow precursor of extranodal T-cell lymphoma. *Blood.* 2003;102(10):3797–3799.

References for chapter 5:

1. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood.* 2016;127(20):2375–2390.
2. Olsen E, Vonderheid E, Pimpinelli N, et al. Revisions to the staging and classification of mycosis fungoides and Sezary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the cutaneous lymphoma task force of the European Organization of Research and Treatment of Cancer (EORTC). *Blood.* 2007;110(6):1713–1722.
3. Kim YH, Willemze R, Pimpinelli N, et al. TNM classification system for primary cutaneous lymphomas other than mycosis fungoides and Sezary syndrome: a proposal of the International Society for Cutaneous Lymphomas (ISCL) and the Cutaneous

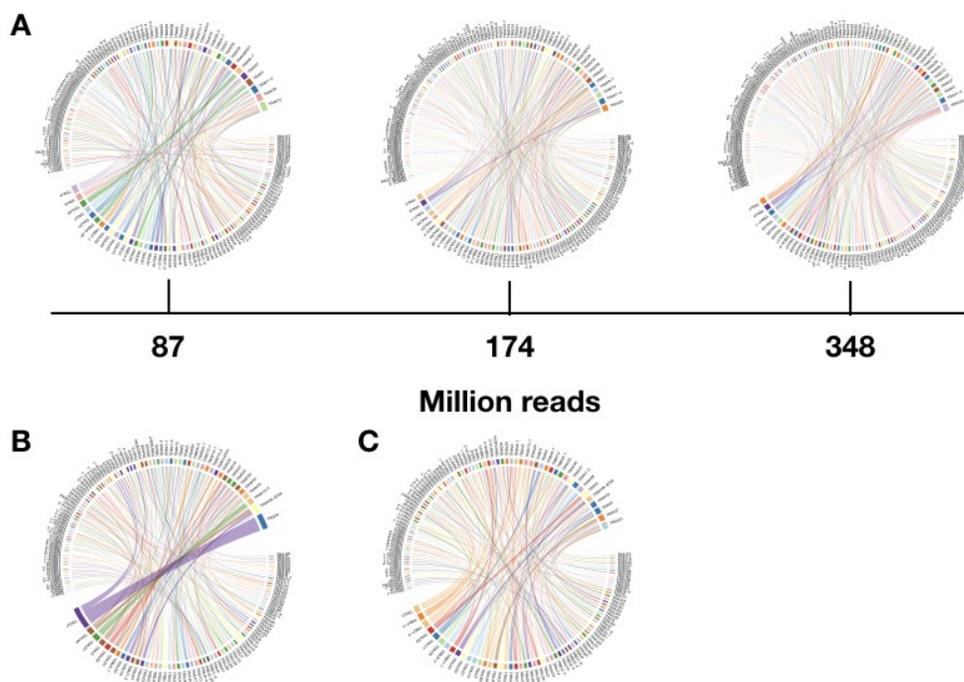
- Lymphoma Task Force of the European Organization of Research and Treatment of Cancer (EORTC). *Blood*. 2007;110(2):479–484.
4. Willemze R. WHO-EORTC classification for cutaneous lymphomas. *Blood*. 2005;105(10):3768–3785.
 5. McGranahan N, Swanton C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell*. 2017;168(4):613–628.
 6. Vogelstein B, Kinzler KW. The Path to Cancer --Three Strikes and You're Out. *N. Engl. J. Med*. 2015;373(20):1895–1898.
 7. Ellsworth DL, Blackburn HL, Shriver CD, et al. Single-cell sequencing and tumorigenesis: improved understanding of tumor evolution and metastasis. *Clin. Transl. Med*. 2017;6(1):15.
 8. Heyde A, Reiter JG, Naxerova K, Nowak MA. Consecutive seeding and transfer of genetic diversity in metastasis. *Proc. Natl. Acad. Sci. U. S. A*. 2019;116(28):14129–14137.
 9. Klemke CD, Booken N, Weiss C, et al. Histopathological and immunophenotypical criteria for the diagnosis of Sézary syndrome in differentiation from other erythrodermic skin diseases: a European Organisation for Research and Treatment of Cancer (EORTC) Cutaneous Lymphoma Task Force Study of 97 cases. *Br. J. Dermatol*. 2015;173(1):93–105.
 10. Massone C, Kodama K, Kerl H, Cerroni L. Histopathologic features of early (patch) lesions of mycosis fungoides: a morphologic study on 745 biopsy specimens from 427 patients. *Am. J. Surg. Pathol*. 2005;29(4):550–560.

11. Chu AC, Morgan EW, MacDonald DM. Ultrastructural identification of T lymphocytes in tissue sections of mycosis fungoides. *J. Invest. Dermatol.* 1980;74(1):17–20.
12. Buus TB, Willerslev-Olsen A, Fredholm S, et al. Single-cell heterogeneity in Sézary syndrome. *Blood Adv.* 2018;2(16):2115–2126.
13. Gaydosik AM, Tabib T, Geskin LJ, et al. Single-Cell Lymphocyte Heterogeneity in Advanced Cutaneous T-cell Lymphoma Skin Tumors. *Clin. Cancer Res.* 2019;25(14):4443–4454.
14. Litvinov IV, Tetzlaff MT, Thibault P, et al. Gene expression analysis in Cutaneous T-Cell Lymphomas (CTCL) highlights disease heterogeneity and potential diagnostic and prognostic indicators. *Oncoimmunology.* 2017;6(5):e1306618.
15. Iyer A, Hennessey D, O’Keefe S, et al. Clonotypic heterogeneity in cutaneous T-cell lymphoma (mycosis fungoides) revealed by comprehensive whole-exome sequencing. *Blood Adv.* 2019;3(7):1175–1184.
16. Iyer A, Hennessey D, O’Keefe S, et al. Skin colonization by circulating neoplastic clones in cutaneous T-cell lymphoma. *Blood.* 2019;blood.2019002516.
17. Iyer A, Hennessey D, O’Keefe S, et al. Branched evolution and genomic intratumor heterogeneity in the pathogenesis of cutaneous T-cell lymphoma.
18. Bolotin DA, Poslavsky S, Mitrophanov I, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods.* 2015;12:380.
19. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinformatics.* 2013;43:11.10.1–33.

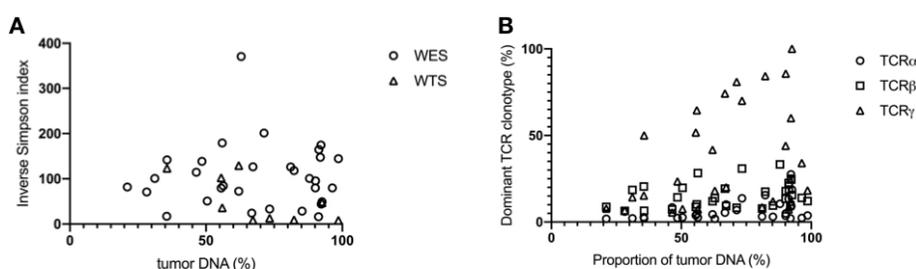
20. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* 2013;31:213.
21. Kim S, Scheffler K, Halpern AL, et al. Strelka2: fast and accurate calling of germline and somatic variants. *Nat. Methods.* 2018;15(8):591–594.
22. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol.* 2016;17(1):122.
23. Ha G, Roth A, Khattra J, et al. TITAN: inference of copy number architectures in clonal cell populations from tumor whole-genome sequence data. *Genome Res.* 2014;24(11):1881–1893.
24. Deshwar AG, Vembu S, Yung CK, et al. PhyloWGS: Reconstructing subclonal composition and evolution from whole-genome sequencing of tumors. *Genome Biol.* 2015;16(1):35.
25. Krangel MS. Mechanics of T cell receptor gene rearrangement. *Curr. Opin. Immunol.* 2009;21(2):133–139.
26. Williams MJ, Werner B, Barnes CP, Graham TA, Sottoriva A. Identification of neutral tumor evolution across cancer types. *Nat. Genet.* 2016;48(3):238–244.
27. Bissell MJ, Labarge MA. Context, tissue plasticity, and cancer: are tumor stem cells also regulated by the microenvironment? *Cancer Cell.* 2005;7(1):17–23.
28. Anderson ARA, Weaver AM, Cummings PT, Quaranta V. Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell.* 2006;127(5):905–915.
29. Polyak K, Haviv I, Campbell IG. Co-evolution of tumor cells and their microenvironment. *Trends Genet.* 2009;25(1):30–38.

30. Zhang L, Zhang S, Yao J, et al. Microenvironment-induced PTEN loss by exosomal microRNA primes brain metastasis outgrowth. *Nature*. 2015;527(7576):100–104.
31. Sottoriva A, Kang H, Ma Z, et al. A Big Bang model of human colorectal tumor growth. *Nat. Genet*. 2015;47(3):209–216.
32. Zhou H, Neelakantan D, Ford HL. Clonal cooperativity in heterogenous cancers. *Semin. Cell Dev. Biol*. 2017;64:79–89.
33. Templeton AR. The reality and importance of founder speciation in evolution. *Bioessays*. 2008;30(5):470–479.
34. Schluter D. Evidence for ecological speciation and its alternative. *Science*. 2009;323(5915):737–741.
35. Vonderheid EC, Pavlov I, Delgado JC, et al. Prognostic factors and risk stratification in early mycosis fungoides. *Leuk. Lymphoma*. 2014;55(1):44–50.
36. Gniadecki R, Lukowsky A, Rossen K, et al. Bone marrow precursor of extranodal T-cell lymphoma. *Blood*. 2003;102(10):3797–3799.
37. Hamrouni A, Fogh H, Zak Z, Ødum N, Gniadecki R. Clonotypic Diversity of the T-cell Receptor Corroborates the Immature Precursor Origin of Cutaneous T-cell Lymphoma. *Clin. Cancer Res*. 2019;25(10):3104–3114.

Appendix A



Appendix Figure A1: VJ gene usage. (A) Peripheral blood mononuclear cells (MF5 PBMC) sample was used for whole exome capture and sequenced at a maximum depth of 400 million reads. Sequenced samples were analyzed for V and J gene combination and no preferential combination was identified as opposed to the high frequency VJ combination identified in (B) that can be associated with the presence of dominant clone in a MF5_1T sample. (C) VJ gene usage of MF5_2P.



Appendix Figure A2: Correlation of inverse Simpson index and dominant clone frequency against tumor enrichment. Microdissected island containing atypical lymphoma cells were subjected to WES/WTS. (A) Inverse Simpson index reflects the TCR repertoire richness. (B) Indicates the most frequent most clonotype for TCR α , - β and - γ . Proportion of tumor-derived DNA in the sample was calculated based on copy number aberration analysis

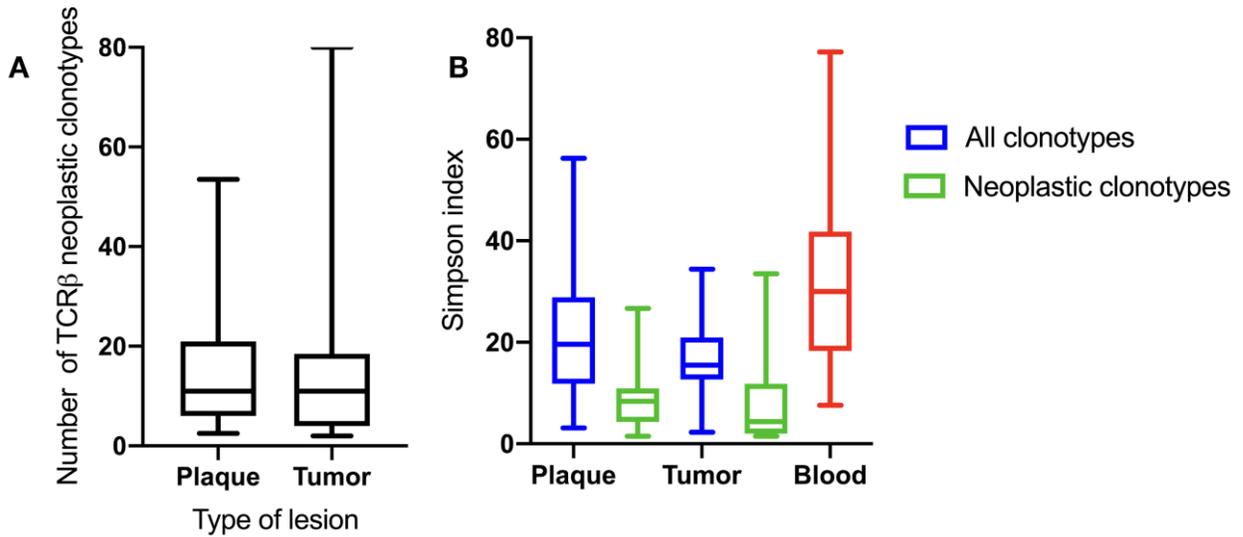
from WES. Symbols represent values for individual samples.

Appendix Table A1: Samples included in the study, with patient age, sex and diagnosis.

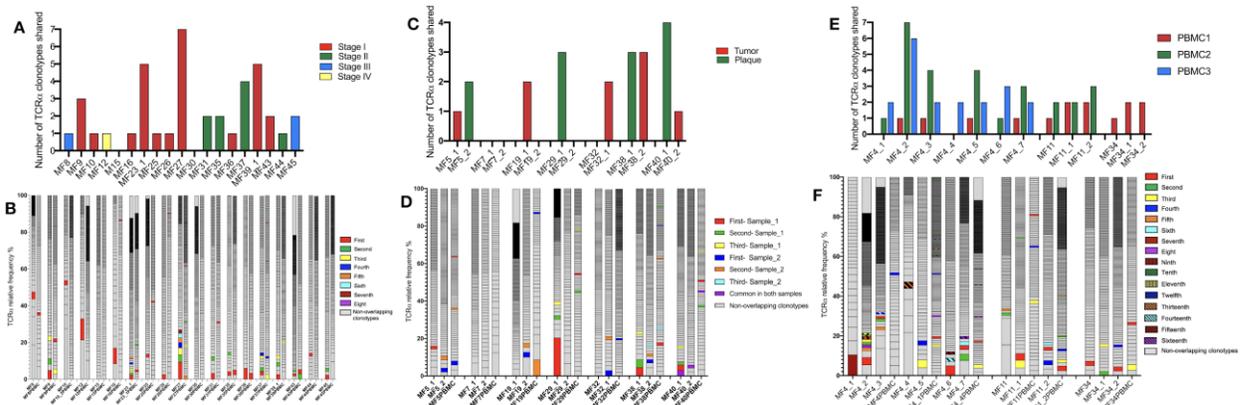
Patient ID (age [years], sex [M-male, F-female])	Sample ID	Lesion type	Diagnosis and stage
MF1 (78, F)	MF1T	Tumor	Mycosis Fungoides IIB
MF2 (83, M)	MF2T	Tumor	Mycosis Fungoides IIB
MF4 (69, M)	MF4_1P	Plaque	Mycosis Fungoides IIB
	MF4_2T	Tumor	
	MF4_3P	Plaque	
MF5 (44, F)	MF5_1T	Tumor	Folliculotropic Mycosis Fungoides IIB
	MF5_2P	Plaque	
MF7 (62, M)	MF7_1T	Tumor	Mycosis Fungoides IVA
	MF7_2P	Plaque	
MF8 (54, F)	MF8P	Plaque	Folliculotropic Mycosis Fungoides IIIB
MF9 (42, F)	MF9P	Plaque	Mycosis Fungoides IA
MF10 (56, M)	MF10P	Plaque	Mycosis Fungoides IB
MF11 (56, M)	MF11T	Tumor	Mycosis Fungoides IIB
	MF11_1P	Plaque	
MF12 (66, M)	MF12P	Plaque	Mycosis Fungoides IVA
MF15 (65, M)	MF15P	Plaque	Mycosis Fungoides IB

MF16 (68, M)	MF16P	Plaque	Mycosis Fungoides IB
MF19 (74, M)	MF19_1T	Tumor	Mycosis Fungoides IIB
	MF19_2P	Plaque	
MF20 (70, M)	MF20P	Plaque	Mycosis Fungoides IB
MF25 (48, F)	MF25P	Plaque	Mycosis Fungoides IB
MF26 (76, M)	MF26P	Plaque	Mycosis Fungoides IB
MF27 (71, M)	MF27P	Plaque	Mycosis Fungoides IA
MF30 (62, M)	MF30P	Plaque	Mycosis Fungoides IB
MF31 (67, M)	MF31T	Tumor	Folliculotropic Mycosis Fungoides IIB
MF33 (75, M)	MF33T	Tumor	Mycosis Fungoides IIB
MF35 (54, M)	MF35P	Plaque	Folliculotropic Mycosis Fungoides IIB
MF36 (64, M)	MF36P	Plaque	Mycosis Fungoides IA
MF37 (63, M)	MF37		Mycosis Fungoides IIB
MF39 (71, M)	MF39_1P	Plaque	Mycosis Fungoides IB
MF43 (60, M)	MF43P	Plaque	Folliculotropic Mycosis Fungoides IA
MF44 (85, M)	MF44T	Tumor	Mycosis Fungoides IIB
MF45 (77, M)	MF45P	Plaque	Mycosis Fungoides IIIA

Appendix B

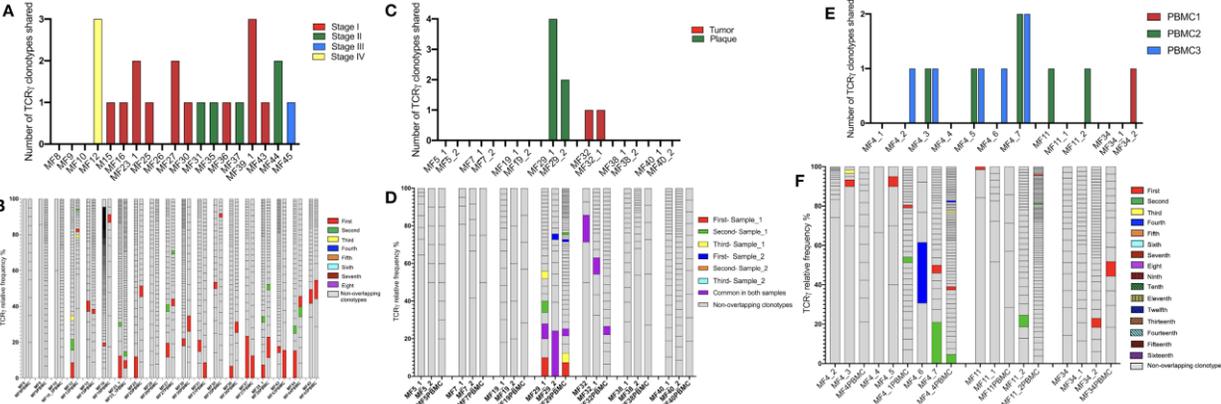


Appendix Figure B1: Comparison of the number of neoplastic TCR β clonotypes and the Simpson diversity in different stages of skin lesions and blood. Box plots represent median value with 25th-75th percentiles; the whiskers show the largest value within 1.5 times the interquartile range below the 25th percentile or above the 75th percentile. **(A)** represents the number of neoplastic TCR β clonotypes in stages of skin lesions (plaque and tumor). **(B)** Indicates the Simpson diversity index of neoplastic TCR β clonotypes and all TCR β clonotypes in skin lesions and blood.

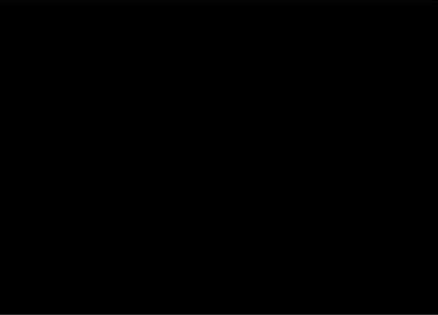


Appendix Figure B2: Shared neoplastic TCR α clonotypes in the skin and the peripheral blood. The sequences of the TCR α clonotypes in the blood were matched with the sequences of the top 10 neoplastic TCR α clonotypes identified in the corresponding skin sample to identify the neoplastic clonotypes in the circulation. The number and frequency of those shared neoplastic clonotypes are shown separately for the three groups of patients as defined in Fig 1: **(A, B)** 19 patients with a single biopsy, **(C-F)** patients with multiple skin biopsies, of whom in 7 patients the biopsies were obtained at a single time point **(C, D)** whereas 3 patients

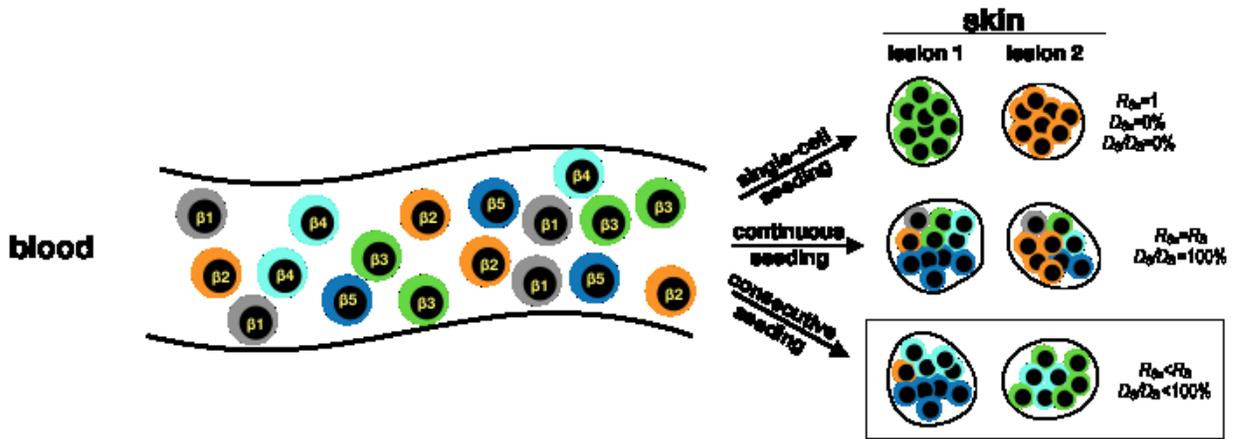
were sampled longitudinally (E, F). In B, D, F the first ranking shared clonotype in the skin is indicated in red and the subsequent shared clonotypes are color-coded as indicated in the legend. The non-overlapping clonotypes are indicated in gray.



Appendix Figure B3: Shared neoplastic TCR γ clonotypes in the skin and the peripheral blood. The sequences of the TCR γ clonotypes in the blood were matched with the sequences of the top 10 neoplastic TCR γ clonotypes identified in the corresponding skin sample to identify the neoplastic clonotypes in the circulation. The number and frequency of those shared neoplastic clonotypes are shown separately for the three groups of patients as defined in Fig 1: (A, B) 19 patients with a single biopsy, C-F: patients with multiple skin biopsies, of whom in 7 patients the biopsies were obtained at a single time point (C, D) whereas 3 patients were sampled longitudinally (E, F). In B, D, F the first ranking shared clonotype in the skin is indicated in red and the subsequent shared clonotypes are color-coded as indicated in the legend. The non-overlapping clonotypes are indicated in gray.



Appendix Figure B4: Correlation between the number of TCR α and TCR β clonotypes in the skin lesions of MF. Pooled data from all 46 samples. The regression coefficient is 0.786. Note that all clonotypes are shown.



Appendix Figure B5: Models of cancer seeding. The simplest scenario is the **single-cell seeding (upper row)** when single neoplastic clonotypes (as detected by unique TCR β sequences β_1 - β_5 marked with different colors) enter the target tissue (skin) from the circulation and develop into the lesions. In this case, each lesion represents a single clone (clonotypic richness of any single n -lesion $R_{S_n}=1$) and the Simpson diversity index (the probability of detecting two different clonotypes in a single skin lesion, D_{S_n}) is 0%. Note, that discrete lesions may originate from different clones. During **continuous seeding (middle row)**, all malignant clonotypes will be detectable in any given skin lesion, either because the metastasis is mediated by clusters of malignant cells or because cooperation between all malignant clones is required for tumor growth. In this situation clonotypic richness of each lesion will be the same and equal to the clonotypic richness of all lesions combined ($\sum R_{S_n}$) and average richness of the lesions (R_S) and to the neoplastic cells in the blood ($R_{S_n}=R_{S_{n+1}} \dots = \sum R_{S_n}=R_S=R_B$). The **consecutive seeding** is an intermediate situation when only a portion of circulating clones seeds the lesions (lower row, frame). Here, the clonotypic richness of any single skin lesion will be lower than the richness of all lesions combined or the richness of the entire repertoire of clonotypes in the blood ($R_{S_n} < \sum R_{S_n} \leq R_B$) and the average Simpson index of skin lesions (D_S) will be lower than the entire blood ($D_S < D_B$, $D_S/D_B < 100\%$). This model is compatible with the data reported in this study. Note, that the richness and Simpson index of the neoplastic clonotypes in the blood (R_B , D_B) cannot be determined precisely experimentally because of insufficient sampling due to a low frequency of circulating, neoplastic clones. The model is modified from Ref⁴³ For simplicity, the traffic of cancer clones is marked as unidirectional movement from the blood to the skin, however it is likely that a recirculation of malignant clones between the blood and the tumors takes place as well.^{40,56}

Appendix Table B1: Patient characteristics and samples included in the study

Patient ID (age [years], sex [M-male, F-female])	Sample ID	Lesion type	Diagnosis and stage ¹	Stage progression ²	TTP ³ or PFS ⁴ (months) ³
MF4 (69, M)	MF4_1P	Plaque	Mycosis Fungoides IIB	no	30
	MF4_2T	Tumor			
	MF4_3P	Plaque			
	MF4_4T	Tumor			
	MF4_5P	Plaque			
	MF4_6T	Tumor			
	MF4_7T	Tumor			
MF5 (44, F)	MF5_1T	Tumor	Folliculotropic Mycosis Fungoides IIB	no	31
	MF5_2P	Plaque			
MF7 (62, M)	MF7_1T	Tumor	Mycosis Fungoides IVA2	no	17 DOD ⁵
	MF7_2P	Plaque			
MF8 (54, F)	MF8P	Plaque	Folliculotropic Mycosis Fungoides IIIB	yes	3
MF9 (42, F)	MF9P	Plaque	Mycosis Fungoides IA	no	29
MF10 (56, M)	MF10P	Plaque	Mycosis Fungoides IB	no	29
MF11 (56, M)	MF11T	Tumor	Mycosis	no	29

	MF11_1P	Plaque	Fungoides IIB		
	MF11_2P	Plaque			
MF12 (66, M)	MF12P	Plaque	Mycosis Fungoides IVA2	no	27
MF15 (65, M)	MF15P	Plaque	Mycosis Fungoides IB	no	27
MF16 (68, M)	MF16P	Plaque	Mycosis Fungoides IB	no	26
MF19 (74, M)	MF19_1T	Tumor	Mycosis Fungoides IIB	no	13 DOD ⁵
	MF19_2P	Plaque			
MF23_1 (69, F)	MF23_1P	Plaque	Mycosis Fungoides, IA	no	22
MF25 (48, F)	MF25P	Plaque	Mycosis Fungoides IB	no	21
MF26 (76, M)	MF26P	Plaque	Mycosis Fungoides IB	yes	6
MF27 (71, M)	MF27P	Plaque	Mycosis Fungoides IA	no	16
MF29	MF29_1P	Plaque	Mycosis Fungoides IA	no	15
	MF29_2P	Plaque			
MF30 (62, M)	MF30P	Plaque	Mycosis Fungoides IB	yes	2
MF31 (67, M)	MF31T	Tumor	Folliculotropic Mycosis Fungoides IIB	no	19
MF32	MF32T	Tumor	Mycosis Fungoides IVA	no	27
	MF32_1T	Tumor			

MF34	MF34T	Tumor	Mycosis Fungoides IIB	no	20
	MF34_1P	Plaque			
	MF34_2T	Tumor			
MF35 (54, M)	MF35P	Plaque	Folliculotropic Mycosis Fungoides IIB	no	11
MF36 (64, M)	MF36P	Plaque	Mycosis Fungoides IA	no	12 DOD ⁵
MF37 (63, M)	MF37	Tumor	Mycosis Fungoides IIB	no	16
MF38	MF38_1P	Plaque	Mycosis Fungoides IIB	no	15
	MF38_2T	Tumor			
MF39 (71, M)	MF39_1P	Plaque	Mycosis Fungoides IB	yes	12 DOD ⁵
MF40	MF40_1P	Plaque	Mycosis Fungoides IIB	no	29
	MF40_2T	Tumor			
MF43 (60, M)	MF43P	Plaque	Folliculotropic Mycosis Fungoides IA	no	15
MF44 (85, M)	MF44T	Tumor	Mycosis Fungoides IIB	no	23 DOD ⁵
MF45 (77, M)	MF45P	Plaque	Mycosis Fungoides IIIA	yes	9 DOD ⁵

¹established at the time of the biopsy; ²observed between the time of the biopsy and last follow-up; ³TTP - only for patients who progressed in stage during follow-up (months); ⁴PFS - progression free survival, for the patients who remained in the same stage during follow-up (months); ⁵DOD - dead of disease

Appendix Table B2: CDR3aa sequences and the predicted V and J combinations for the TCR clonotypes that are shared between at least two skin biopsies (including across different MF patients) or skin biopsies and peripheral blood.

CDR3 sequence	Samples	V usage	J usage
CDNNNDMRF	MF4_2, MF4_3, MF4_1PBMC, MF4_6, MF4_7, MF4_4PBMC, MF9, MF9PBMC, MF11_1, MF11PBMC, MF11_2PBMC, MF23_1, MF23_1PB, MF29_1, MF29PBMC, MF34_1, MF34PBMC, MF38_1, MF38PBMC, MF40_1, MF40_2, MF40PBMC	TRAV16	TRAJ43
CAG_AGA	MF4_2, MF4_3, MF4PBMC, MF4_1PBMC, MF4_7, MF4_4PBMC, MF11_1, MF11PBMC, MF11_2, MF11_2PBMC, MF27, MF27PBMC, MF32_1, MF32_1PBMC, MF34_1, MF34PBMC, MF39_1, MF39PBMC, MF40_1, MF40PBMC, MF43, MF43PBMC	TRAV35	TRAJ4
CGCENSGGSNYKLTF	MF5_2, MF5PBMC, MF11_2, MF11PBMC, MF11_2PBMC, MF16, MF16PBMC, MF19_1, MF19PBMC, MF27, MF27PBMC, MF37, MF37PBMC, MF38_1, MF38PBMC,	TRAV3	TRAJ53
CYEASHSVNTGTASKLTF	MF4_2, MF4_3, MF4_1PBMC, MF9, MF9PBMC, MF23_1, MF23_1PBMC, MF38_1, MF38PBMC	TRAV31	TRAJ44
CMIVGSPELTF	MF4_2, MF4_1PBMC, MF26, MF26PBMC, MF36, MF36PBMC, MF43, MF43PBMC	TRAV35	TRAJ10
CAGPCVFP*RREICYL_L NDALMWAKVNPFSPL	MF4_2, MF4_5, MF4_1PBMC, MF4_4PBMC, MF8, MF8PBMC, MF31, MF31PBMC	TRAV25	TRAJ48
CAV*G_LTNFS	MF5_2, MF5PBMC, MF11,	TRAV36	TRAJ21

	MF11PBMC, MF11_2PBMC, MF27, MF27PBMC,		
IGEWELVLPPTK_LTRF LPAEQLPF	MF23_1, MF23_1PBMC, MF25, MF25PBMC, MF45, MF45PBMC	TRAV13-2	TRAJ46
CALVTAFLEPFF	MF4_7, MF4_1PBMC, MF4_4PBMC, MF27, MF27PBMC	TRAV39	TRAJ12
CGCDLHHRGSWK_VQG DEGLDKLHF	MF4_2, MF4_4PBMC, MF5_1, MF5PBMC	TRAV8-4	TRAJ34
CALRCLKSILNYNDHSI GTF	MF4_2, MF4_1PBMC, MF37, MF37PBMC,	TRAV41	TRAJ3
CAVRSNWLECVRDGW EGGRYF*DGGSPCLEES IHQLHPQCA	MF35, MF35PBMC, MF39_1, MF39PBMC	TRAV21	TRAJ5
CAVKNTGLTAPC_SSSR YLPSRDF	MF4_5, MF4PBMC, MF4_1PBMC, MF4_4PBMC	TRAV3	TRAJ50
WTPY*E_F*ETLF	MF4_1, MF4_4PBMC	TRAV8-3	TRAJ8
CVASGC*PRFPPLHSVSP LSTPGAGTAPALIPQSS*S LRQTF	MF4_2, MF4_1PBMC	TRAV12-1	TRAJ48
CRE_IIF	MF4_3, MF4_1PBMC	TRAV7	TRAJ30
CAVRDVTMRPL*L_PK FKRQPQQREAVF	MF4_4, MF4_4PBMC	TRAV1-2	TRAJ51
CSPEGPWPSLEPRT*VHL T*S_GIRIPEFSQPGVQNS RVRL	MF4_6, MF4_4PBMC	TRAV34	TRAJ24
CAARHPPSWRQIIGSPRT EPLGVP_LPRPGQGRRA AGLLHLQHQLW	MF4_6, MF4_4PBMC	TRAV13-1	TRAJ33
CAVEG_EHLPF	MF4_4, MF4_4PBMC	TRAV36	TRAJ5
CALS_RR*G	MF9, MF9PBMC	TRAV19	TRAJ4
FERNAILFIESLKACFTC LFLKV*_EGVQQGCDR SSKN*NSFIYGHILF	MF10, MF10_2PBMC	TRAV5	TRAJ14

CVWRQGVPPESPGLVRI SP_GTQKRKGQSREWGS DHSQQ	MF11, MF11_2PBMC	TRAV29	TRAJ57
CGSNRYSS*LFH*CQ**S LKWRSC_LSIDKVLTVFL FQG*IGSESSLVF	MF11_2, MF11_2PBMC	TRAV34	TRAJ47
CVCECS*CVYICVCVCV YV_VFVTDLIPVLSAGSI LTC	MF12, MF12PBMC	TRAV10	TRAJ58
WQGLLLTGEERTSWERT	MF19_1, MF19PBMC	TRAV27	TRAJ18
SDSFRQSLLGSRGRSSL SLAKSVSTTNIAG*ASV W	MF23_1, MF23_1PBMC	TRAV36	TRAJ55
CAL_AF	MF23_1, MF23_1PBMC	TRAV15	TRAJ27
CALSGTVAGF	MF27, MF27PBMC	TRAV16	TRAJ16
CALRDRVG	MF27, MF27PBMC	TRAV18	TRAJ36
YAVSGLIFSASLGFLQS_ RKLCAYSGAGSYQLTF	MF27, MF27PBMC	TRAV8-4	TRAJ28
SAVGPAQDPPCHRQWG PCVHP_TLRPSGTRPPM WTGVWVGSP	MF29_1, MF29PBMC	TRAV25	TRAJ30
CLVGLLMVTAHTTSPWL RVLIWRACRGIPGMRA SG	MF29_1, MF29PBMC	TRAV4	TRAJ30
CAGPCVFP*RREICYL_L NDALMWAKVNPFSPL	MF31, MF31PBMC	TRAV25	TRAJ44
CALRDRVGGTAARAQ_T NPGERRWGLGVSFR	MF31, MF31PBMC	TRAV18	TRAJ14
CCVGI_TNFS	MF32_1, MF32_1PBMC	TRAV28	TRAJ21
FMVNVGTGELARQHIAL SSL*PLHFFGQVKR**DL *L*LDARHTAS*FLKFFF	MF34, MF34_2, MF34PBMC	TRAV3	TRAJ36
CAGQHSAPQT_LQPVLK LAA	MF34_2, MF34PBMC	TRAV35	TRAJ8

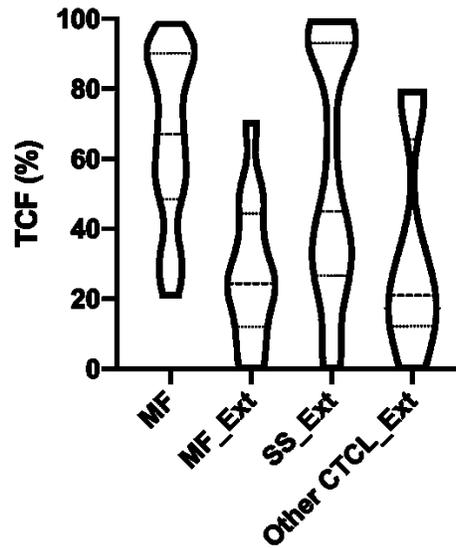
CVVSEMFLSLLISKIFQG V*PQDITH_GPYAGGSLN RSN*EK*KTESKILIF	MF37, MF37PBMC	TRAV8-2	TRAJ37
VFYVSSHLLGVPLSYVPA LFS DG_LLLREGVALTVI VYQNLLYLTF	MF37, MF37PBMC	TRAV8-2	TRAJ2
CVVSGVSSGLLVPRLGIG WFEQTLGSGLI	MF38_2, MF38PBMC	TRAV10	TRAJ4
SASCTMRSISSLLRRPLS LVMVILLALPVLLSVAV TFRMPLASMSKVTSIW	MF38_2, MF38PBMC	TRAV3	TRAJ33
CAESTHCFSGTCILYPNL HLGLKPHSISFTF	MF38_2, MF38PBMC	TRAV5	TRAJ13
CAATTSGTYKYIF	MF39_1, MF39PBMC	TRAV21	TRAJ40
CVGVFQHGKVEIANDQ GN_HHPQLRGLHRHRAA HWGC	MF40_1, MF40PBMC	TRAV12-1	TRAJ54
*PGGCLSLFIRDWPRNL L_GHWGLRLQALASSRP RGI	MF40_1, MF40PBMC	TRAV1-2	TRAJ9
CLLGDFPSLGLFLMWW* IHGS*RP_CGFERP*AGA PG*T*IYLLVPPF	MF44, MF44PBMC	TRAV40	TRAJ22
CAGPCSWTRY_AVSSH LTF	MF45, MF45PBMC	TRAV25	TRAJ46
		TCR beta clonotypes	
CAASRGC_AKNIQYF	MF4_2, MF4_3, MF4_5, MF4_1PBMC, MF4_6, MF4_7, MF4_4PBMC, MF5_1, MF5PBMC, MF11, MF11_1, MF11_2, MF11_2PBMC, MF12, MF12PBMC, MF23_1, MF23_1PBMC, MF25, MF25PBMC, MF26, MF26PBMC, MF27, MF27PBMC, MF30, MF30PBMC, MF31, MF31PBMC, MF32, MF32_1, MF32PBMC,	TRBV18	TRBJ2-4

	MF34, MF34_1, MF34_2, MF34PBMC, MF35, MF35PBMC, MF36, MF36PBMC, MF37, MF37PBMC, MF38_1, MF38_2, MF38PBMC, MF39_1, MF39PBMC, MF40_1, MF40_2, MF40PBMC, MF44, MF44PBMC, MF45, MF45PBMC		
CVS_GVL	MF25, MF25PBMC, MF31, MF31PBMC, MF40_1, MF40PBMC	TRBV5-3	TRBJ2-4
CASSEATALHG	MF30, MF30PBMC, MF34_2, MF34PBMC, MF37, MF37PBMC	TRBV6-1	TRBJ2-2P
CAKQLNEDCSKTQPCDH TKGLECNF	MF4_5, MF4_1PBMC, MF4_4PBMC	TRBV23-1	TRBJ2-4
GPGTRLLVLGERGLLGR GRGR_WVWFLRGVPGL CSGANVLTF	MF4_6, MF4_7, MF4_1PBMC	TRBV13	TRBJ2-6
CGGGARKTGQSPRQPRP GPSF	MF10, MF10_2PBMC, MF11_2, MF11_2PBMC	TRBV22-1	TRBJ2-4
CLARRQR_TTVVVF	MF4_5, MF4_1PBMC, MF35, MF35PBMC	TRBV5-6	TRBJ2-2
CELECIWPGSRNFECRR GISETVDKTAWKKTERC *SQTG	MF4_2, MF4_4PBMC	TRBV30	TRBJ2-7
CGRCSALP**RCGGRIPC SPRSS_NIVVEARGLVPT HTA*LVPQHF	MF4_2, MF4_4PBMC	TRBV8-2	TRBJ1-5
CAT_AR	MF4_3, MF4_1PBMC	TRBV27	TRBJ2-2P
CAQQPVQSELVQRCQQL QSRLSTLKI_KRRGEFLP* EAQSHTLSFPVETFAAA	MF4_4, MF4_4PBMC	TRBV5-4	TRBJ2-2P
CASCPH_VSCRRP	MF4_6, MF4_1PBMC	TRBV4-2	TRBJ2-4
CAS*LWW_*KSV**A	MF4_6, MF4_4PBMC	TRBV6-2	TRBJ1-3
*AAAWPAAAG_LRRVG	MF9, MF9PBMC	TRBV5-5	TRBJ2-2P

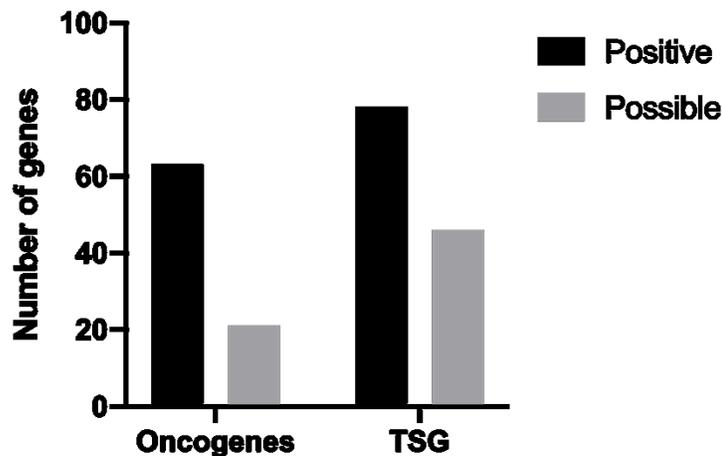
WLAG			
FLRLLKYIRKLSG_QLRV SSSLNKPS	MF11, MF11PBMC	TRBV22-1	TRBJ1-4
CARVPRAV_NTGELFF	MF11, MF11_2PBMC	TRBV5-1	TRBJ2-2
CASSYSNQPQHF	MF12, MF12PBMC	TRBV19	TRBJ1-5
CAI*S*HPAS_KSSHAYSL FF	MF23_1, MF23PBMC	TRBV10-3	TRBJ2-2
CGSSEEGT	MF30, MF30PBMC	TRBV8-2	TRBJ1-6
CASSF	MF37, MF37PBMC	TRBV12-1	TRBJ2-2
CTTMR*QSR*VRARAGG RAAC_GGDFRLSMRFPA PGRSTEHF	MF38_1, MF38PBMC	TRBV5-6	TRBJ2-7
CASSLD*REETDTQYF	MF39_1, MF39PBMC	TRBV5-6	TRBJ2-3
CASSEPGRNQPQHF	MF39_1, MF39PBMC	TRBV6-1	TRBJ1-5
CASSQDTALQSHCIPVH KPPGSARKLQGSV_APA PRAPVSIP*WPLMEFQSV VQPASAPS	MF40_2, MF40PBMC	TRBV3-1	TRBJ2-2
		TCR gamma clonotypes	
CAAWDYH_GWFKIF	MF4_1, MF4_7, MF4_4PBMC, MF12, MF12PBMC, MF27, MF27PBMC, MF29_1, MF29PBMC, MF30, MF30PBMC, MF31, MF31PBMC, MF32, MF32_1, MF32_1PBMC, MF35, MF35PBMC, MF36, MF36PBMC, MF37, MF37PBMC, MF39_1, MF39PBMC, MF43, MF43PBMC, MF44, MF44PBMC, MF45, MF45PBMC	TRGV10	TRGJP1
CATAAGLL_WL*SLLP*	MF4_3, MF4_5, MF4_7, MF4_1PBMC, MF4_4PBMC	TRGV5P	TRGJ1
STAVGLA_SQAVQDN	MF4_2, MF4_4PBMC, MF11_2, MF11_2PBMC, MF27,	TRGVB	TRGJP1

	MF27PBMC		
CATAAGLLVVVVF	MF4_6, MF4_4PBMC, MF11, MF11_2PBMC, MF23_1, MF23_1PBMC	TRGV5P	TRGJP1
CATWENR_GWFKIF	MF12, MF12PBMC, MF29_1, MF29PBMC	TRGV8	TRGJP1
CAAWDYGTGWFKIF	MF12, MF12PBMC, MF29_1, MF29_2, MF29PBMC, MF44, MF44PBMC	TRGV10	TRGJP1
CAAWDYTK_TTGWFKIF	MF15, MF15PBMC	TRGV10	TRGJP1
CATWVLP_GWFKIF	MF16, MF16PBMC	TRGV5	TRGJP1
CATRT_YKKLF	MF23_1, MF23_1PBMC, MF29_1, MF29PBMC	TRGV8	TRGJ1
CATWD_TRELF	MF25, MF25PBMC	TRGV2	TRGJ1
CACWIRH_GDWIKTF	MF29_2, MF29PBMC	TRGV11	TRGJP2
STAVGLA_KSGSSR*	MF34_2, MF34PBMC	TRGVB	TRGJP1
CATWDG_YYKKLF	MF39_1, MF39PBMC	TRGV4	TRGJ1
CAAWDYGTGWFKIF	MF44, MF44PBMC	TRGV10	TRGJP1

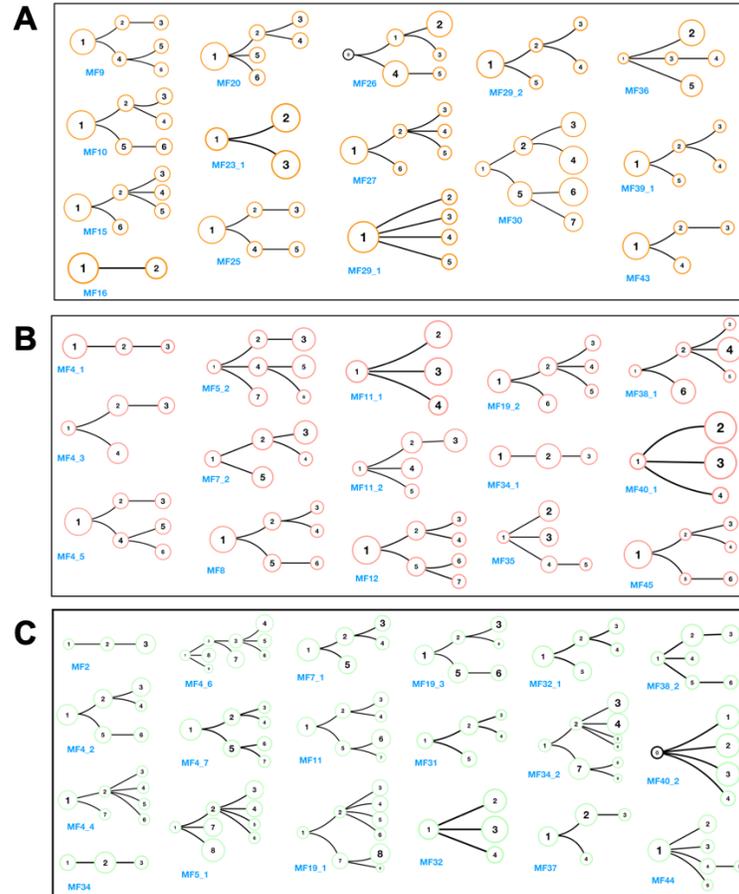
Appendix C



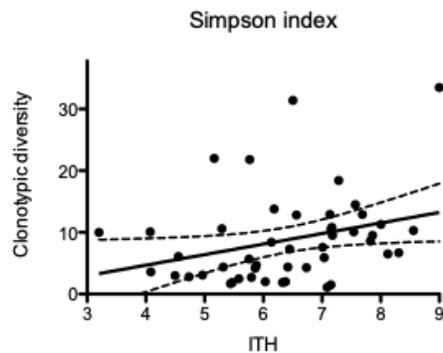
Appendix Figure C1: Tumor cell fraction (TCF) in analyzed CTCL samples. The TCF was calculated for all samples in our study and the samples published in three other CTCL studies (see supplementary Table S2). The samples from the previously CTCL studies were classified as mycosis fungoides (MF), Sezary syndrome (SS) and other CTCL based on the information provided in the original publications (labelled with _Ext for external datasets). The violin plot represents the overall distribution of TCF in the samples.



Appendix Figure C2: Functional classification of putative driver genes. Mutations identified in putative driver genes were classified as oncogenes or Tumor suppressor genes (TSG) based on their function with previously reported studies. “Possible” classifies the genes that present a functional ambiguity.

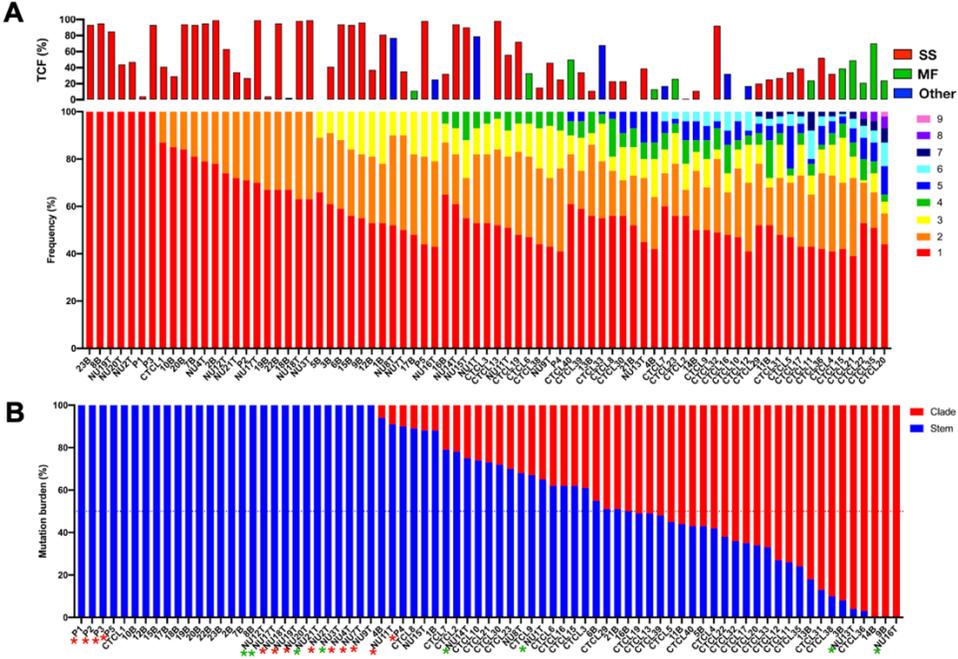


Appendix Figure C3: Phylogenetic structures for samples in ESP, LSP and TMR. (A-C) Presents the phylogenetic trees for every sample in ESP, LSP and TMR respectively. The sample IDs are presented in blue text below each phylogenetic tree.



Appendix Figure C4: Simpson heterogeneity index for clonotypic and intratumoral heterogeneity (ITH) in MF. Simpson index for clonotypic heterogeneity was calculated as previously described¹, the ITH Simpson index was calculated and 1D where $D = N(N-1)/n(n-1)$

(N - total number of subclones, n - the frequency of the individual subclone). A regression line with 95% CI is plotted, $R^2=0.11$, $p=0.028$.



Appendix Figure C5: Intratumoral heterogeneity in CTCL. Data obtained from previous sequencing studies (supplementary Table S2) were analyzed by the same bioinformatic pipeline as in Figure 4A and Figure 5A. **(A)** Rainbow graph representing the number and frequency of the subclones identified in each sample. The samples are arranged by increasing number of subclones, followed by the relative frequency of the most abundant subclone. The top bar graph shows TCF for each sample; the colour of the bars indicates the disease (MF, mycosis fungoides, SS, Sézary syndrome, Other, CTCL unspecified). **(B)** Distribution of mutations in the stem and clades. Percentage of all SNV mutations in the stem (blue) and clade (red) of the phylogenetic trees. Asterisk next to the sample names, in red indicate the MF samples and the green indicates CTCL not specified.

Appendix Table C1: Patient characteristics and samples included in the study

Patient ID (age [years], sex [M-male, F-female])	Sample ID	Lesion type	Diagnosis and stage
MF2(83, M)	MF2	Tumor	Mycosis Fungoides IIB
MF4 (69, M)	MF4_1P	Plaque	Mycosis Fungoides IIB
	MF4_2T	Tumor	
	MF4_3P	Plaque	
	MF4_4T	Tumor	
	MF4_5P	Plaque	
	MF4_6T	Tumor	
	MF4_7T	Tumor	
MF5 (44, F)	MF5_1T	Tumor	Folliculotropic Mycosis Fungoides IIB
	MF5_2P	Plaque	
MF7 (62, M)	MF7_1T	Tumor	Mycosis Fungoides IVA2
	MF7_2P	Plaque	
MF8 (54, F)	MF8P	Plaque	Folliculotropic Mycosis Fungoides IIIB
MF9 (42, F)	MF9P	Plaque	Mycosis Fungoides IA
MF10 (56, M)	MF10P	Plaque	Mycosis Fungoides IB

MF11 (56, M)	MF11T	Tumor	Mycosis Fungoides IIB
	MF11_1P	Plaque	
	MF11_2P	Plaque	
MF12 (66, M)	MF12P	Plaque	Mycosis Fungoides IVA2
MF15 (65, M)	MF15P	Plaque	Mycosis Fungoides IB
MF16 (68, M)	MF16P	Plaque	Mycosis Fungoides IB
MF19 (74, M)	MF19_1T	Tumor	Mycosis Fungoides IIB
	MF19_2P	Plaque	
	MF19_3T	Tumor	
MF20 (70, M)	MF20	Plaque	Mycosis Fungoides IB
MF23_1 (69, F)	MF23_1P	Plaque	Mycosis Fungoides, IA
MF25 (48, F)	MF25P	Plaque	Mycosis Fungoides IB
MF26 (76, M)	MF26P	Plaque	Mycosis Fungoides IB
MF27 (71, M)	MF27P	Plaque	Mycosis Fungoides IA
MF29	MF29_1P	Plaque	Mycosis Fungoides IA
	MF29_2P	Plaque	
MF30 (62, M)	MF30P	Plaque	Mycosis Fungoides IB
MF31 (67, M)	MF31T	Tumor	Folliculotropic Mycosis Fungoides IIB

MF32(49, M)	MF32T	Tumor	Mycosis Fungoides IVA
	MF32_1T	Tumor	
MF34(65, M)	MF34T	Tumor	Mycosis Fungoides IIB
	MF34_1P	Plaque	
	MF34_2T	Tumor	
MF35 (54, M)	MF35P	Plaque	Folliculotropic Mycosis Fungoides IIB
MF36 (64, M)	MF36P	Plaque	Mycosis Fungoides IA
MF37 (63, M)	MF37	Tumor	Mycosis Fungoides IIB
MF38(76, M)	MF38_1P	Plaque	Mycosis Fungoides IIB
	MF38_2T	Tumor	
MF39 (71, M)	MF39_1P	Plaque	Mycosis Fungoides IB
MF40(59, F)	MF40_1P	Plaque	Mycosis Fungoides IIB
	MF40_2T	Tumor	
MF43 (60, M)	MF43P	Plaque	Folliculotropic Mycosis Fungoides IA
MF44 (85, M)	MF44T	Tumor	Mycosis Fungoides IIB
MF45 (77, M)	MF45P	Plaque	Mycosis Fungoides IIIA

Appendix Table C2: List of previous CTCL studies used for metaanalysis.

Study	Sample type	Number of samples
McGrit.et.al ²	MF	5
Choi.et.al ³	SS	31
Da Silva Almeida.e ³ t.al ⁴	MF, SS and other CTCL	41

Appendix Table C3: Mutations in putative driver presented in <20% of the samples.

Genes	TMR (%)	LSP (%)	ESP (%)	Gene description	Pathway
TRAF3	2	2	2	TNF receptor associated factor 3	Apoptosis
BCL2	2	0	0	BCL2 apoptosis regulator	Apoptosis
BCL2L11	2	0	0	BCL2 like 11	Apoptosis
CASP8	2	0	0	caspase 8	Apoptosis
CDK4	4	0	2	cyclin dependent kinase 4	Cell cycle
BTG2	0	2	0	BTG anti-proliferation factor 2	Cell cycle
CDKN1A	2	0	0	cyclin dependent kinase inhibitor 1A	Cell cycle
ZMYM3	6	8	4	zinc finger MYM-type containing 3	Chromatin histone modifiers
KANSL1	8	4	4	KAT8 regulatory NSL complex subunit 1	Chromatin histone modifiers
ARID5B	6	2	6	AT-rich interaction domain 5B	Chromatin histone modifiers
KMT2A	8	4	2	lysine methyltransferase 2A	Chromatin histone modifiers

SIN3A	4	2	2	SIN3 transcription regulator family member A	Chromatin histone modifiers
NIPBL	6	6	2	NIPBL cohesin loading factor	Chromatin other
ASXL1	8	0	2	ASXL transcriptional regulator 1	Chromatin other
AJUBA	4	4	0	ajuba LIM protein	Chromatin other
ATF7IP	2	0	2	activating transcription factor 7 interacting protein	Chromatin other
H3F3C	0	4	0	H3.5 histone	Chromatin other
NPM1	0	2	0	nucleophosmin 1	Chromatin other
SMARCA4*	8	4	2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	Chromatin SWI/SNF complex
PBRM1	8	2	2	polybromo 1	Chromatin SWI/SNF complex
ATRX*	6	2	2	ATRX chromatin remodeler	Chromatin SWI/SNF complex
SMARCA1	2	0	2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 1	Chromatin SWI/SNF complex
MSH6	2	6	10	mutS homolog 6	Genome integrity
POLQ	6	8	2	DNA polymerase theta	Genome integrity
SMC1A	4	6	2	structural maintenance of chromosomes 1A	Genome integrity
BRCA2	2	4	4	BRCA2 DNA repair associated	Genome integrity
ERCC2	0	2	8	ERCC excision repair 2, TFIIH core complex helicase subunit	Genome integrity

PPM1D	6	4	0	protein phosphatase, Mg ²⁺ /Mn ²⁺ dependent 1D	Genome integrity
RFC1	4	4	0	replication factor C subunit 1	Genome integrity
BRCA1	2	2	2	BRCA1 DNA repair associated	Genome integrity
CHEK2	2	2	2	checkpoint kinase 2	Genome integrity
STAG2	4	2	0	stromal antigen 2	Genome integrity
ATR	0	2	0	ATR serine/threonine kinase	Genome integrity
SETD2*	6	10	2	SET domain containing 2, histone lysine methyltransferase	Histone modification
SETBP1	6	10	0	SET binding protein 1	Histone modification
RNF111	8	2	2	ring finger protein 111	Immune signaling
IRF6	8	2	0	interferon regulatory factor 6	Immune signaling
HLA-A	0	2	4	major histocompatibility complex, class I, A	Immune signaling
IL7R	2	2	2	interleukin 7 receptor	Immune signaling
HGF	2	2	0	interleukin 6	Immune signaling
B2M*	0	0	2	beta-2-microglobulin	Immune signaling
MAP3K4	6	8	0	mitogen-activated protein kinase kinase kinase 4	MAPK signaling
RPS6KA3	2	2	2	ribosomal protein S6 kinase A3	MAPK signaling
RRAS2	2	0	0	RAS related 2	MAPK signaling
IDH2*	4	2	0	isocitrate dehydrogenase (NADP(+)) 2	Metabolism

IDH1	4	0	0	isocitrate dehydrogenase (NADP(+)) 1	Metabolism
DMD	8	4	6	dystrophin	Other
SPTA1	10	4	4	spectrin alpha, erythrocytic 1	Other
MUC6	4	6	6	mucin 6, oligomeric mucus/gel- forming	Other
GRIN2D	8	2	4	glutamate ionotropic receptor NMDA type subunit 2D	Other
SPTAN1	8	4	0	spectrin alpha, non-erythrocytic 1	Other
FLNA	2	4	4	filamin A	Other
MYH9	4	4	2	myosin heavy chain 9	Other
GABRA6	6	2	0	gamma-aminobutyric acid type A receptor alpha6 subunit	Other
KIF1A	2	2	4	kinesin family member 1A	Other
ALB	2	2	2	Fas binding factor 1	Other
TXNIP	4	2	0	thioredoxin interacting protein	Other
CNBD1	2	2	0	cyclic nucleotide binding domain containing 1	Other
POLRMT	0	2	0	RNA polymerase mitochondrial	Other
COL5A1	14	18	14	collagen type V alpha 1 chain	Other
APOB	12	14	2	apolipoprotein B	Other
CACNA1A	10	8	8	calcium voltage-gated channel subunit alpha1 A	Other

KEL	10	6	6	Kell metallo-endopeptidase (Kell blood group)	Other
PTPRC	8	6	4	protein tyrosine phosphatase receptor type C	Other signaling
ARHGAP35	4	8	4	Rho GTPase activating protein 35	Other signaling
GNAS	10	2	4	GNAS complex locus	Other signaling
PTPDC1	6	4	2	protein tyrosine phosphatase domain containing 1	Other signaling
PTPN11	10	2	0	protein tyrosine phosphatase non-receptor type 11	Other signaling
CDH1	6	0	4	cadherin 1	Other signaling
KEAP1	2	4	2	kelch like ECH associated protein 1	Other signaling
SOS1	0	6	2	SOS Ras/Rac guanine nucleotide exchange factor 1	Other signaling
LEMD2	6	0	0	LEM domain containing 2	Other signaling
NF2	4	2	0	neurofibromin 2	Other signaling
PLCB4	4	0	2	phospholipase C beta 4	Other signaling
GNA11	2	0	2	G protein subunit alpha 11	Other signaling
GNA13	2	2	0	G protein subunit alpha 13	Other signaling
GNAQ	2	0	2	G protein subunit alpha q	Other signaling
PLXNB2	0	2	2	plexin B2	Other signaling
RHOB	2	0	2	ras homolog family member B	Other signaling
DIAPH2	0	2	0	diaphanous related formin 2	Other signaling

GPS2	0	2	0	G protein pathway suppressor 2	Other signaling
MAP2K4	0	0	2	mitogen-activated protein kinase kinase 4	Other signaling
PIM1*	2	0	0	Pim-1 proto-oncogene, serine/threonine kinase	Other signaling
PRKAR1A	2	0	0	protein kinase cAMP-dependent type I regulatory subunit alpha	Other signaling
RAC1	2	0	0	Rac family small GTPase 1	Other signaling
FAT1	22	10	6	FAT atypical cadherin 1	Other signaling
PIK3CG*	8	4	4	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit gamma	PI3K signaling
PIK3R2	6	0	0	phosphoinositide-3-kinase regulatory subunit 2	PI3K signaling
AKT1	2	0	2	AKT serine/threonine kinase 1	PI3K signaling
PPP2R1A	0	2	2	protein phosphatase 2 scaffold subunit Aalpha	PI3K signaling
PIK3CA	2	0	0	phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha	PI3K signaling
USP9X	0	6	8	ubiquitin specific peptidase 9 X-linked	Protein homeostasis/ubiquitination
CYLD	6	2	2	CYLD lysine 63 deubiquitinase	Protein homeostasis/ubiquitination
CUL1	6	0	2	cullin 1	Protein homeostasis/ubiquitination
EEF2	2	0	4	eukaryotic translation elongation factor 2	Protein homeostasis/ubiquitination

BAP1	0	4	0	BRCA1 associated protein 1	Protein homeostasis/ubiquitination
FBXW7	2	2	0	F-box and WD repeat domain containing 7	Protein homeostasis/ubiquitination
SPOP	0	4	0	speckle type BTB/POZ protein	Protein homeostasis/ubiquitination
VHL	0	4	0	von Hippel-Lindau tumor suppressor	Protein homeostasis/ubiquitination
CUL3	0	2	0	cullin 3	Protein homeostasis/ubiquitination
EEF1A1	2	0	0	eukaryotic translation elongation factor 1 alpha 1	Protein homeostasis/ubiquitination
ZFP36L2	4	8	4	ZFP36 ring finger protein like 2	RNA abundance
DHX9	2	8	4	DExH-box helicase 9	RNA abundance
NUP93	8	4	2	nucleoporin 93	RNA abundance
CSDE1	6	2	2	cold shock domain containing E1	RNA abundance
NUP133	2	6	0	nucleoporin 133	RNA abundance
RBM10	2	4	2	RNA binding motif protein 10	RNA abundance
ZC3H12A	2	4	2	zinc finger CCCH-type containing 12A	RNA abundance
SCAF4	2	4	0	SR-related CTD associated factor 4	RNA abundance
DDX3X*	2	0	0	DEAD-box helicase 3 X-linked	RNA abundance
XPO1	0	2	0	exportin 1	RNA abundance
ZFP36L1	0	0	2	ZFP36 ring finger protein like 1	RNA abundance

ERBB2	8	2	4	erb-b2 receptor tyrosine kinase 2	RTK signaling
EGFR	4	8	0	epidermal growth factor receptor	RTK signaling
EPHA3	6	0	6	EPH receptor A3	RTK signaling
FGFR2	4	4	4	fibroblast growth factor receptor 2	RTK signaling
PDGFRA	6	4	2	platelet derived growth factor receptor alpha	RTK signaling
EPHA2	4	2	4	EPH receptor A2	RTK signaling
FGFR1	4	6	0	fibroblast growth factor receptor 1	RTK signaling
FLT3	4	4	0	fms related tyrosine kinase 3	RTK signaling
MET	4	2	2	MET proto-oncogene, receptor tyrosine kinase	RTK signaling
ERBB4	2	4	0	erb-b2 receptor tyrosine kinase 4	RTK signaling
KIT	2	4	0	KIT proto-oncogene, receptor tyrosine kinase	RTK signaling
RET	2	4	0	ret proto-oncogene	RTK signaling
RIT1	2	0	0	Ras like without CAAX 1	RTK signaling
SRSF2	8	8	0	serine and arginine rich splicing factor 2	Splicing
SF3B1	10	2	0	splicing factor 3b subunit 1	Splicing
DAZAP1	2	4	2	DAZ associated protein 1	Splicing
THRAP3	4	4	0	thyroid hormone receptor associated protein 3	Splicing
ACVR1B	2	2	6	activin A receptor type 1B	TGFB signaling

SMAD2	4	0	2	SMAD family member 2	TGFB signaling
SMAD4	4	2	0	SMAD family member 4	TGFB signaling
ACVR1	4	0	0	activin A receptor type 1	TGFB signaling
DACH1	2	2	0	dachshund family transcription factor 1	TGFB signaling
ACVR2A	2	0	0	activin A receptor type 2A	TGFB signaling
TGFBR2	2	0	0	transforming growth factor beta receptor 2	TGFB signaling
TSC2	8	2	0	TSC complex subunit 2	TOR signaling
TSC1	2	0	0	TSC complex subunit 1	TOR signaling
MED12	8	8	2	mediator complex subunit 12	Transcription factor
GATA3	2	8	6	GATA binding protein 3	Transcription factor
FOXA1	4	6	4	forkhead box A1	Transcription factor
PGR	10	2	2	progesterone receptor	Transcription factor
TGIF1	6	4	4	TGFB induced factor homeobox 1	Transcription factor
ZNF750	6	6	2	zinc finger protein 750	Transcription factor
MGA	10	0	2	MAX dimerization protein MGA	Transcription factor
RUNX1	6	4	2	RUNX family transcription factor 1	Transcription factor
WT1	4	6	2	WT1 transcription factor	Transcription factor
KLF5	6	2	2	Kruppel like factor 5	Transcription factor
MYCN	4	2	4	MYCN proto-oncogene, bHLH transcription factor	Transcription factor

EPAS1	4	2	2	endothelial PAS domain protein 1	Transcription factor
TBX3	2	2	4	T-box transcription factor 3	Transcription factor
SOX17	2	2	2	SRY-box transcription factor 17	Transcription factor
UNCX	4	0	2	UNC homeobox	Transcription factor
CREB3L3	0	4	0	cAMP responsive element binding protein 3 like 3	Transcription factor
SOX9	2	2	0	SRY-box transcription factor 9	Transcription factor
TAF1	2	0	2	TATA-box binding protein associated factor 1	Transcription factor
ZCCHC12	2	2	0	zinc finger CCHC-type containing 12	Transcription factor
CBFB	0	2	0	core-binding factor subunit beta	Transcription factor
CEBPA	2	0	0	CCAAT enhancer binding protein alpha	Transcription factor
FOXQ1	2	0	0	forkhead box Q1	Transcription factor
GTF2I	2	0	0	general transcription factor Iii	Transcription factor
MYC*	0	2	0	MYC proto-oncogene, bHLH transcription factor	Transcription factor
NFE2L2	0	0	2	nuclear factor, erythroid 2 like 2	Transcription factor
PSIP1	2	0	0	PC4 and SFRS1 interacting protein 1	Transcription factor
TCF12	0	2	0	transcription factor 12	Transcription factor
ZMYM2	0	0	2	zinc finger MYM-type containing 2	Transcription factor
AXIN1	4	8	2	axin 1	Wnt/B-catenin signaling
CTNNB1	6	2	4	catenin beta 1	Wnt/B-catenin signaling

TCF7L2	2	6	4	transcription factor 7 like 2	Wnt/B-catenin signaling
APC	2	4	2	APC regulator of WNT signaling pathway	Wnt/B-catenin signaling
AXIN2	0	0	2	axin 2	Wnt/B-catenin signaling

* Genes previously reported mutated in CTCL

References

1. Iyer A, Hennessey D, O’Keefe S, et al. Skin colonization by circulating neoplastic clones in cutaneous T-cell lymphoma. *Blood*. 2019;
2. McGirt LY, Jia P, Baerenwald DA, et al. Whole-genome sequencing reveals oncogenic mutations in mycosis fungoides. *Blood*. 2015;126(4):508–519.
3. Choi J, Goh G, Walradt T, et al. Genomic landscape of cutaneous T cell lymphoma. *Nature Genetics*. 2015;47(9):1011–1019.
4. da Silva Almeida AC, Abate F, Khiabani H, et al. The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome. *Nat. Genet*. 2015;47(12):1465–1470.

Appendix D

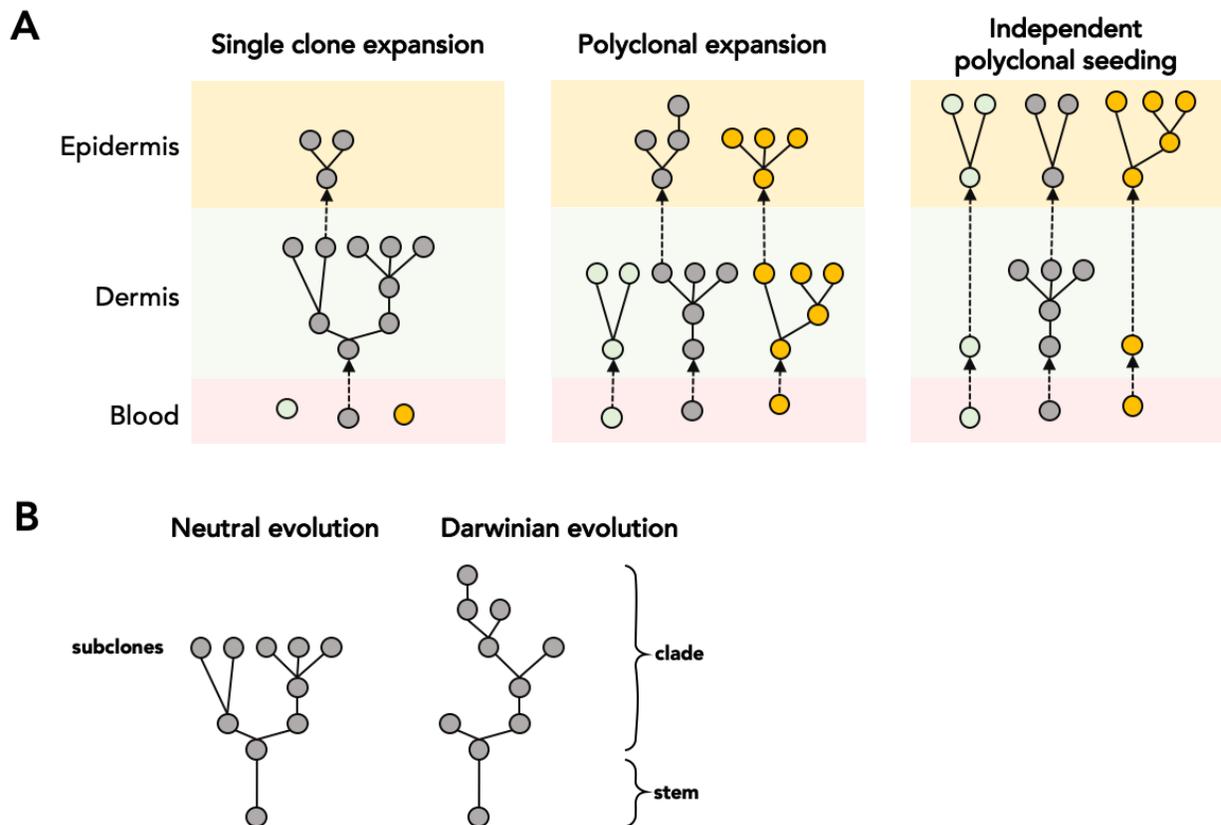


Figure D1: Possible scenarios of tumor evolution and generating intratumoral heterogeneity in MF.

(A) Our previous work showed that lesions of MF are initiated by polyclonal circulating malignant T-cells homing to the skin where they undergo expansion and accumulation of mutations. Various clones (defined as T-cells sharing identical TCR β clonotype) are highlighted by different colours. Expanding clones accumulate mutations and form subclones forming a phylogenetic structure. In lesions funded by a single T-cell clone (left), the entire lesion will comprise the same clonotype and the epidermal subclones will form a branch of the phylogenetic tree. If the lesion is initiated by diverse subclones (middle) that primarily proliferate in the dermis and secondarily infiltrate the epidermis the malignant cells in the epidermis and dermis would be polyclonal but epidermal malignant T-cells will form branches derived from dermal subclones. Finally, in case of independent seeding of the dermal and epidermal niches (right), both compartments will harbour cells showing non-overlapping (or partially overlapping) clonotypes and independent patterns of mutational subclones. (B) Different shapes of phylogenetic tree characteristic for neutral evolution and Darwinian evolution by natural selection. In neutral evolution the shape of the phylogenetic tree is symmetrical (left) in contrast to Darwinian evolution where the extinction of the subclones

will prune some branches (right). To avoid confusion, we use the term “clone” as the group of T-cells of identical clonotype (i.e. sharing a common ancestry) rather than to mutationally identical groups of cells which we refer to as “subclones”. Clades are collections of several subclones.

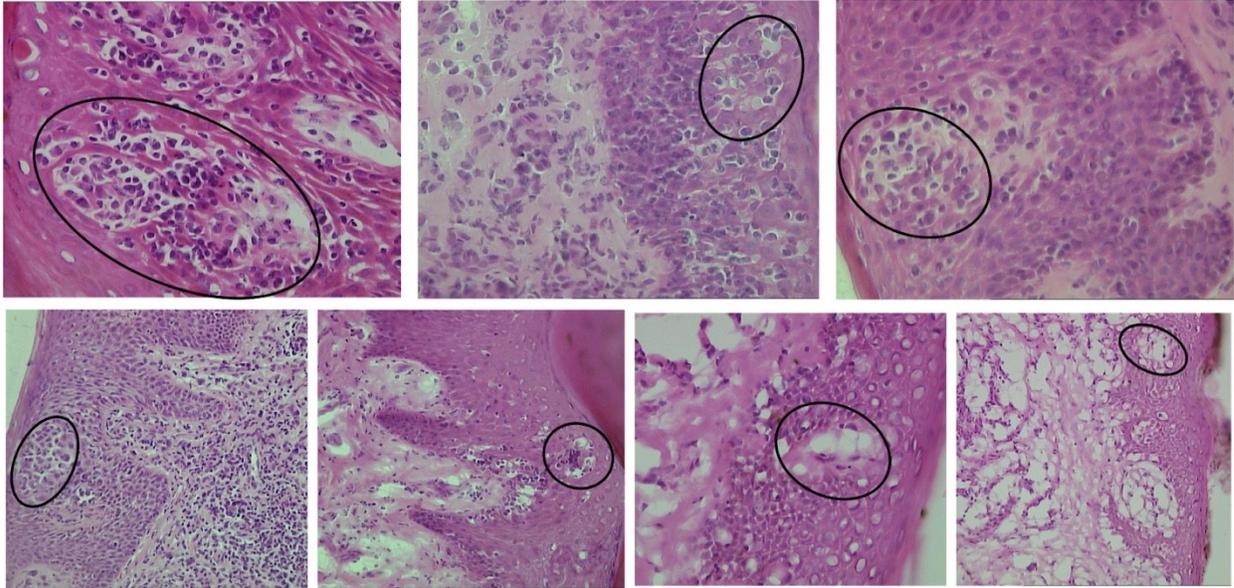


Figure D2: Histological identification of Pautrier microabscess in MF.

Skin biopsies of MF lesions were sectioned at 10 μ thickness and stained with H and E staining protocol. Cluster of atypical lymphoid cells with enlarged hyperchromatic nuclei clustered in epidermis were used as histological markers to identify Pautrier microabscesses.

Representative images (magnification of 20x or 40x) of the H and E stained issues are presented for the 7 patients in the study. Black circles indicate the tissue microdissected for WES.

Appendix Table D1: Patient characteristics and samples included in the study

Patient ID (age [years], sex [M-male, F-female])	Sample ID	Lesion type	Diagnosis and stage
MF17 (70, M)	MF17E	Plaque	Mycosis Fungoides, IB
	MF17D		
MF18 (78, M)	MF18E	Plaque	Mycosis Fungoides, IB
	MF18D		
MF22 (56, F)	MF22E	Plaque	Mycosis Fungoides, IA
	MF22D		
MF23 (69, F)	MF23E	Plaque	Mycosis Fungoides, IA
	MF23D		
MF28 (65, M)	MF28E	Plaque	Mycosis Fungoides, IB
	MF28D		
MF41(77, F)	MF41E	Plaque	Mycosis Fungoides, IIB
	MF41D		
MF42 (82, M)	MF42E	Plaque	Mycosis Fungoides, IIIB
	MF42D		