Published in final edited form as:

Sci Transl Med. 2018 May 09; 10(440): . doi:10.1126/scitranslmed.aar5894.

High-Throughput Sequencing of the T cell Receptor β gene identifies aggressive early-stage Mycosis Fungoides

Adele De Masson¹, John T. O'Malley¹, Christopher P. Elco¹, Sarah S. Garcia¹, Sherrie J. Divito¹, Elizabeth L. Lowry¹, Marianne Tawa¹, David C. Fisher², Phillip M. Devlin³, Jessica E. Teague¹, Nicole R. Leboeuf¹, Ilan R. Kirsch⁴, Harlan Robins⁴, Rachael A. Clark¹, and Thomas S. Kupper^{1,*}

¹Department of Dermatology, Brigham and Women's Hospital, and the Center for Cutaneous Oncology, Dana-Farber / Brigham and Women's Cancer Center, Harvard Medical School, Boston MA 02115, USA

²Department of Hematologic Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston MA 02115, USA

³Department of Radiation Oncology, Dana-Farber / Brigham and Women's Cancer Center, Harvard Medical School, Boston MA 02115, USA

⁴Adaptive Biotechnologies, Seattle WA 98102, USA.

*To whom correspondence should be addressed: Thomas S. Kupper or Rachael A. Clark, tkupper@bwh.harvard.edu, rclark@bwh.harvard.edu.

Author Contributions

ADM, JTO, RAC and TSK designed the study, performed data analysis and interpretation, and wrote the manuscript. ADM, ELL, JET, JTO, CPE did laboratory experimentation and analysis. HR and IRK contributed to data analysis of high-throughput sequencing. SG, JTO, SJD, MT, DCF, PMD, ELL, NRL and TSK acquired clinical samples and patient data and provided clinical interpretation. ADM and JTO prepared the figures.

All authors approved the final manuscript

** share first authorship

† share last authorship

Competing Interests

HR and IRK are employed by and own equity in Adaptive Biotechnologies. I.R.K., H.S.R., T.S.K., and R.A.C. are named as inventors on two pending patent applications regarding the use of HTS in cutaneous lymphoma. I.R.K. and H.S.R. hold stock in Adaptive Biotechnologies, and I.R.K. is employee of Adaptive Biotechnologies. T.S.K. serves on the Scientific Advisory Board (Hematology) of Adaptive Biotechnologies but does not own stock or receive compensation.

Data and materials availability

Supplementary figures

Figure S1. TCR V β high-throughput sequencing allows specific quantification of the frequency of the malignant T cell clone within a V β gene family

Figure S2. Continuous relationship between the tumor clone frequency and the hazard ratio for progression-free and overall survival

Figure S3. Prognostic value of the Cutaneous Lymphoma International Prognostic Index (CLIPi) in early-stage mycosis fungoides

Figure S4. Tumor clone frequency in patches versus plaques

Figure S5. Prognosis in early-stage patients according to the body surface area involved and the presence of plaques

Figure S6. Prognosis in stage IA patients

Figure S7. Reproducibility of the tumor clone frequency as measured by high throughput sequencing of the TCR β gene in different lesions in the same patient

Figure S8. Progression-free and overall survival in pre-treated and treatment-naïve early-stage mycosis fungoides patients with a tumor clone frequency >25%

Supplementary Tables

Table S1. ISCL/EORTC classification and staging of mycosis fungoides and Sézary syndrome

Table S2. Clinical characteristics of 208 patients with cutaneous T cell lymphoma in the discovery set

Table S3. Clinical characteristics of 101 patients with cutaneous T cell lymphoma in the validation set

Table S4. Multivariable analysis on progression-free survival in early-stage patients

Abstract

Mycosis fungoides (MF), the most common cutaneous T cell lymphoma (CTCL) is a malignancy of skin-tropic memory T cells. Most MF cases present as early stage (Stage I A/B, limited to skin), and these patients typically have a chronic, indolent clinical course. A small subset of early-stage cases, however, develop progressive and fatal disease. Because outcomes can be so different, early identification of this high-risk population is an urgent unmet clinical need. We evaluated the use of next-generation high-throughput DNA sequencing of the T cell receptor β gene (*TCRB*) in lesional skin biopsies to predict progression and survival in a discovery cohort of 208 patients with CTCL (177 with MF) from a 15-year longitudinal observational clinical study. We compared these data to the results in an independent validation cohort of 101 CTCL patients (87 with MF). The tumor clone frequency (TCF) in lesional skin, measured by high-throughput sequencing of the *TCRB* gene, was an independent prognostic factor of both progression-free and overall survival in patients with CTCL, and MF in particular. In early-stage patients, a TCF>25% in skin had a higher HR for PFS than any other established prognostic factor (stage IB versus IA, presence of plaques, high blood lactate dehydrogenase concentration, large-cell transformation, or age). The TCF is therefore a biomarker that accurately predicts disease progression in early-stage MF. Early identification of patients at high risk for progression could help identify candidates who may benefit from allogeneic hematopoietic stem cell transplantation before their disease becomes treatment-refractory.

One Sentence Summary:

The malignant T cell clone frequency in cutaneous T cell lymphoma lesions is an independent biomarker for early disease progression and death.

Introduction

Cutaneous T cell Lymphomas (CTCL) are uncommon non-Hodgkin lymphomas of mature skin-tropic memory T cells. Mycosis Fungoides (MF) is the most common and prevalent CTCL, and typically presents as inflammatory patches and plaques on the skin. Diagnosis is often difficult, and has relied on a combination of clinical, histopathological, and molecular findings (1). The average time from appearance of lesions to definitive diagnosis has been estimated to be 3–6 years (2). Recently, the advent of next-generation high-throughput DNA sequencing has revolutionized the diagnosis of MF. MF is nearly always a malignancy of CD4⁺ T cells with an $\alpha\beta$ T cell receptor, encoded by the *TCRA* and *TCRB* genes (3). High-throughput sequencing of the *TCRB* gene can not only identify the unique T cell clone in MF, but can precisely determine the tumor clone frequency (TCF) in the entire T cell infiltrate (3, 4).

A major challenge in the management of MF patients is the identification of early-stage patients at high risk for progression to advanced disease. More than 80% of early-stage patients will have an indolent life-long course free of disease progression, regardless of treatment modality (5). As a result, early-stage patients are treated and maintained with conservative skin-directed therapies unless their disease worsens (6). However, a subset of patients develops highly aggressive, treatment-resistant disease that can be fatal. Although

greater percent skin surface area involvement is associated with a somewhat higher risk of progression, the majority of early-stage MF patients have indolent courses (5). In contrast, advanced-stage patients (stage IIB or higher) have dismal prognoses, with life expectancies ranging from 1.5 to 4 years. Recently, allogeneic hematopoietic stem cell transplantation has emerged as a potentially life-saving intervention in advanced-stage CTCL patients (7). Outcomes from this procedure are somewhat better in patients with Sézary syndrome (SS, a leukemic form of CTCL) than with MF, but regardless, successful outcomes are observed only in patients who are in complete (or near complete) remission at the time of transplantation (8). Unfortunately, such significant remissions are typically impossible to achieve in advanced MF (9). Therefore, prospective identification of MF patients with early-stage disease who are at high risk for disease progression as potential candidates for allogeneic hematopoietic stem cell transplantation is a major unmet clinical need.

Much effort has been devoted to identifying early-stage patients at high risk for disease progression. Previous studies have identified clinical variables associated with decreased progression-free survival (PFS) (5, 10). A Cutaneous Lymphoma International Prognostic Index (CLIPi) has been developed and applied to patients with both early and late-stage disease (11). Although useful in late stage disease, when applied to independent cohorts of early-stage patients, this index has been of limited utility (12). Several studies have identified candidate biomarkers using transcriptional profiling that may improve the prognostic predictions in CTCL (13–15), but these are cumbersome to use in clinical practice and none has been fully validated. Clinically useful and validated risk factors for progression in early-stage disease patients are still based on the physical exam. They include body surface area involvement (with CTCL disease stages T1/IA and T2/IB involving <10% and 10% body surface area, respectively), and the presence of skin plaques (subclass b) vs. patches (subclass a) (Table S1) (10). Although useful, these variables can be subjective, arbitrary, and imprecise; for example, stage T2/IB disease covers from 10% to 79% body surface area, and patients may have a mixture of patches and plaques in different proportions. An objective and quantitative biomarker that addresses likelihood of disease progression does not currently exist.

Recently, we showed that high-throughput sequencing of the TCR genes (*TCRB*, *TCRG*) provides a superior tool for the diagnosis of CTCL by precise identification of the malignant T cell clone (3). Because each T cell clone possesses a unique TCR complementarity-determining region 3 (CDR3) sequence (3), DNA sequencing allows the precise identification and absolute quantification of both malignant and benign T cell clones in CTCL (3, 4). Skin lesions of MF patients are infiltrated by large numbers of non-malignant memory T cells, and it is often impossible to distinguish the malignant T cell clone from activated benign infiltrating T cells in early-stage lesions by histopathology (16). The high-throughput sequencing test greatly facilitates the diagnosis of early-stage disease (3), allows tracking of specific T cell clones over time and in different tissues (4, 17, 18) and detects residual disease after treatment with high sensitivity (19, 20). For the past three decades, the most commonly used diagnostic assay for clonality in CTCL patients employs polymerase chain reaction (PCR) amplification of the rearranged TCR gene, typically *TCRG*, followed by denaturing gradient gel electrophoresis and gel scanning or Biomed GeneScan analysis (21). These non-quantitative tests have false negative rates of at least 25% and a false

positive rate of 15% in the setting of MF (21, 22) and are particularly unreliable in early-stage MF. In the present study, we asked whether high-throughput sequencing of *TCRB* in DNA extracted from lesional skin could predict clinical outcome in large cohorts of CTCL patients.

Results

High-throughput *TCRB* sequencing in lesional skin of 309 patients with cutaneous T cell lymphomas

We performed high throughput sequencing of the *TCRB* gene in lesional skin of 309 patients with cutaneous T cell lymphomas in the DFCI-02016 longitudinal study at Dana Farber between 2002 and 2016. The clinical characteristics of the 309 patients in the discovery cohort (n=208) and validation cohort (n=101) are detailed in Tables S2 and S3, respectively. The distribution of the types of CTCL in both cohorts is shown in Figure 1A, and included primarily patients with MF and SS. The distribution of TCR V β family usage by the tumor clone is depicted in Figure 1B. The most frequently used TCR V β family in CTCL patients was TCRBV20 which represented 13% of all T cell clones. Although our patients had MF, this observation is similar to published data in patients with Sézary syndrome (23). TRBV20 is associated with *Staphylococcus aureus* infection (24), which commonly colonizes the skin of CTCL patients and has been associated with superantigen-driven TCR stimulation in a subset of patients (24). We measured the tumor clone frequency (TCF) in each sample as follows: $TCF = (v_1 / \sum v_n) \times 100$, where v_1 is the number of reads of the most abundant *TCRB* sequence, and v_n is the number of all rearranged *TCRB* sequence reads. This method does not take into account reactive T cells with two rearranged TCRB alleles, but these cells represent a minority of $\alpha\beta$ T cells. As such, the TCF is a conservative estimate of the tumor clone frequency. Examples in two patients with stage 1B MF are depicted in Figure 1C. Histopathological analyses demonstrated that a high TCF was not associated with higher absolute numbers of mononuclear cells in the skin infiltrate (Figure 1D). There was no statistically significant difference in terms of TCF between patients with skin category T1 (<10% body surface area involved with patches and plaques), T2 (>10% body surface area involved with patches and plaques), and T4 (erythroderma) distribution. Only category T3 patients showed a small but statistically significant increase in TCF ($p < 0.05$). Thus, TCF did not increase as a function of skin category T value alone (Figure 1E).

Immunostaining versus high throughput *TCRB* sequencing

Counting T cells by immunostaining with antibodies to V β gene products has been used to identify clonal populations in skin, since all malignant clonal cells express the same V β gene product. Therefore, we asked whether immunostaining could substitute for high-throughput sequencing of the *TCRB*. However, antibodies are available for only about 50% of V β families. Moreover, immunostaining for V β is inherently imprecise in the identification and quantification of a specific T cell clone. In part, this is because a given *TCRVB* exon can rearrange and pair with one of 13 *TCRJB* exons during intrathymic T cell maturation. In one patient (339) analysed with high-throughput sequencing, 28.4% of skin T cells were TCRVB20+, but only 39.8% of these TCRVB20 T cells shared the specific CDR3 sequence of the malignant clone (CSALGLSSYNEQFF) (Figure S1A). Thus, staining with the anti-

V β 20 antibody (Figure S1B) overestimated the true clonal frequency, because it also stained benign infiltrating T cells expressing TCRVB20 (60.2% of T cells expressing TCRVB20, Figure S1C). In another patient (425), 68.4% of T cells were TCRVB05, but this population included 94% of the malignant clone (64.5% of total T cells) (TCRVB05–1/J01–02, CDR3 sequence CASSLGGTGGYTF, FigureS1A,B,C). Here, antibody staining more accurately estimated the malignant clone, but was still variable from histological section to section. These approaches appear to be fundamentally inferior at quantifying the malignant clone when compared to the highly quantitative metric of TCF.

Prognostic impact of clinical, histological and molecular parameters

We first tested the association of the TCF in lesional skin, as measured by high-throughput sequencing of *TCRB*, with prognosis in all CTCL patients in our discovery cohort. A TCF>25% in skin was significantly associated with reduced PFS ($p<0.001$) and overall survival (OS) ($p<0.001$) in 208 patients with CTCL (Figure 2A). This was confirmed in a validation cohort of 101 patients ($p<0.001$) (Figure 2B). The TCF in skin was significantly associated with the PFS ($p<0.001$) and OS ($p<0.001$) in patients with MF, in which the disease primarily affects the skin (Figure 2C–D). By contrast, in patients with SS (in which there is considerable blood involvement) there was no significant association of the TCF in skin with PFS or OS (Figure 2E). This prompted us to restrict our subsequent analyses to patients with MF ($n=177$, discovery cohort). We did not address the predictive value of TCF in the blood of SS patients in this study, as we were focused on the utility of skin biopsy alone.

Gender, age, blood lactate dehydrogenase (LDH) concentration, folliculotropism, large-cell transformation, and the presence of a clone in skin detected by PCR have all been associated with disease progression in MF (5). We therefore compared these variables to the TCF in our discovery cohort for their association with PFS and OS. Age>60 years ($p<0.01$), elevated LDH ($p<0.01$), the existence of large-cell transformation ($p<0.001$) and the TCF>25% ($p<0.001$) were significantly associated with PFS and OS ($p<0.001$, $p=0.001$, $p=0.001$, $p<0.001$ respectively) in univariable analysis (Table 1). In a multivariable analysis that included age, advanced tumor stage, LDH concentration, large-cell transformation, and the type of treatment received as confounding factors, the TCF was still significantly associated with PFS ($p<0.001$) and OS ($p<0.001$). A TCF of 25% was found to be the best cutoff as determined by the concordance index (25) in univariable analysis on PFS and OS. There was a continuous relationship between the TCF threshold and the hazard ratios for OS and PFS, until a TCF of 25% where a plateau was reached (Figure S2).

Prognosis in early-stage MF

We hypothesized that the predictive value of the TCF in skin might be more useful in patients with early-stage MF (skin-limited disease), where outcome is uncertain, than in patients with advanced-stage disease, where life expectancy is invariably reduced. This prompted us to study the prognostic value of the TCF in skin in patients with early-stage MF and compare it to existing prognostic factors in these patients. There was a significant interaction between the TCF and the disease stage (early versus advanced-stage, $p<0.01$). Most patients with MF present with early-stage disease, and up to 20% will experience

disease progression and/or death within 10 years (5). In early-stage patients with skin-limited disease, the body surface area involved in the disease is considered the primary prognostic factor, with T1/IA involving <10% and T2/IB >10% of the body surface area (5, 26) (Table S1). Variables significantly and independently associated with prognosis in the entire cohort (Table 1) were studied in a sub-cohort of early-stage patients. Univariable analysis on 141 early-stage MF patients in the discovery set revealed that TCF>25% ($p<0.001$), disease stage (T2/IB versus T1/IA) ($p<0.01$), age (>60 years) ($p<0.05$), and the presence of plaques ($p<0.05$) were each significantly associated with PFS, but the hazard ratio for TCF was the highest (Table 2). To compare the prognostic value of the TCF to the reference prognostic index in patients with early-stage CTCL, the Cutaneous Lymphoma International Prognostic Index (CLIPi) was calculated in these patients as described in (11). An intermediate or high CLIPi score was associated with a lower PFS in the discovery set, but again the hazard ratios (HR) were lower than the HR of the TCF (Table 2 and Figure S3).

A TCF>25% showed the highest HR for PFS both in the discovery set (HR, 4.9, 95% CI, 2.5–9.7, $p<0.001$) and the validation set (HR, 10, 95% CI, 3.4–31, $p<0.001$) (Figure 3A–B). In contrast, the HR of stage (T2/IB vs. T1/IA) was lower in both the discovery set (HR, 2.5, 95% CI, 1.3–4.9, $p<0.01$) and the validation set (HR, 6.4, 95% CI, 1.5–28, $p=0.01$). In the discovery set, 89% (95% CI, 76–95%) of stage T2/IB patients with a malignant clone<25% of the skin T cells were alive without disease progression 4 years later, vs. 30% (95% CI, 7–58) of stage T2/IB patients with a malignant clone>25% (Figure 3C). In the validation set, 85% (95% CI, 50–96%) of stage T2/IB patients with a malignant clone <25% were alive and progression-free 4 years later, vs. 19% (95% CI, 5–40%) of patients with a malignant clone>25% (Figure 3D). The TCF in skin was also significantly associated with OS ($p<0.01$) (Figure 3C).

There was a significantly higher TCF in patients with plaques compared to patches ($p<0.001$). However, a number of patients with patches had a high TCF in skin, and conversely, many patients with plaques had a low TCF (Figure S4). To directly compare the prognostic value of TCF to both skin stage (T1/IA vs. T2/IB) and the presence of plaques (b) vs. patches (a), we assessed these variables on the discovery and validation sets. Figure S5 compares PFS and OS of T1a, T1b, T2a, and T2b patients, and confirms that both skin stage (T2/IB versus T1/IA), and the presence of plaques are associated with decreased PFS ($p<0.01$ and $p<0.05$ for skin stage and plaques, respectively) and that the skin stage is associated with decreased OS ($p<0.01$) in Cox univariable analysis. However, when PFS and OS in Stage IB/T2 patients were assessed according to the presence or absence of plaques (IB/T2a vs. IB/T2b) or the TCF>25% (Figure 3C–D), the latter was far more predictive. For TCF>25%, stage IB/T2 patients had decreased PFS (HR=13, 95%CI, 5–36, $p<0.001$ in the discovery set; HR=11, 95% CI, 2.5–48, $p=0.001$ in the validation set) and OS (HR=9.0, 95% CI, 3.0–27, $p<0.001$) (Figure 3C–D). For T2b vs. T2a, the HR for PFS was 2.2 (95% CI, 0.9–5.3, $p=0.08$) in the discovery set and 1.7 (95% CI, 0.6–4.8, $p=0.3$) in the validation set, and the HR for OS was 2.0 (95%CI, 0.7–6.5, $p=0.2$) (Figure 3C–D). The TCF was significantly associated with PFS (HR=5.8, 95% CI 1.8–19, $p=0.004$) in a multivariable model that included the age, stage T2 versus T1, and the presence of plaques as covariates (Table S4). This was confirmed on the independent validation cohort (HR=13.6, 95% CI

1.2–154, $p=0.03$). Therefore, in stage IB/T2 patients, $\text{TCF}>25\%$ was highly predictive of PFS and OS, and was far more predictive than the presence of plaques vs patches. In stage IB patients who experienced progression or death during the follow-up, there was an inverse correlation between the TCF and the time to progression or death ($\rho -0.6$, $p<0.001$) (Figure 3E).

A TCF in skin $>25\%$ was associated with a positive predictive value of 92% for 5-year disease progression or death, and a negative predictive value of 83% (Figure 3F). As previously shown (5), stage IA/T1 patients (who have limited skin involvement with $<10\%$ of the body surface area involved) had an excellent prognosis regardless of the TCF (Figure S6). These data indicate that the frequency of the malignant T cell clone in skin (TCF) is the single best predictive test for identifying patients at risk for disease progression, particularly in Stage IB patients, who appear to be the only early stage patients who progress. The variability of the TCF between different lesions of the same type (e.g., patches or plaques) in the same patient was low, as depicted in Figure S7. There was no significant difference between early-stage patients with a $\text{TCF}>25\%$ and $\leq 25\%$ in terms of treatments received prior to sequencing. Patients with a $\text{TCF}>25\%$ had a poor progression-free and overall survival with no significant difference between treatment-naïve and pre-treated patients. The TCF was significantly associated with PFS (HR=4, 95% CI 1.3–12, $p=0.01$) and OS (HR 8.9, 95% CI 2–39, $p=0.004$) in treatment-naïve early-stage patients (Figure S8).

T cell microenvironment

We speculated that the poor prognosis associated with a higher TCF in skin might be linked to a defective antitumor immune response, an intrinsic aggressiveness of the malignant cells themselves, or a combination of these variables. To investigate the causal mechanism, we assessed the immune microenvironment in patients with a high TCF in skin versus patients with a low TCF in skin. As the TCF accounts for the number and diversity of non-malignant T cells present in the lesional skin biopsy, differences in the TCF might simply represent different numbers of reactive, non-malignant T cells in the setting of similar numbers of clonal T cells.

An increased number of reactive CD8⁺ T cells in the skin of patients with CD4⁺ MF has been previously associated with improved prognosis (29, 30). We performed CD8 and granzyme B staining in the skin of early-stage MF patients with a high TCF ($>30\%$ T cells) and a low TCF in skin ($<10\%$) (Figure 4A). TCFs of 10% and 30% were chosen as cutoffs because they were close to the 25th and 75th percentiles in this population. There was no significant difference in the percentage of CD8⁺ T cells in skin between these two groups. The percentage of Granzyme B-positive cells was not significantly lower in patients with a high TCF (Figure 4B). In fact, patients with a high TCF in skin had a more clonal reactive T cell infiltrate (Figure 4C), a feature which has been suggested to represent the capacity to mount an antitumor immune response in skin (27). Thus, a defective antitumor immune response alone, by these criteria, does not seem to be the primary mechanism of the progression in patients with a high TCF.

Gene expression profiling and exome sequencing

We obtained gene expression data on 78 biomarkers in CTCL that were obtained in lesional skin biopsies from 157 patients with early- and advanced-stage MF and SS. These biomarkers were selected based on previous studies showing their overexpression in CTCL compared to normal skin (13–15), or their copy number variations in CTCL (23, 28–30). The unsupervised analysis of the dataset revealed that patient transcriptomes clustered in 3 groups according to gene expression (Figure 5A), in accordance with previously published work (13–15). Patients in cluster 1 had overexpression of numerous T cell specific genes, such as cell surface markers (CD4, CCR4, CCR7, CD28, CD52, PDCD1), genes in the IL21/JAK/STAT pathway (IL21, IL2RG, JAK3), and genes in the TCR signaling pathway (ITK, LCK, PRKCQ, SH2D1A, FYB, LAT, PTPRCAP, RAC2, GIMAP4, T3JAM, CARD11, SIT1, PIK3CD, VAV1, LEF1). We next asked if this gene expression pattern simply reflected differences in T cell abundance between cluster 1 versus the two other clusters. Although there were no statistically significant differences in the absolute abundance of total T cells in any of the clusters studied (Figure 5B), the abundance of the malignant clone (relative to total lesional T cells) was significantly higher in patients in cluster 1 (poor prognosis) compared to the 2 other clusters ($p < 0.05$ compared to cluster 3 and $p < 0.01$ compared to cluster 2, Mann-Whitney U-test with Bonferroni correction) (Figure 5C). There was a significantly lower PFS in patients in cluster 1 versus the 2 other clusters ($p < 0.05$ compared to cluster 3 and $p < 0.001$ compared to cluster 2, log-rank test with Bonferroni correction) (Figure 5D). Thus, we identified a cluster of patients (cluster 1) with a distinct gene expression profile, high tumor clone frequency in skin and poor prognosis. The overexpression of genes in the JAK-STAT and TCR signaling pathways in patients with a high TCF are consistent with the role of these pathways in T cell proliferation and survival.

Whole exome sequencing of tumors has yielded valuable data in a variety of cancers, but has been difficult to perform in patches or plaques of MF because of the paucity of tumor cells relative to total nucleated cells. We thus conducted whole exome sequencing on microdissected skin T cells in 19 patients with skin-limited MF, using peripheral blood mononuclear cells as a comparator. The mean target coverage was 70× in tumor samples and 103× in peripheral blood mononuclear cells. The number of somatic mutations was significantly correlated to the TCF in skin ($r = 0.5$, $p = 0.04$) (Figure 5E). There was a higher number of somatic mutations in patients in stage IIB compared to patients in stage IB and IA, but all demonstrated abundant somatic mutations. This indicates that the poor prognosis associated with a high TCF in skin is accompanied by a high frequency of genetic abnormalities of the malignant cells.

Discussion

In this report, we showed that an increased frequency of the malignant T cell clone in skin was strongly correlated with reduced PFS and OS in patients with CTCL, particularly in patients with early-stage MF with a T2 distribution. Histological examination of patches and plaques for the density of the T cell infiltrate could not distinguish lesions with high TCF vs low TCF. TCF was the single best independent predictor of PFS and OS in MF, and particularly in early stage MF. Moreover, it is readily obtained from sequencing a small skin

biopsy. Prior to this study, assessment of body surface involvement (T1 vs. T2) has been the best means of predicting which patients might progress. Although our data are in agreement with the well-accepted observation that Stage IB patients (T2) are more likely to progress than Stage IA (T1) patients, the TCF outperformed the predictive value of T category. Similarly, the presence of skin plaques (versus patches) and the currently used CLUPI criteria were less discriminative and predictive than the TCF in our early-stage patients. Taken together, our findings suggest that a TCF>25% in MF skin lesions is the most sensitive and specific method available to identify early-stage patients at the highest risk for disease progression. For patients with T2/Stage IB disease, the positive predictive value is 92% and the negative predictive value is 83% for 5-year PFS.

Risk stratification is one of the goals of precision oncology, and there is great interest in biomarkers that predict aggressive disease in malignancies in which a majority of patients have indolent disease, while a smaller subset have aggressive disease. Identifying patients at risk for disease progression is particularly important in CTCL, a disease in which two patients with similar physical exams and histopathological morphology can have markedly different outcomes. High-throughput sequencing of the *TCRB* gene provides a quantitative measurement of the malignant T cell clonal burden in CTCL lesions. Moreover, it is straightforward and readily accomplished using available platforms.

It is possible that differences in the TCF might simply represent different numbers of reactive, non-malignant T cells in the setting of similar numbers of clonal T cells. However, the total number of T cells alone was not able to accurately discriminate patients at high risk of progression. An increased number of reactive CD8+ T cells in the skin of patients with CD4+ MF has been previously associated with improved prognosis (31, 32). Our data however indicate that a defective antitumor immune response does not appear to be the primary mechanism associated with progression in patients with a high TCF in skin. The subset of patients with a high TCF displayed a specific gene expression profile in lesional biopsies, and their malignant T cells harboured a higher number of somatic mutations. These alterations were already detectable in stage IB patients with a TCF>25% in skin, which is consistent with the poor prognosis associated with these features. The limitations of our study include the fact that this is a retrospective analysis of patients from a single Cancer Center. Patients' samples were collected at the time of presentation to the DFCI Cutaneous Lymphoma Clinic. Many patients were treatment-naïve, but others were referred from outside facilities for disease management, often after relapse or progression during treatment. This precluded us from taking samples exclusively upon diagnosis and before treatment. However, treatment appeared to have no effect on the predictive nature of TCF. In patients with a TCF>25%, there was no significant difference in terms of progression-free or overall survival between patients who had received treatments prior to inclusion and treatment-naïve patients (Figure S8). The usually indolent nature of MF and the long follow-up period required mean that prospective studies will take many years; however, such studies are essential to validate our findings.

The prognostic value of a TCF>25% in skin for progression-free and overall survival may help identify the subset of patients who may ultimately benefit from allogeneic stem cell transplant earlier in their disease course. Because the tumour cells are restricted to the skin

in MF, their overall tumour burden is a product of the TCF and the body surface area involved. It is therefore not surprising that a high TCF is associated with a poor prognosis in these patients. Overexpression of genes involved in TCR signalling was found in MF patients with a high TCF, and in peripheral T cell lymphomas this has led to therapeutic strategies that target TCR signalling (33). Our gene expression data suggest that therapeutic inhibition of TCR signalling in CTCL may be a useful strategy, although other pathways might be involved in CTCL cell survival and proliferation. Finally, the role of immunomodulatory therapies in this subset of high-risk patients should be assessed. A previously published study showed efficacy of topical resiquimod, a Toll-like receptor agonist, including in patients with a high pre-treatment TCF in skin (19). In summary, our data suggest that a TCF of >25%, as determined by high throughput DNA sequencing of the *TCRB* gene, is a strong predictor of disease progression and survival in patients with MF limited to the skin.

Materials and Methods

Study Design

This is an experimental laboratory study performed on human tissue samples. All studies were performed in accordance with the Declaration of Helsinki. Lesional skin from patients with CTCL was obtained from patients seen at the Dana-Farber/Brigham and Women's Cancer Center Cutaneous Lymphoma Program and included in the DFCI-02016 observational cohort, after informed consent. Eligibility criteria included a confirmed diagnosis of CTCL after review of the clinical, molecular and histological data, as well as adequate remaining research or clinical specimens for high-throughput sequencing. CTCL patients described in this article met the WHO-EORTC (World Health Organization–European Organization for Research and Treatment of Cancer) criteria for Sézary syndrome or MF (22). All tissues were collected with previous approval from the Partners, Dana-Farber Institutional Review Boards. All samples with enough available material were analysed by high-throughput sequencing of the *TCRB* gene, which were masked to clinical outcomes. Staging and disease progression were evaluated according to the ISCL/EORTC criteria (10, 26). Analyses of HTS studies were done in an investigator-blinded fashion. Immunostaining studies were performed using in vitro assays without blinding or randomization. Study components were not predefined.

Patients

The primary discovery cohort comprised 208 patients with CTCL seen at the Dana-Farber Cancer Institute's Cutaneous Lymphoma Clinic from 2002 to 2016 (Table S2). This discovery set included 177 patients with MF with a median follow-up of 8 years. Samples were typically collected at the time of diagnosis or at the time of referral to the DFCI Cutaneous Lymphoma Clinic for management of established disease.

The independent validation set included 101 distinct CTCL patients recently included in the same study, including 87 patients with MF (Table S3). The data were collected on December 23, 2016.

Nucleic acid extraction

DNA and RNA were extracted from four 20 µm-thick formalin-fixed, paraffin-embedded tissue scrolls from a lesional skin biopsy using the Allprep DNA/RNA FFPE isolation kit (Qiagen) as per the manufacturer's instructions. DNA and RNA amounts were measured using a BioDrop spectrophotometer (Denville Scientific Inc.). For fresh frozen samples from the validation set, DNA was isolated from thirty cryosections of 10 µm thickness. DNA extraction was carried out using the QIAamp DNA Mini Kit (Qiagen) kit as per manufacturer's instructions with overnight tissue digestion.

High throughput sequencing of the *TCRB* gene

Immunosequencing.—For each sample, DNA was extracted from skin biopsies. We then shipped it on dry ice to Adaptive Biotechnologies. TCRβ CDR3 regions were amplified and sequenced using ImmunoSEQ (Adaptive Biotechnologies). The ImmunoSeq platform is available as a kit or service (<https://www.adaptivebiotech.com/immunoseq>). All TCR-β characterization was performed by Adaptive Biotechnologies using the ImmunoSeq TCR-β 'survey level' human assay (4, 34) which has previously described in detail (3).

Clonal detection.—The putative malignant clone was defined by sequence abundance. A clone can have either one or two rearranged TCR alleles. For most of the clones, both TCRG alleles are rearranged, and for TCRB, a minority have both alleles rearranged. For consistency, a clone's abundance was defined by summing the abundance of the most frequent single allele for TCRB. The putative malignant clone was defined by relative abundance of its unique CDR3 sequence (3). The percent of T cells consisting of the malignant clone was determined by dividing the abundance of the malignant clone (number of reads) by the total number of T cell reads. This method does not take into account reactive T cells with two rearranged TCRB alleles, but these cells represent a minority of αβ Tcells.

Reactive T cell diversity and clonality measurements.—The diversity of the reactive T cell clones was studied using the Shannon's index. Shannon's entropy quantifies the uncertainty in predicting the sequence identity of a random sequence from a dataset. The Shannon's index of the reactive clones (H) was calculated according to the following formula: $H = -\sum_{i=1}^S f_i \log(f_i)$ where i represents each individual reactive clone and f the frequency of this rearrangement among all productive rearrangements in the sample, excluding the malignant clone. To allow for comparisons between samples differing in the total number of reads, entropy was normalized by division of log 2 of the number of unique productive sequences. Clonality is the reciprocal of normalized Shannon's entropy (Clonality=1-normalized entropy) with values ranging from 0 (most diverse) to 1 (least diverse).

Cryosection immunostaining and cell counting

CTCL skin samples were co-immunostained for anti-Vβ2 (Beckman-Coulter, clone:MPB2D5) or anti-Vβ5.1 (Beckman-Coulter, clone:IMMU 157) conjugated to R-phycoerythrin with anti-CD3 conjugated to Alexa Fluor 647 (Biolegend, clone:UCHT1) with 3 five-minute wash steps in TBS-saponin before mounting. Single color controls

confirmed specificity of staining and no bleed through into the other channel. The samples were analyzed using an Olympus BX43 microscope with the objective lens of 10×/0.40, 20×/0.75 and 40×/0.95 Olympus UPlanFL (Olympus). Images were acquired with the Mantra Quantitative Pathology Imaging System, and analyzed using inForm software (Perkin-Elmer) and the manual counting feature from Adobe Photoshop CS5 (Adobe). We analyzed 10× images of nonoverlapping fields.

Transcriptional analyses

At least 140 ng of mRNA per sample were analysed by NanoString gene expression profiling using a custom codeset including 78 probes directed against potential biomarkers identified in previous gene expression studies by our group (13–15) or in exome sequencing studies by others (23, 28–30), and 3 housekeeping genes. The Nanostring technology uses molecular barcode and single molecule imaging for the direct hybridization and detection of hundreds of unique transcripts in a single reaction. Each color-coded barcode is attached to a single target-specific probe corresponding to an analyte of interest. Combined together with invariant controls, the probes form a multiplexed CodeSet. The samples are run on the nCounter platform. Gene expression data were background subtracted and normalized to positive controls and housekeeping genes using the Nanostring nSolver software (<https://www.nanostring.com/products/analysis-software/nsolver>). Gene expression values were expressed as log2 fold changes (FC) of the average gene expression of the considered gene in the whole study group. Gene expression assays were performed blinded to the patient's outcome.

Exome sequencing of microdissected skin T cells Sample preparation and expression microdissection

The total lesional skin biopsies of 19 patients with skin-limited mycosis fungoides were embedded in Optimal Cutting Temperature compound (OCT) and stored frozen at –80°C. Six-micrometers slides were then sectioned using a microtome-cryostat and stained for CD3 by immunohistochemistry. Briefly, slides were blocked with bovine serum albumin, incubated with rabbit recombinant anti-human CD3 antibody (Sp7, Abcam) followed by secondary antibody coupled to horseradish peroxidase (HRP, Envision+ Dual Link, DAKO) and revelation with diaminobenzidine (DAB, Vector Laboratories).. Slides were dehydrated in alcohol and xylene. A membrane was placed on the tissue and a flashlamp was applied. The flashlamp is an intense pulsed light (IPL) that emits a bright range of wavelengths from ultraviolet to visible light and infrared, but ultraviolet light is filtered out and does not reach the tissue. The light excites and heats the stained cells that transfer to the membrane. The membrane was then placed in lysis buffer and DNA extracted using a QIAmp DNA microkit (Qiagen). The DNA quantity and integrity were measured by using a Bioanalyzer. A matched blood sample from the same patient, without blood involvement as confirmed by high-throughput sequencing of the TCRβ gene, was used as a germline control.

Library preparation and sequencing

This study was done in collaboration with the Center for Cancer Genome Discovery at Dana Farber. Prior to library construction, DNA was fragmented (Covaris sonication) to 250 bp and further purified using Agentcourt AMPure XP beads. Size-selected DNA was then

ligated to specific adaptors during manual library construction (Modified (low input) KAPA Library Prep). Each library was made with sample-specific barcodes, quantified using the MiSeq, and libraries were pooled at equal mass (1×2 -plex) to a total of 750 ng for Exome v5 enrichment using the Agilent SureSelect hybrid capture kit. The capture was then sequenced on HiSeq 2500 and 3000.

Variant analysis

Mutation analysis for single nucleotide variants (SNV) was performed using MuTect v1.1.4 and annotated by Variant Effect Predictor (VEP). We used the SomaticIndelDetector tool that is part of the GATK for indel calling. MuTect was run in paired mode pairing the tumor sample to the matched normal.

Quality control

80% of the targets were covered at least $20\times$. Fingerprinting analysis was performed using 44 polymorphic loci to identify if the aggregation pairing strategy was performed appropriately. Picard Tools GenotypeConcordance was used to calculate the concordance that a given test sample matches the sample being considered. This was performed on all pairwise combinations of samples in the cohort. The output of the pair-wise comparisons was then mapped to a concordance matrix, where concordance values above 4 standard deviations of the median concordance value for the cohort indicated a high likelihood that the samples match.

Statistical analyses

Patient clinical information was collected at the reference date of December 23, 2016. Progression-free survival (PFS) and overall survival (OS) were estimated by the Kaplan-Meier method. PFS was defined as the time between sampling and death from any cause or progression of the lymphoma disease, or the last disease evaluation time where no disease progression was observed. OS was defined as the time between sampling and death from any cause (OS), or censoring at last follow-up. Age, gender, disease stage, serum lactate dehydrogenases (LDH) concentrations, the existence of folliculotropism or large-cell transformation, the presence of a clonal pattern in skin as assessed by polymerase chain reaction (PCR) of the T cell receptor (TCR) γ gene *TCRG*, the malignant clone frequency as assessed by high throughput sequencing of the *TCRB* gene, and the Cutaneous Lymphoma International Prognostic Index (CLIPi) were assessed for their association with PFS and OS in univariable Cox regression analysis. The Cutaneous Lymphoma International Prognostic Index (CLIPi) was calculated based on the presence of the following factors: age>60 years, male sex, plaques, folliculotropism and clinical adenopathy N1/Nx; low risk=0–1, intermediate risk=2, high risk=3–5 prognostic factors (11). Missing values for LDH were imputed in 5 patients (mean imputation based on disease stage) and all cases were used in the final analysis. For PCR, the analysis was carried out on complete cases ($n=114$). A stepwise selection process was applied, with all variables significant ($p<0.05$) in univariable analysis retained in the initial multivariable model, followed by backward elimination. Interactions between the clone frequency and other variables were tested. The proportionality assumption was tested for each variable in the final model and the model was stratified on variables that violated the proportionality assumption. The model was then

tested on an independent cohort. The cutoff value for the malignant clone frequency was selected in order to maintain the concordance index and AIC of the univariable model using the corresponding continuous variable. Sixty years was chosen as the cutoff for age because it was the most frequently selected cutoff in previously published studies in the field. OS was not tested in the validation set due to recent sampling of the patients leading to an insufficient number of events. Comparisons of the T cell percentages and TCF between the 3 clusters were performed using the Mann-Whitney U test followed by Bonferroni correction for multiple testing with $p < 0.05$ considered significant. Heatmaps and hierarchical clustering (complete linkage) were performed with Genesis software (Institute for Genomics and Bioinformatics, Graz Institute of Technology, Graz, Austria, freely available at http://genome.tugraz.at/genesiscient/genesiscient_description.shtml). Medians and interquartile ranges are indicated on dot plots. Statistical analyses were performed with R 3.1.1 and Graphpad Prism.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr Ivan Litvinov for his help with data collection and the statistical reviewers from Harvard Catalyst for helpful statistical advice and reviewing of our paper, as well as Aaron Thorne, Anwesha Nag and Bruce Wollison from the Dana Farber Center for Cancer Genome Discovery for contributing the exome sequencing data.

Funding

This study was supported by generous charitable contributions from Edward P. Lawrence, Esq. and from the Lubin Family Foundation. Support was also obtained from National Institute of Health (NIH) Specialized Program of Research Excellence (SPORE P50 CA9368305) grant (to T.S.K.), NIH R01CA203721 (to R.A.C. and T.S.K.), T32 AR-07098 to T.S.K. (J.T.O.) and NIH R01 AR063962 (to R.A.C.) and NIH P30 AR069625 (to R.A.C.). Support also came from grants from the Societe Francaise de Dermatologie, CEDEF, Association pour la Recherche contre le Cancer, Fondation Rene Touraine and the Philippe Foundation (to A.d.M.) and a CLARIONS grant from the Cutaneous Lymphoma Foundation (to J.T.O.). This work was conducted with support from Harvard Catalyst | The Harvard Clinical and Translational Science Center (National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health Award UL1 TR001102 to J.T.O.) and financial contributions from Harvard University and its affiliated academic healthcare centers.

References and notes

1. Girardi M, Heald PW, Wilson LD, The pathogenesis of mycosis fungoides, *N. Engl. J. Med* 350, 1978–1988 (2004). [PubMed: 15128898]
2. van Doorn R, Van Haselen CW, van Voorst Vader PC, Geerts ML, Heule F, de Rie M, Steijlen PM, Dekker SK, van Vloten WA, Willemze R, Mycosis fungoides: disease evolution and prognosis of 309 Dutch patients, *Arch. Dermatol* 136, 504–510 (2000). [PubMed: 10768649]
3. Kirsch IR, Watanabe R, O'Malley JT, Williamson DW, Scott L-L, Elco CP, Teague JE, Gehad A, Lowry EL, LeBoeuf NR, Krueger JG, Robins HS, Kupper TS, Clark RA, TCR sequencing facilitates diagnosis and identifies mature T cells as the cell of origin in CTCL, *Sci. Transl. Med* 7, 308ra158 (2015).
4. Robins HS, Campregher PV, Srivastava SK, Wachter A, Turtle CJ, Kahsai O, Riddell SR, Warren EH, Carlson CS, Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells, *Blood* 114, 4099–4107 (2009). [PubMed: 19706884]
5. Agar NS, Wedgeworth E, Crichton S, Mitchell TJ, Cox M, Ferreira S, Robson A, Calonje E, Stefanato CM, Wain EM, Wilkins B, Fields PA, Dean A, Webb K, Scarisbrick J, Morris S, Whittaker SJ, 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* 28, 4730–4739 (2010). [PubMed: 20855822]

6. Edelson RL, Outsmarting cutaneous T-cell lymphoma cells by decoding the language they speak: focusing past and present insights on future prospects, *Clin. Lymphoma Myeloma Leuk* 10 Suppl 2, S59–62 (2010). [PubMed: 20826398]
7. Virmani P, Zain J, Rosen ST, Myskowski PL, Querfeld C, Hematopoietic Stem Cell Transplant for Mycosis Fungoides and Sézary Syndrome, *Dermatol. Clin* 33, 807–818 (2015). [PubMed: 26433851]
8. de Masson A, Beylot-Barry M, Bouaziz J-D, Peffault de Latour R, Aubin F, Garciaz S, d'Incan M, Dereure O, Dalle S, Dompnmartin A, Suarez F, Battistella M, Vignon-Pennamen M-D, Rivet J, Adamski H, Brice P, Francois S, Lissandre S, Turlure P, Wierzbicka-Hainaut E, Brissot E, Dulery R, Servais S, Ravinet A, Tabrizi R, Ingen-Housz-Oro S, Joly P, Socie G, Bagot M, French Study Group on Cutaneous Lymphomas and Societe Frangaise de Greffe de Moelle et Therapie Cellulaire, Allogeneic stem cell transplantation for advanced cutaneous T-cell lymphomas: a study from the French Society of Bone Marrow Transplantation and French Study Group on Cutaneous Lymphomas, *Haematologica* 99, 527–534 (2014). [PubMed: 24213148]
9. Hughes CFM, Khot A, McCormack C, Lade S, Westerman DA, Twigger R, Buelens O, Newland K, Tam C, Dickinson M, Ryan G, Ritchie D, Wood C, Prince HM, Lack of durable disease control with chemotherapy for mycosis fungoides and Sézary syndrome: a comparative study of systemic therapy, *Blood* 125, 71–81 (2015). [PubMed: 25336628]
10. Olsen E, Vonderheid E, Pimpinelli N, Willemze R, Kim Y, Knobler R, Zackheim H, Duvic M, Estrach T, Lamberg S, Wood G, Dummer R, Ranki A, Burg G, Heald P, Pittelkow M, Bernengo M-G, Sterry W, Laroche L, Trautinger F, Whittaker S, ISCL/EORTC, 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 Cancer (EORTC), *Blood* 110, 1713–1722 (2007). [PubMed: 17540844]
11. Benton EC, Crichton S, Talpur R, Agar NS, Fields PA, Wedgeworth E, Mitchell TJ, Cox M, Ferreira S, Liu P, Robson A, Calonje E, Stefanato CM, Wilkins B, Scarisbrick J, Wain EM, Child F, Morris S, Duvic M, Whittaker SJ, A cutaneous lymphoma international prognostic index (CLIPi) for mycosis fungoides and Sézary syndrome, *Eur. J. Cancer* 1990 49, 2859–2868 (2013).
12. Sanz-Bueno J, Lora D, Monsalvez V, Maronas-Jimenez L, Postigo C, Rodriguez-Peralto JL, Ortiz-Romero PL, The new Cutaneous Lymphoma International Prognostic index (CLIPi) for early mycosis fungoides failed to identify prognostic groups in a cohort of Spanish patients, *Br. J. Dermatol* 175, 794–796 (2016). [PubMed: 26990536]
13. Litvinov IV, Jones DA, Sasseville D, Kupper TS, Transcriptional profiles predict disease outcome in patients with cutaneous T-cell lymphoma, *Clin. Cancer Res* 16, 2106–2114 (2010). [PubMed: 20233883]
14. Litvinov IV, Netchiporouk E, Cordeiro B, Doré M-A, Moreau L, Pehr K, Gilbert M, Zhou Y, Sasseville D, Kupper TS, The Use of Transcriptional Profiling to Improve Personalized Diagnosis and Management of Cutaneous T-cell Lymphoma (CTCL), *Clin. Cancer Res* 21, 2820–2829 (2015). [PubMed: 25779945]
15. Shin J, Monti S, Aires DJ, Duvic M, Golub T, Jones DA, Kupper TS, Lesional gene expression profiling in cutaneous T-cell lymphoma reveals natural clusters associated with disease outcome, *Blood* 110, 3015–3027 (2007). [PubMed: 17638852]
16. Pimpinelli N, Olsen EA, Santucci M, Vonderheid E, Haeffner AC, Stevens S, Burg G, Cerroni L, Dreno B, Glusac E, Guitart J, Heald PW, Kempf W, Knobler R, Lessin S, Sander C, Smoller BS, Telang G, Whittaker S, Iwatsuki K, Obitz E, Takigawa M, Turner ML, Wood GS, International Society for Cutaneous Lymphoma, Defining early mycosis fungoides, *J. Am. Acad. Dermatol* 53, 1053–1063 (2005). [PubMed: 16310068]
17. Gaide O, Emerson RO, Jiang X, Gulati N, Nizza S, Desmarais C, Robins H, Krueger JG, Clark RA, Kupper TS, Common clonal origin of central and resident memory T cells following skin immunization, *Nat. Med* 21, 647–653 (2015). [PubMed: 25962122]

18. Wu D, Sherwood A, Fromm JR, Winter SS, Dunsmore KP, Loh ML, Greisman HA, Sabath DE, Wood BL, Robins H, High-throughput sequencing detects minimal residual disease in acute T lymphoblastic leukemia, *Sci. Transl. Med* 4, 134ra63 (2012).
19. Rook AH, Gelfand JM, Gelfand JC, Wysocka M, Troxel AB, Benoit B, Surber C, Elenitsas R, Buchanan MA, Leahy DS, Watanabe R, Kirsch IR, Kim EJ, Clark RA, Topical resiquimod can induce disease regression and enhance T-cell effector functions in cutaneous T-cell lymphoma, *Blood* 126, 1452–1461 (2015). [PubMed: 26228486]
20. Weng W-K, Armstrong R, Arai S, Desmarais C, Hoppe R, Kim YH, Minimal residual disease monitoring with high-throughput sequencing of T cell receptors in cutaneous T cell lymphoma, *Sci. Transl. Med* 5, 214ra171 (2013).
21. Ponti R, Fierro MT, Quaglino P, Lisa B, di CF, Paola O, Michela F, Paolo A, Comessatti M, Novelli M, Bernengo G, TCRgamma-chain gene rearrangement by PCR-based GeneScan: diagnostic accuracy improvement and clonal heterogeneity analysis in multiple cutaneous T-cell lymphoma samples, *J. Invest. Dermatol* 128, 1030–1038 (2008). [PubMed: 17989737]
22. Willemze R, Jaffe ES, Burg G, Cerroni L, Berti E, Swerdlow SH, Ralfkiaer E, Chimenti S, Diaz-Perez JL, Duncan LM, Grange F, Harris NL, Kempf W, Kerl H, Kurrer M, Knobler R, Pimpinelli N, Sander C, Santucci M, Sterry W, Vermeer MH, Wechsler J, Whittaker S, Meijer CJLM, WHO-EORTC classification for cutaneous lymphomas, *Blood* 105, 3768–3785 (2005). [PubMed: 15692063]
23. Wang L, Ni X, Covington KR, Yang BY, Shiu J, Zhang X, Xi L, Meng Q, Langridge T, Drummond J, Donehower LA, Doddapaneni H, Muzny DM, Gibbs RA, Wheeler DA, Duvic M, Genomic profiling of Sézary syndrome identifies alterations of key T cell signaling and differentiation genes, *Nat. Genet* 47, 1426–1434 (2015). [PubMed: 26551670]
24. Jackow CM, Cather JC, Hearne V, Asano AT, Musser JM, Duvic M, Association of erythrodermic cutaneous T-cell lymphoma, superantigen-positive *Staphylococcus aureus*, and oligoclonal T-cell receptor V beta gene expansion, *Blood* 89, 32–40 (1997). [PubMed: 8978274]
25. Harrell FE, Lee KL, Mark DB, Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors, *Stat. Med* 15, 361–387 (1996). [PubMed: 8668867]
26. Olsen EA, Whittaker S, Kim YH, Duvic M, Prince HM, Lessin SR, Wood GS, Willemze R, Demierre M-F, Pimpinelli N, Bernengo MG, Ortiz-Romero PL, Bagot M, Estrach T, Guitart J, Knobler R, Sanches JA, Iwatsuki K, Sugaya M, Dummer R, Pittelkow M, Hoppe R, Parker S, Geskin L, Pinter-Brown L, Girardi M, Burg G, Ranki A, Vermeer M, Horwitz S, Heald P, Rosen S, Cerroni L, Dreno B, Vonderheid EC, Clinical end points and response criteria in mycosis fungoides and Sézary syndrome: a consensus statement of the International Society for Cutaneous Lymphomas, the United States Cutaneous Lymphoma Consortium, and the Cutaneous Lymphoma Task Force of the European Organisation for Research and Treatment of Cancer, *J. Clin. Oncol* 29, 2598–2607 (2011). [PubMed: 21576639]
27. Tumeh PC, Harview CL, Yearley JH, Shintaku IP, Taylor EJM, Robert L, Chmielowski B, Spasic M, Henry G, Ciobanu V, West AN, Carmona M, Kivork C, Seja E, Cherry G, Gutierrez AJ, Grogan TR, Mateus C, Tomasic G, Glaspy JA, Emerson RO, Robins H, Pierce RH, Elashoff DA, Robert C, Ribas A, PD-1 blockade induces responses by inhibiting adaptive immune resistance, *Nature* 515, 568–571 (2014). [PubMed: 25428505]
28. Ungewickell A, Bhaduri A, Rios E, Reuter J, Lee CS, Mah A, Zehnder A, Ohgami R, Kulkarni S, Armstrong R, Weng W-K, Gratzinger D, Tavallaee M, Rook A, Snyder M, Kim Y, Khavari PA, Genomic analysis of mycosis fungoides and Sézary syndrome identifies recurrent alterations in TNFR2, *Nat. Genet* 47, 1056–1060 (2015). [PubMed: 26258847]
29. da Silva Almeida AC, Abate F, Khiabani H, Martinez-Escala E, Guitart J, Tensen CP, Vermeer MH, Rabadan R, Ferrando A, Palomero T, The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome, *Nat. Genet* 47, 1465–1470 (2015). [PubMed: 26551667]
30. Choi J, Goh G, Walradt T, Hong BS, Bunick CG, Chen K, Bjornson RD, Maman Y, Wang T, Tordoff J, Carlson K, Overton JD, Liu KJ, Lewis JM, Devine L, Barbarotta L, Foss FM, Subtil A, Vonderheid EC, Edelson RL, Schatz DG, Boggon TJ, Girardi M, Lifton RP, Genomic landscape of cutaneous T cell lymphoma, *Nat. Genet* 47, 1011–1019 (2015). [PubMed: 26192916]

31. Hoppe RT, Medeiros LJ, Warnke RA, Wood GS, CD8-positive tumor-infiltrating lymphocytes influence the long-term survival of patients with mycosis fungoides, *J. Am. Acad. Dermatol* 32, 448–453 (1995). [PubMed: 7868714]
32. Vermeer MH, van Doorn R, Dukers D, Bekkenk MW, Meijer CJ, Willemze R, CD8+ T cells in cutaneous T-cell lymphoma: expression of cytotoxic proteins, Fas Ligand, and killing inhibitory receptors and their relationship with clinical behavior, *J. Clin. Oncol* 19, 4322–4329 (2001). [PubMed: 11731515]
33. Vallois D, Dobay MPD, Morin RD, Lemonnier F, Missiaglia E, Juilland M, Iwaszkiewicz J, Fataccioli V, Bisig B, Roberti A, Grewal J, Bruneau J, Fabiani B, Martin A, Bonnet C, Michielin O, Jais J-P, Figeac M, Bernard OA, Delorenzi M, Haioun C, Tournilhac O, Thome M, Gascoyne RD, Gaulard P, de Leval L, Activating mutations in genes related to TCR signaling in angioimmunoblastic and other follicular helper T-cell-derived lymphomas, *Blood* 128, 1490–1502 (2016). [PubMed: 27369867]
34. Carlson CS, Emerson RO, Sherwood AM, Desmarais C, Chung M-W, Parsons JM, Steen MS, LaMadrid-Herrmannsfeldt MA, Williamson DW, Livingston RJ, Wu D, Wood BL, Rieder MJ, Robins H, Using synthetic templates to design an unbiased multiplex PCR assay, *Nat. Commun* 4, 2680 (2013) [PubMed: 24157944]

Figure 1. High throughput TCRB sequencing in 309 patients with cutaneous T cell lymphomas
 A. Clinical diagnosis in 309 patients with cutaneous T cell lymphomas in the discovery and validation sets. Other: CD30+ lymphoproliferative disorder and CD8+ aggressive epidermotropic cutaneous T cell lymphoma. Pre-Sézary refers to the evidence of blood abnormalities (B1; elevated absolute CD4 T cell count or CD4/CD8 T cell ratio) that do not meet the criteria for stage B2 or Sézary syndrome (26). B. TCRBV gene family usage by the malignant clone in 309 cases of primary cutaneous T cell lymphomas. C. Example of the measurement of the malignant clone frequency in skin in two patients with stage IB mycosis

fungoides. Clinical pictures and 3D histograms of the TCRB sequencing data in lesional skin in two patients with stage IB mycosis fungoides. On the upper panel, the 3D histogram shows the presence of a tumor clone frequency of 8% (18,131 reads). This patient showed no evidence of disease progression after 4 years follow-up. On the lower panel, the 3D histogram shows a tumor clone frequency of 37% (264,252 reads - the y-axis has been cut at 80,000). This patient died of disease progression after 28 months. D. Hematoxylin-eosin sections of lesional skin biopsies in 4 patients with cutaneous T cell lymphoma and various malignant clone frequencies and outcomes, magnification x10 Upper left. Malignant clone 32% of the T cells. Progression after 2 years Upper right. Malignant clone 41% of the T cells. Progression after 2 months Lower left. Malignant clone 6% of the T cells. No progression in 8 years Lower right. Malignant clone 6% of the T cells. No progression in 9 years E. Malignant clone frequency according to the extent of body surface area involved in patients with mycosis fungoides Medians are indicated by horizontal bars and comparisons are carried out using Mann-Whitney U-test, * $p < 0.05$ considered significant.

Figure 2. The tumor clone frequency in skin as predictor of progression-free and overall survival in patients with cutaneous T cell lymphomas.

A. Kaplan-Meier estimates of progression-free (left panel) and overall survival (right) in 208 patients with cutaneous T cell lymphomas in the discovery set, according to the tumor clone frequency in skin ($<25\%$ versus $>25\%$ of the total T cells in skin). B. Kaplan-Meier estimates of progression-free survival in 101 patients with cutaneous T cell lymphomas in the validation set, according to the tumor clone frequency in skin ($<25\%$ versus $>25\%$ of the total T cells in skin). C. Kaplan-Meier estimates of progression-free (left panel) and overall survival (right) in 177 patients with mycosis fungoides in the discovery set, according to the

tumor clone frequency in skin (<25% versus >25% of the total T cells in skin). D. Kaplan-Meier estimates of progression-free survival in 87 patients with mycosis fungoides in the validation set, according to the tumor clone frequency in skin (<25% versus >25% of the total T cells in skin). p-values in A-D are estimated by Cox univariable analysis. E. Kaplan-Meier estimates of progression-free (left panel) and overall survival (right) in 22 patients with Sézary syndrome in the discovery set, according to the tumor clone frequency in skin (<25% versus >25% of the total T cells in skin).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3. The tumor clone frequency in skin as predictor of progression-free and overall survival in patients with early-stage mycosis fungoides

A. Kaplan-Meier estimates of progression-free (left) and overall survival (right) in 141 patients with early-stage (IA to IIA) mycosis fungoides in the discovery set, according to the tumor clone frequency (<25% versus >25% of the total T cells in skin). B. Kaplan-Meier estimates of progression-free survival in 69 patients with early-stage (IA to IIA) mycosis fungoides in the validation set, according to the tumor clone frequency (<25% versus >25% of the total T cells in skin). C. Kaplan-Meier estimates of progression-free (left) and overall survival (right) in 70 patients with stage IB mycosis fungoides in the discovery set,

according to the tumor clone frequency (<25% versus >25% of the total T cells in skin, upper panels) or to the presence of plaques (lower panels). D. Kaplan-Meier estimates of progression-free survival in 42 patients with stage IB mycosis fungoides in the validation set, according to the tumor clone frequency (<25% versus >25% of the total T cells in skin, upper panel) or to the presence of plaques (lower panel). p-values in A-D are estimated by Cox univariable analysis. E. Dot plot and linear regression of the time to progression/death according to the tumor clone frequency in skin in stage IB patients from the discovery and validation sets, who experienced disease progression during the follow-up. Pearson's correlation coefficient and p-value are indicated. F. Receiver operating characteristic curve of the tumor clone frequency in skin (>25%) in patients with stage IB mycosis fungoides in the discovery and validation sets for 5-year progression or death. Progressors are patients who progressed or died within 5 years after the test. Nonprogressors are patients with at least 5 years of follow-up and no event of death or progression in 5 years. The sensitivity is defined as the percentage of patients with a malignant clone >25% of T cells in skin among progressors. The specificity is defined as the percentage of patients with a malignant clone <25% of T cells in skin among nonprogressors.

Figure 4. Samples with a high tumor clone frequency are not associated with a decreased anti-tumor immune response

A. Example of CD8+ and granzyme immunostaining in lesional skin in 2 lesional CTCL skin biopsies. B. Percentage of CD8+ T cell % (left) and granzyme B positive cell % in lesional skin of CTCL patients with a low tumor clone frequency (<10% T cells) and high tumor clone frequency (>30% T cells). (Mann Whitney U-test, * $p < 0.05$ ** $p < 0.01$). C. Reactive T cell clonality (left) and entropy (right) in lesional skin of CTCL patients with a low tumor clone frequency (<10% T cells) and high tumor clone frequency (>30% T cells) (Mann Whitney U-test, * $p < 0.05$).

$p < 0.05$ ** $p < 0.01$ C. Dot plots of the tumor clone frequency (TCF) in patients in cluster 1, 2 and 3. Means were compared by Mann-Whitney U-test with Bonferroni adjustment for multiple testing, * $p < 0.05$ ** $p < 0.01$ D. Kaplan-Meier estimates of progression-free survival in 157 patients with cutaneous T cell lymphomas in the training group, according to the gene expression clustering. Log-rank test with Bonferroni adjustment for multiple testing, * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$ E. Whole exome sequencing data of microdissected skin T cells in patients with mycosis fungoides. Number of somatic mutations according to the clinical stage (left). Mann-Whitney U-test, * $p < 0.05$ Number of somatic mutations according to the malignant clone frequency in skin (right). Spearman correlation, $p < 0.05$ considered significant

Table 1.

Uni- and multivariable analysis on progression-free and overall survival in 177 patients with mycosis fungoides in the discovery set

		Discovery set (n=177)											
		Progression-free survival						Overall survival					
		Univariable			Multivariable			Univariable			Multivariable		
		HR	95%CI	p	HR	95%CI	p	HR	95%CI	p	HR	95%CI	p
Tumor clone frequency	>15%	2.1	1.3–3.5	.002				2.5	1.4–4.7	.003			
	>25%	4.0	2.4–6.8	<.001	3.3	1.9–5.9	<.001	4.8	2.6–9.0	<.001	5.1	2.5–10	<.001
	>35%	4.6	2.6–8.1	<.001				5.2	2.7–10	<.001			
LDH levels	Elevated versus normal	2.3	1.2–4.3	.008				3.2	1.6–6.5	.001			
Large-cell transformation	Presence	3.4	1.8–6.4	<.001	1.8	.9–3.7	.1	3.3	1.6–7.0	.001	1.3	.5–3.0	.6
Age	>60 years	2.2	1.3–3.7	.002	2.3	1.3–4.0	.003	3.4	1.7–6.7	.0005	3.5	1.7–7.4	<.001
Treatments	Phototherapy	.9	.5–1.5	.7				.8	.4–1.5	.5			
	Radiation therapy	1.9	1.1–3.4	.02	1.4	.7–2.5	.3	1.8	.9–3.8	.1			
	Systemic treatments	3.4	2.1–5.5	<.001				3.9	2.1–7.1	<.001			
PCR	Clonal pattern	1.5	.7–3.4	.3				1.0	.5–4.2	.4			
Gender	Male vs. female	1.5	.9–2.5	.1				1.2	.7–2.4	.5			
Folliculotropism	Presence	1.5	.9–2.5	.1				1.2	.6–2.3	.6			

Abbreviations: LDH, lactate dehydrogenases; PCR, polymerase chain reaction; CI, confidence interval; HR, hazard ratio. Treatments used before first evidence of progression, death or censoring. Phototherapy includes PUVA (psoralen+Ultraviolet A) and UVB therapy. Radiation therapy includes electron-beam therapy and brachytherapy. Systemic treatments include α -interferon, oral bexarotene, folate inhibitors, systemic histone deacetylase inhibitors and monoclonal antibodies. The multivariable model was stratified on LDH levels and the use of systemic treatments.

Table 2.

Prognostic factors of progression-free survival in 141 patients with early-stage disease from the training set and 69 patients in the validation set

	Progression-free survival					
	Discovery set (n=141)			Validation set (n=69)		
	HR	95% CI	p	HR	95% CI	p
Tumor clone frequency >25%	4.9	2.5 – 9.7	<.001	10	3.4 – 31	<.001
Stage (IB versus IA)	2.5	1.3 – 4.9	.008	6.4	1.5 – 28	.01
Presence of plaques	2.2	1.1 – 4.2	.02	1.9	0.8 – 4.4	.15
Elevated LDH levels	1.2	0.4 – 3.1	.8	1.0	.3 – 3.4	1
Large-cell transformation	1.5	0.4 – 6.4	.6	1.6	.2 – 12	.6
Age>60	2.0	1.0 – 3.7	.04	1.3	0.5 – 3.0	.6
CLIPi score *						
Intermediate risk (versus low)	2.2	1.0 – 5.0	.05	.4	0.1 – 2.0	.2
High risk (versus low)	3.5	1.6 – 7.7	.002	2.3	0.9 – 5.9	.07

Abbreviations: HR, hazard ratio; CI, confidence interval; LDH, lactate dehydrogenases; CLIPi, Cutaneous Lymphoma International Prognostic Index.

* The Cutaneous Lymphoma International Prognostic Index (CLIPi) was calculated based on the presence of the following factors: age>60 years, male sex, plaques, folliculotropism and clinical adenopathy N1/Nx; low risk=0–1, intermediate risk=2, high risk=3–5 prognostic factors.