npg

ORIGINAL ARTICLE

Minimal residual disease quantification using consensus primers and high-throughput IGH sequencing predicts post-transplant relapse in chronic lymphocytic leukemia

AC Logan[1], B Zhang[2], B Narasimhan[3], V Carlton[4], J Zheng[4], M Moorhead[4], MR Krampf[1], CD Jones[2], AN Waqar[2], M Faham[4], JL Zehnder[2] and DB Miklos[1]

Quantification of minimal residual disease (MRD) following allogeneic hematopoietic cell transplantation (allo-HCT) predicts post-transplant relapse in patients with chronic lymphocytic leukemia (CLL). We utilized an MRD-quantification method that amplifies immunoglobulin heavy chain (IGH) loci using consensus V and J segment primers followed by high-throughput sequencing (HTS), enabling quantification with a detection limit of one CLL cell per million mononuclear cells. Using this IGH–HTS approach, we analyzed MRD patterns in over 400 samples from 40 CLL patients who underwent reduced-intensity allo-HCT. Nine patients relapsed within 12 months post-HCT. Of the 31 patients in remission at 12 months post-HCT, disease-free survival was 86% in patients with MRD $<10^{-4}$ and 20% in those with MRD $\geqslant 10^{-4}$ (relapse hazard ratio (HR) 9.0; 95% confidence interval (CI) 2.5–32; $P < 0.0001$), with median follow-up of 36 months. Additionally, MRD predicted relapse at other time points, including 9, 18 and 24 months post-HCT. MRD doubling time <12 months with disease burden $\geqslant 10^{-5}$ was associated with relapse within 12 months of MRD assessment in 50% of patients, and within 24 months in 90% of patients. This IGH–HTS method may facilitate routine MRD quantification in clinical trials.

INTRODUCTION

Chronic lymphocytic leukemia (CLL) is a clonal B-cell malignancy with a broad spectrum of clinical presentations. Those with high-risk disease are frequently treated with allogeneic hematopoietic cell transplantation (allo-HCT) after relapse or early treatment failure.[1,2] An important goal in the management of CLL patients after allo-HCT is the ability to monitor minimal residual disease (MRD) in order to predict clinical relapse.[3] Available methods for quantifying MRD following treatment largely fall into two categories: (1) those based on multiparameter flow cytometry[4,5] and, (2) those based on genetic methods for quantifying clonally rearranged immunoglobulin genes associated with the CLL clone. Both methodologies have been shown to predict future relapse when disease burden is $\geqslant 10^{-4}$ 12 months following allo-HCT.[3,4,6,7]

Until recently, quantification of a CLL-specific immunoglobulin heavy chain (IGH) rearrangement (that is, IGH clonotype) required a lengthy process of developing patient-specific primers and quantitative PCR probes to perform allele-specific oligonucleotide polymerase chain reaction (ASO-PCR).[8,9] It is costly to maintain the expertise and technical capacity to develop, validate and routinely run patient-specific assays. Thus, a universal PCR-based assay permitting patient-specific decision-making is preferable, and will increase the availability of MRD quantification for patients with CLL and potentially other B-cell malignancies.

To facilitate accessibility to universally applicable, high-sensitivity MRD quantification in CLL, our group has pioneered methods for using degenerate consensus primers (that is, sets of primers that can be used for any patient) to amplify all IGH genes in a mixture of polyclonal lymphoid cells, followed by massively parallel high-throughput sequencing (HTS) of the resulting immunoreceptor amplimers.[10,11] In a previous study, we used BIOMED-2 consensus primers[12] and 454 pyrosequencing to sequence the IGH locus from six somatically unmutated CLL patients after allo-HCT, and we developed a bioinformatics algorithm for processing the resulting sequence data to quantify MRD in a prospective manner with 10^{-5} sensitivity.[11] The IGH–HTS approach demonstrated sensitivity equivalent to ASO-PCR, without the need for patient-specific reagents or procedures.[11]

We now apply the LymphoSIGHT method,[13] an IGH–HTS MRD platform with a validated detection limit of 10^{-6} and quantitative range above 10^{-5}, to the prediction of relapse in 40 patients who underwent reduced-intensity allo-HCT for high-risk CLL, and we evaluate characteristics of molecular disease progression to better predict the timing of clinical relapse.

PATIENTS AND METHODS

Patient characteristics

Patients evaluated in this study represent consecutive CLL patients transplanted at Stanford between July 2005 and December 2010 on

[1]Department of Medicine, Division of Blood and Marrow Transplantation, Stanford University School of Medicine, Stanford, CA, USA; [2]Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA; [3]Department of Health Research and Policy-Biostatistics, Stanford University School of Medicine, Stanford, CA, USA and [4]Sequenta, Inc, South San Francisco, CA, USA. Correspondence: Dr DB Miklos, Department of Medicine, Division of Blood and Marrow Transplantation, Stanford University School of Medicine, 269 West Campus Drive, CCSR 2205, MC 5462, Stanford, CA 94305, USA.
E-mail: dmiklos@stanford.edu

clinical trials evaluating outcomes of related and unrelated donor allo-HCT, following preparation with total lymphoid irradiation (TLI) and antithymocyte globulin (ATG). TLI/ATG conditioning consisted of 800 cGy TLI (80 cGy for 10 days) and rabbit ATG (thymoglobulin; Genzyme, Cambridge, MA, USA) 1.5 mg kg^{-1} for 5 days, as previously described.[14,15] Grafts were comprised of unmanipulated granulocyte colony-stimulating factor (G-CSF)-mobilized peripheral blood collected by apheresis. Some patients received post-HCT rituximab 375 mg m^{-2} weekly, beginning on day +56 for a total of four doses.[16] Primary graft-versus-host disease (GVHD) prophylaxis consisted of cyclosporine (tapered off beginning 6 months post-HCT) and mycophenolate mofetil (tapered at 1–2 months post-HCT, depending on whether the donor was related or unrelated). All patient and donor samples were obtained with explicit authorization and monitoring by the Stanford University School of Medicine's Institutional Review Board.

MRD samples
Peripheral blood mononuclear cells (PBMC) were prospectively cryopreserved prior to initiation of the transplant-conditioning regimen (day-12), and at months +1, +3, +6, +9, +12, +18, +24, +30, +36, and other times as deemed clinically appropriate. PBMC were generated from 10 ml of whole blood as described in Supplementary Methods.

IGH–HTS
IGH molecules were quantified using the LymphoSIGHT IGH–HTS method as previously described.[13,17] PBMC genomic DNA was spiked with a specific number of copies of a reference IGH sequence, and this mixture was then amplified with three sets of multiplexed degenerate primers in the IGH V region and a single J primer to generate the IGH library (Figure 1). The use of multiple primer sets mitigates the risk of non-amplification with any one primer set due to the presence of somatic mutations in those regions. See Supplementary Methods for a description of the primer design, amplification sequencing reactions and bioinformatic clonotype quantification.

IGH ASO-PCR
Quantitative real-time ASO-PCR was performed as previously described.[18] The CLL clone-specific IGH gene sequence was amplified using BIOMED-2 primers[12] and identified with heteroduplex analysis of the diagnostic DNA sample followed by Sanger's sequencing of the clonal band.[19] This CLL clonal sequence was then used to design a patient-specific real-time quantitative PCR assay as described in Supplementary Methods.

Definition of outcomes
Our primary outcome of interest was disease-free survival (DFS), which was defined as the number of days between graft infusion (day 0) and clinical relapse, as determined by accepted clinical criteria.[3] Overall survival (OS) was defined as the number of days between graft infusion and death from any cause. Mixed chimerism was defined as <95% donor cells using short-tandem repeat analysis.[15] Acute GVHD grades II–IV were graded clinically according to accepted criteria.[20] Chronic GVHD was based on the National Institute of Health's consensus guidelines.[21] Thresholds for MRD positivity were explored and are described in the Results.

Statistical analysis
DFS and OS were estimated by the Kaplan–Meier method.[22] Patients relapsing or dying from other causes prior to specified landmarks were censored from analyses where appropriate. Factors included in univariate analysis for prediction of post-transplant relapse included MRD status at 12 months, fludarabine refractoriness, 17p deletion, 11q deletion, post-transplant rituximab exposure, CD3 chimerism at month +1, CD19 chimerism at month +1, CD3 chimerism at month +12 and CD19 chimerism at month +12. GraphPad Prism (GraphPad Software, La Jolla, CA, USA) was used to generate Kaplan–Meier curves and other figures after primary analysis in R[23] or Spotfire (Tibco, Somerville, MA, USA).

RESULTS
Technical performance of IGH–HTS MRD quantification
Some performance characteristics of the LymphoSIGHT quantification method have been previously evaluated, and if sufficient material is available, the assay can detect residual disease at the 10^{-6} level in dilution series with acute lymphoblastic leukemia samples.[13] As this is the first published use in CLL, we again tested
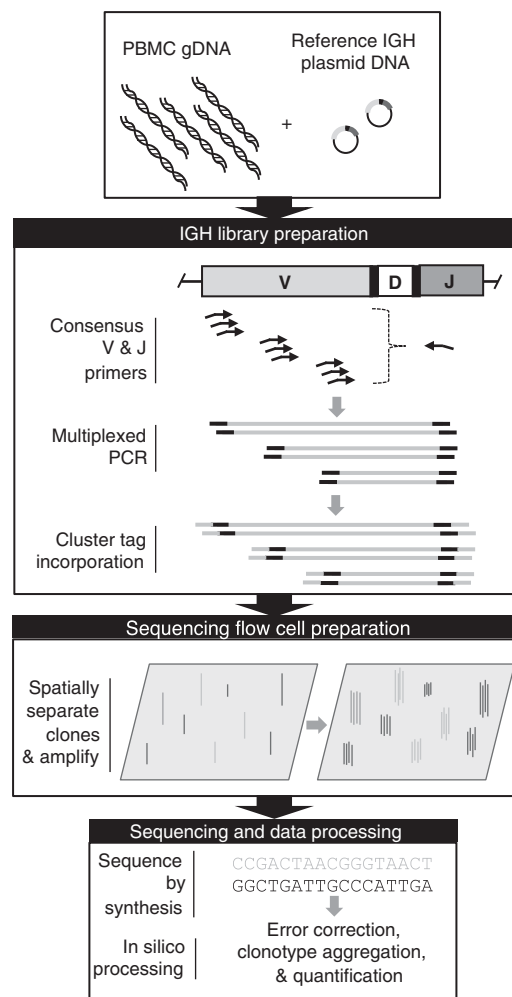

Figure 1. Overview of the LymphoSIGHT IGH–HTS method. PBMC genomic DNA (gDNA) was mixed with a set of reference IGH plasmids at known concentration. IGH alleles in this mixture were then amplified with consensus IGH V and J primers that append the annealing sites for secondary PCR primers. The PCR products from the first amplification were then amplified in a second PCR reaction using primers which append sample indices and cluster-formation sequences (cluster tag incorporation). These amplimers were then annealed to an Illumina Genome Analyzer-sequencing lane and locally amplified in situ via bridging PCR. Sequencing primers annealing to the second-round primer at each end of the IGH amplimers were used to sequence by synthesis the ends of paired strands. These data were then processed bioinformatically to map IGH V and J sequences to IMGT germline sequences, further analyzed to aggregate clonotypes to remove artifacts, and finally quantified.

the 10^{-6} detection limit with a dilution curve generated from fluorescence-activated cell sorting-purified CLL cells (CD19$^+$ CD5$^+$ CD23$^+$) from a patient that were mixed with a normal healthy donor's PBMC sample having a B-cell content of 10% (based on CD19 expression). Serial dilutions of leukemic cells, ranging from 10^{-3} to 10^{-6} CLL cells per mononuclear cell, were prepared. These mixtures were then amplified, sequenced and quantified, demonstrating the recovery of the expected quantity of CLL-specific clonotypes in each reaction (Figure 2a).

A bioinformatics algorithm was used to determine how to treat two apparently related clonotype species based on their relative abundance. As detailed in the Supplementary Methods, low-frequency clonotypes that differ by up to 1% from another clonotype with significantly higher frequency are deemed to be the result of technical artifact, and the sequence differences are
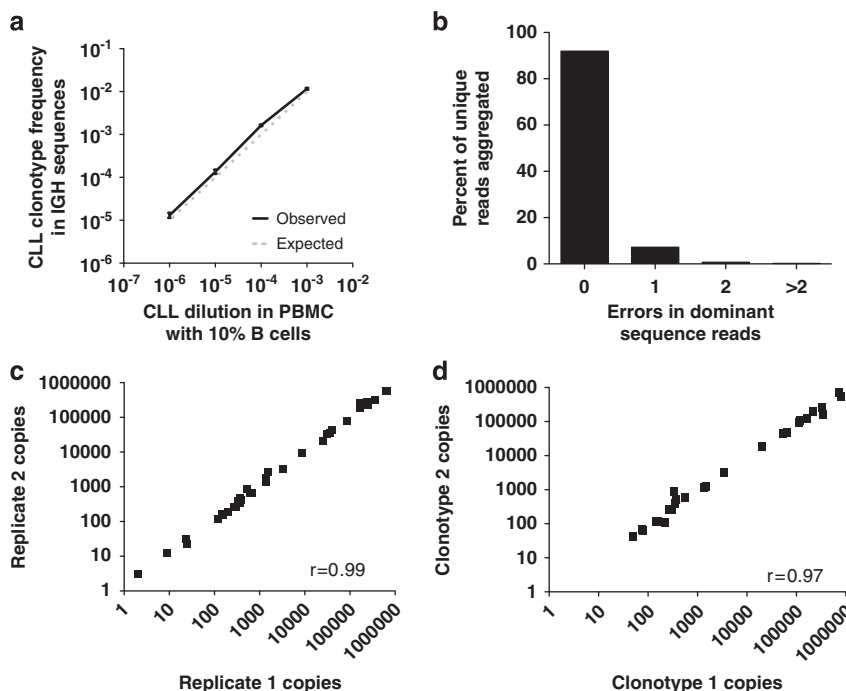
Figure 2. LymphoSIGHT MRD quantification technical performance. (**a**) CLL cells were purified by fluorescence-activated cell sorting and then diluted into normal human PBMC with 10% B-cell content to levels of 10^{-3}, 10^{-4}, 10^{-5} and 10^{-6} CLL cells per leukocyte. DNA was harvested as described in the Patients and methods, and consensus PCR followed by Illumina HTS was performed in quadruplicate and processed bioinformatically for clonotype quantification. Quantifications of the diluted CLL clonotype in each mixture were performed and are graphed $+/-$ s.e.m. (error bars) as a percentage of total IGH gene sequences. The number of CLL clonotypes expected to be recovered at each dilution are shown (dashed line). (**b**) Error frequencies in CLL clonotype reads in diagnostic samples are shown. (**c**) The quantification of CLL clonotypes in samples subjected to two entirely separate genomic DNA-to-sequence replicates from several patients are shown. The correlation between specific CLL clonotype quantification between replicate 1 and 2 was high ($r = 0.99$). (**d**) The correlation between quantification of both IGH alleles in samples from five patients with biallelically rearranged CLL was high ($r = 0.97$).

corrected for clonotype aggregation and quantification. More than 90% of the CLL clonotypes were error-free and the majority of remaining clonotypes contained just one base error needing correction, with a minute percentage ($<1\%$) containing two or more errors (Figure 2b).

To determine assay reproducibility, we analyzed a subset of patient samples in two independent genomic DNA-to-sequence replicates. Samples spanning the entire dynamic range of the test were randomly selected and the inter-replicate correlation was very high ($r = 0.99$) (Figure 2c). We also assessed the quantification of both IGH genes in samples from five biallelically rearranged CLL patients. The clonotype quantification algorithm categorized sequences as 'non-productive' if they possess a premature stop codon within the coding sequences. Non-productive IGH sequences arise biologically from the failed IGH rearrangements and are sometimes accompanied by another productively rearranged IGH allele. By quantifying the non-productively rearranged allele in biallelically rearranged CLL, we found that quantifications of the two alleles correlated very well ($r = 0.97$) (Figure 2d).

Twenty-five of the patients we studied had an ASO-PCR reaction developed for the purpose of monitoring MRD. The correlation between MRD quantification by ASO-PCR and IGH–HTS was good (Pearson $r = 0.64$, $P < 0.0001$; Figure 3). Sixteen samples (9.2% of the 174 samples evaluated by both ASO-PCR and IGH–HTS) deemed negative by ASO-PCR were found to have detectable disease by IGH–HTS. The burden of disease in these samples was mostly $<10^{-4}$, demonstrating improved sensitivity using IGH–HTS.

Patient and sample characteristics for IGH–HTS MRD analysis

All patients evaluated in this retrospective MRD-quantification study were allografted for high-risk CLL using reduced-intensity
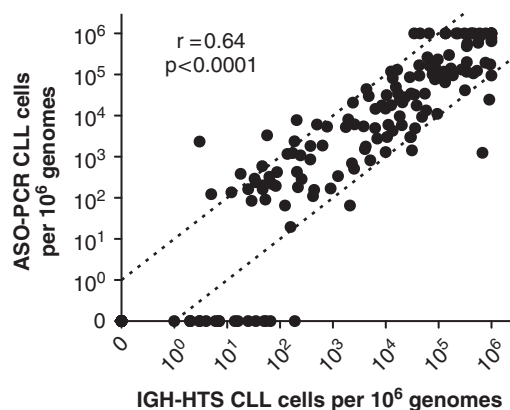

Figure 3. Correlation between IGH–HTS and ASO-PCR. Post-transplant PBMC or whole-blood samples were processed into DNA, and analyzed for MRD by IGH–HTS and ASO-PCR. The correlation between quantification of CLL clonotypes per one million leukocyte genomes using the two techniques was high ($r = 0.64$; $P < 0.0001$). Samples with <1 log differences in quantification using the two methods are shown between the dashed lines. Ninety-six percent of samples exhibited <1 log differences between the two assays (82% <0.5 log, 14% 0.5–1 log).

conditioning with TLI and antithymocyte globulin (TLI/ATG).[14,15] Patient characteristics are summarized in Table 1 and are detailed in Supplementary Table 1. Forty-two sequentially transplanted CLL patients were selected for this study; however, two were removed due to unavailability of a sample with sufficient disease burden to determine their CLL clonotype. Twenty-one patients were treated

Table 1. Summary of patient characteristics

	n (%)
Total patients	40
Recipient gender, male	29 (73%)
Age, median (range)	56 (34–67)
CLL characteristics	
Fludarabine refractoriness	24 (60%)
17p del	12 (30%)
11q del	10 (25%)
Unmutated IGH	36 (90%)
Status at HCT	
CR (any, range 1-2)	8 (20%)
PR (any, range 1-7)	30 (75%)
PD	2 (5%)
Donor type	
Related donor	20 (50%)
Unrelated donor	20 (50%)
Post-HCT rituximab	21 (53%)
Follow-up, median months (range)	50 (24–88)

Figure 4. Twelve-month post-transplant landmark MRD analysis. Kaplan–Meier analysis of DFS is shown for 15 MRD-negative ($<10^{-6}$) and 16 MRD-positive ($\geqslant 10^{-6}$) patients. The curves are significantly different (HR 7.9; 95% CI 2.3–26; $P = 0.0002$).

on a clinical trial in which they received rituximab on days $+56$, $+63$, $+70$ and $+77$ to evaluate the utility of this approach for prevention of chronic GVHD.[16]

Eight patients (20%) proceeded to transplant in complete remission (CR), and five (13%) were in molecular remission prior to transplant, whereas two had progressive disease (PD) and 30 achieved a partial response (PR) to the last therapy. Grade II–IV acute GVHD occurred in 1/40 (2.5%) and new onset-chronic GVHD (not associated with donor lymphocyte infusion for relapse) occurred in 10/40 (25%). Nine of the 17 patients (53%) who received donor lymphocyte infusion for relapse subsequently developed GVHD. Individual transplant characteristics are shown in Supplementary Table 1. DFS and OS are 50 and 75%, respectively, with median follow-up time of 50 months (24–88 months) (Supplementary Figure 1). Relapse beyond 6 months was not a predictor of OS based on available follow-up ($P = 0.14$); thus, this study was designed to assess the relationship between MRD progression and DFS.

A total of 403 PBMC samples were retrieved from cryopreservation and analyzed for this study. MRD was quantified using DNA templates such that each cell may only be counted once. Samples analyzed by IGH–HTS contained a median of 2.7×10^6 PBMC genomes (2.6×10^4–2.4×10^7), with a median of 7.6×10^5 IGH molecules (12–2.8×10^7).

Twelve-month landmark IGH–HTS MRD predicts relapse

Nine of the 40 patients (22.5%) relapsed prior to 12 months post-HCT. We performed a 12-month landmark analysis on the remaining 31 patients to determine the usefulness of MRD status at that time point for predicting subsequent relapse. Ten patients had IW-CLL (International Workshop on CLL)-defined MRD positivity ($\geqslant 10^{-4}$), with 20% remaining disease-free, whereas 86% of MRD-negative patients ($<10^{-4}$) remain in long-term remission (relapse HR 9.0, 95% CI 2.5–32, $P < 0.0001$). Using the full detection depth of the LymphoSIGHT assay, 16 patients deemed MRD positive with at least one CLL clone per million PBMC genomes have 37.5% DFS, whereas 15 patients with MRD $<10^{-6}$ have 93.3% DFS (HR 7.9, 95% CI 2.3–26; $P = 0.0002$), with a median follow-up of 50 months (range 24–88) for non-relapsed patients (Figure 4).

IGH–HTS MRD predicts relapse at other post-transplant landmarks

We evaluated whether MRD status at other time points is significantly associated with subsequent relapse. The first 9 months after TLI/ATG-conditioned allo-HCT represent a dynamic
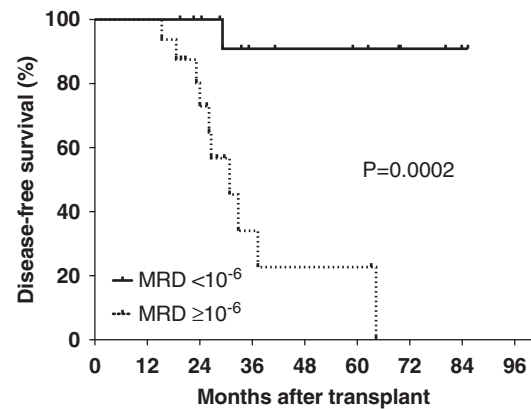
period for IGH clonotype quantifications due to variable immunosuppression tapering and engraftment kinetics, thus prediction of relapse before this time point may be challenging and was not possible in our data set. Beginning at 9 months post transplant, however, detectable MRD at the 10^{-6} level or higher conveys a significant risk for relapse (HR 5.4; 95% CI 1.4–21; $P = 0.01$). As expected, detectable MRD at landmarks following 12 months, including 18 (HR 22; 95% CI 4.8–99; $P < 0.0001$) and 24 months (HR 49; 95% CI 8.8–272; $P < 0.0001$), is significantly associated with relapse as well.

Post-transplant MRD patterns in relapsed and non-relapsed patients

The nine patients who relapsed prior to 12 months post-HCT exhibited rapid changes in MRD (Supplementary Figure 2A). Eleven patients relapsed beyond 12 months post-HCT, and these patients exhibited a wide spectrum of relapse kinetics (Supplementary Figure 2B). We observed that disease burden fluctuated substantially immediately post-transplant, but that the trajectory of disease burden becomes more stable in each patient at a later time (generally, 9–12 months post-HCT and beyond). Non-relapsed patients generally achieved MRD $<10^{-6}$ (Supplementary Figure 2C). One patient (SPN3723) in this group had clearly progressive molecular relapse with disease burden roughly 10^{-2} in the setting of intense immunosuppressive treatment for chronic GVHD and died from treatment complications 18 months following HCT without meeting the criteria for clinical relapse (designated in Supplementary Figure 2C). This patient is excluded from subsequent analyses.

MRD kinetics and time to clinical relapse

We next assessed whether serial sampling could facilitate the prediction of time-to-relapse in patients with MRD detectable post-HCT. We used two metrics for this analysis: the level of disease burden and the change in disease burden over time. We believe the ability to predict rapid relapse (that is, within 12–24 months of an MRD assessment) will be clinically useful, because pre-emptive therapeutic maneuvers such as immunosuppression taper, donor lymphocyte infusion or additional chemo-, immuno- or biological therapy may be applied to delay or prevent clinical relapse.

As CLL doubling time at high disease burdens (that is, assessable by standard blood count methodologies) is associated strongly with prognosis,[24] we first determined whether MRD doubling may be a useful metric. Owing to standard errors of measurement at the limit of detection ($\sim 10^{-6}$), the

LymphoSIGHT method is qualitative below 10^{-5}, but quantitative at 10^{-5} and higher.[13] MRD doubling time was < 12 months in 18/20 (90%) of relapsed patients.

Our analysis above demonstrated that a binary assessment of MRD prior to 9 months post-HCT does not predict relapse. We evaluated whether MRD trajectory improved the utility of these early MRD assessments. Five of the 20 (25%) non-relapsed patients and 6/6 (100%) patients who relapsed before 9 months post-HCT exhibited MRD doubling, with MRD later falling in the non-relapsed patients. Thus, MRD doubling was only associated with relapse in 6/11 (55%) patients during the first 9 months, suggesting that additional measurements or other markers of disease activity will be necessary to identify patients likely to experience relapse in this time-frame with high positive predictive value. Assessment of MRD trajectory after 9 months post-HCT, on the other hand, was more informative, with ten patients exhibiting MRD doubling of less than 12 months, five (50%) of whom relapsed within 12 months of MRD sampling and nine (90%) relapsed within 24 months.

Amongst patients with MRD $\geqslant 10^{-5}$ beyond 9 months post transplant, stratified by MRD doubling (ie, MRD slope $\geqslant 2$-fold per year), enriched for those relapsing within 12 months of sample acquisition (Figure 5a). MRD doubling was associated with relapse within 12 months of sample acquisition for 50% with MRD $\geqslant 10^{-5}$, 63% with MRD $\geqslant 10^{-4}$, 75% with MRD $\geqslant 10^{-3}$, 88% with MRD $\geqslant 10^{-2}$ and 100% with MRD $\geqslant 10^{-1}$, whereas relapse occurred within 24 months of MRD doubling for 90% with MRD $\geqslant 10^{-5}$ and 100% with MRD $\geqslant 10^{-4}$ (Figure 5b). We believe that the combination of disease burden and slope of MRD progression improves the accuracy with which patients at risk of rapid relapse may be identified.
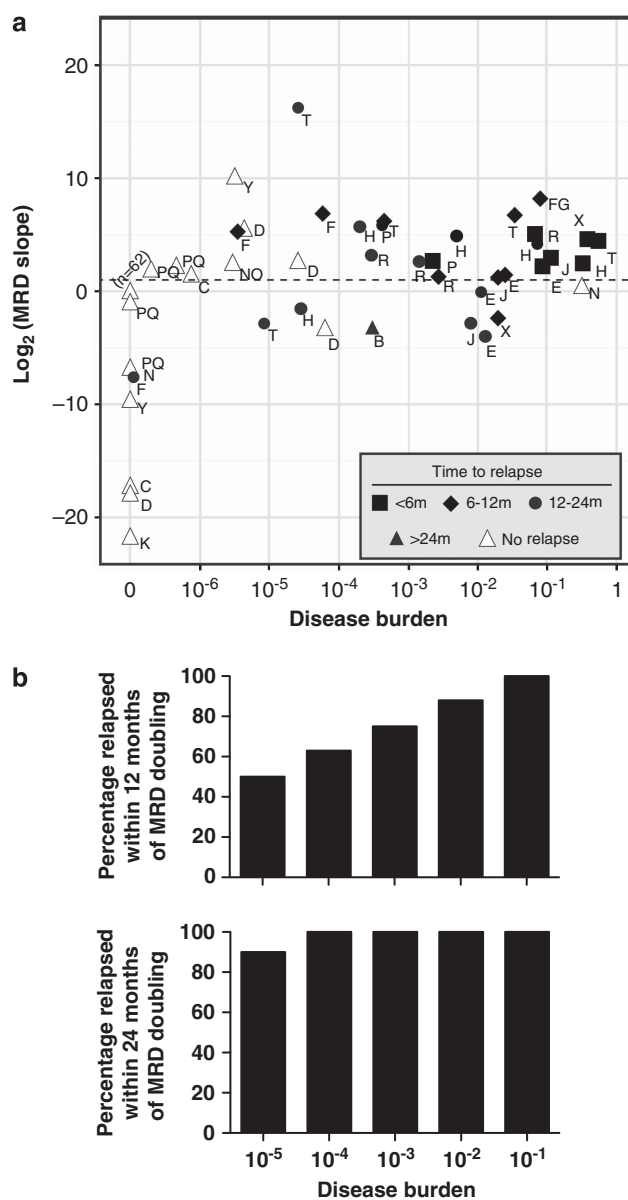
Association of other variables with relapse

Other pre- and post-transplant factors were evaluated to determine whether any were significantly associated with relapse, with the possible implication that a multivariate model could be developed to further improve relapse prediction by IGH–HTS MRD analysis. In univariate analysis, pre-transplant variables including fludarabine refractoriness, 17p deletions, 11q deletions, IGH mutational status and donor relatedness were not significantly associated with outcome. Likewise, post-transplant variables and outcomes, such as use of post-transplant rituximab therapy, CD3/CD19-mixed chimerism 1 month after transplant and CD3-mixed chimerism 12 months after transplant were not significantly associated with relapse. Although persistent mixed CD19 chimerism at 12 months correlated with relapse ($P = 0.002$), this is a surrogate measure of disease burden with a detection limit of 10^{-2} that does not improve relapse prediction beyond what is provided by IGH–HTS quantification.

DISCUSSION

We describe a method for MRD quantification that utilizes consensus primers and HTS to universally amplify and sequence all rearranged IGH genes present in a patient sample without the need for development of patient-specific reagents. The LymphoSIGHT HTS-based MRD-quantification approach employs a series of three multiplexed PCR amplifications to recover the repertoire of IGH molecules in a clinical specimen.[13] IGH amplicons are then sequenced individually, allowing each IGH clonotype to be quantified (Figure 1). Once a patient's disease-associated clonotype is determined—for instance, by sequencing a sample with high disease burden—quantifying the prevalence of that specific sequence within the repertoire of IGH genes observed in follow-up samples becomes primarily a bioinformatic endeavor.

Unique IGH gene rearrangements may be distinguished from theoretical germline VDJ segment alignments by the presence of N-region insertions and deletions generated during VDJ recombination or by somatic hypermutations that are generated in

Figure 5. Patterns of molecular disease progression associated with rapid clinical relapse. (**a**) To account for disease burden and rate of MRD change, disease burden is graphed against the slope ($\log_2$ used to improve visibility) of change at the time of sample collection. The patient from whom each sample was acquired is shown with a letter code (see Supplementary Table 1). Time intervals from sample acquisition to clinical relapse are depicted by graph symbol with intervals < 6 months (squares), 6–12 months (diamonds), 12–24 months (circles) and > 24 months (filled triangles) depicted. Samples from patients who maintained clinical remission are depicted by open triangles. The dashed line depicts an MRD progression slope of 2 ($\log_2 = 1$). (**b**) The percentage of patients who relapsed within 12 and 24 months of MRD doubling is shown for each level of disease burden.

germinal center reactions. Although the majority of CLL patients in this study possessed disease clonotypes without somatic hypermutations (as is typical for CLL patients proceeding to transplant), all disease-associated clonotypes exhibited N-region changes that distinguish them from other B cells using the same germline segments. Consequently, the disease-associated clonotype is unlikely to be confused with other B cells, and the results do not require interpretation as to the relationship of a clonotype quantification and actual disease burden.

We first evaluated several performance characteristics of the LymphoSIGHT IGH–HTS method to assess quantification reproducibility in somatically unmutated CLL. Inter-experimental reproducibility was assessed by processing a subset of patient samples from the primary DNA sample to end-sequence data in two temporally separated assays and found high concordance across the spectrum of assay sensitivity ($r = 0.99$; Figure 2c). By assessing quantification of both IGH alleles in biallelically rearranged CLL, we demonstrated high intra-experimental reproducibility as well ($r = 0.97$; Figure 2d). In comparison with an established method for molecular disease burden quantification—ASO-PCR—we found CLL clonotype quantifications correlated well between the two methods ($r = 0.64$; Figure 3).

A number of limitations may affect quantification using the IGH sequencing assay. First, although the assay has been optimized to minimize this effect, there may be some sequences that amplify preferentially. Moreover, for diseases other than unmutated CLL, mutations may interfere with amplification in one or more of the frames, which may lead to absolute quantification error. Such errors, though, would affect each sample from an individual patient, and apparent changes in disease burden between samples should not arise solely from such amplification biases. Finally, to achieve absolute disease burden quantification, the assay requires measurement of two parameters: the total DNA amount and the number of reference IGH sequence reads that is used to determine the number of CLL-associated and non-CLL IGH sequences. The precision of the assay, including the effects of minute pipetting errors (a reality affecting all molecular quantification approaches), was demonstrated by low average relative standard deviations at clonotype frequencies $\geqslant 3 \times 10^{-5}$.[13] Future studies will be focused on further optimization of these assay parameters.

Understanding the technical performance characteristics and areas for future improvement, we then applied the LymphoSIGHT MRD-quantification approach[13] to 40 patients who underwent reduced-intensity allo-HCT for high-risk CLL. Sufficient cells were generally present in our post-transplant samples derived from 10-ml blood draws to permit detection of residual disease at the 10^{-6} level, which previously has been attained only with allele-specific nested PCR,[25] with a quantitative range above 10^{-5}.

In a 12-month post-HCT landmark analysis, detection of MRD $\geqslant 10^{-6}$ using LymphoSIGHT IGH–HTS was strongly associated with subsequent relapse (relapse HR 7.9; 95% CI 2.3–26; $P = 0.0002$) (Figure 4). Our analysis appears to have similar predictive value as those published by other groups using ASO-PCR or multiparameter flow cytometry to quantify CLL MRD (reviewed by Dreger *et al.*,[26] Varghese *et al.*,[27] and Böttcher *et al.*[6]). For instance, the German CLL Study Group reported that MRD negativity ($< 10^{-4}$) 12 months following allo-HCT strongly correlated with decreased risk of relapse (HR 0.037; 95% CI 0.008–0.18; $P < 0.00001$) in 59 patients, amongst whom 12 (20%) relapsed prior to 12 months, 2 of the 32 (6.3%) MRD-negative patients relapsed and 6 of the 11 (55%) MRD-positive patients relapsed with a median follow-up over 36 months.[6,28] Farina *et al.*[29] reported a 24-month DFS of 100% in patients negative for MRD by a PCR-based assay 12 months following reduced-intensity allo-HCT, whereas 57% of those with detectable molecular disease relapsed with a median follow-up of 26 months ($P = 0.037$). We also found the presence of $\geqslant 10^{-6}$ MRD at 9 months (HR 5.4; 95% CI 1.4–21; $P = 0.01$), 18 months (HR 22; 95% CI 4.8–99; $P < 0.0001$) and 24 months (HR 49; 95% CI 8.8–272; $P < 0.0001$) post-HCT convey increasingly significant risk for relapse.

The availability of generally applicable MRD quantification by this standardized IGH–HTS method may facilitate routine serial measurement of disease burden. This, in turn, necessitates algorithms for predicting time-to-relapse based on changes in MRD burden over time. To improve relapse prediction beyond a simple binary assessment of MRD at a specific time point, we evaluated serial measurements to capture trajectories of molecular disease burden. We found that 90% of patients in our cohort who have relapsed experienced an MRD doubling time of $\leqslant 12$ months. When we combined the analysis of disease burden with slope of MRD progression, we improved our ability to predict relapse within 12–24 months of sampling (Figure 5b). As samples with very low disease burdens are rare, the significance of MRD quantification below 10^{-5} will require validation in larger, multiple-institution patient cohorts.

For most patients in our cohort, samples were acquired roughly every 6 months following the first post-transplant year, which in some cases may have limited our ability to capture MRD doubling in a timely manner prior to clinical relapse. Although our current analytic methodology identifies patients with a substantial risk of relapse within 12 (50% likelihood) or 24 months (90% likelihood) when disease burden is $\geqslant 10^{-5}$ and doubling, we anticipate that increased sampling frequency after patients have demonstrated MRD doubling may permit prediction of time-to-relapse with higher precision.

We believe that sufficient evidence is mounting from the aggregate body of MRD studies to support the development of clinical trials testing therapeutic interventions based on progression of molecular disease burden in CLL.[26,27,30–36] Pre-emptive therapeutic maneuvers triggered by MRD progression will likely need to be applied in a manner that accounts for the specific risk of individual therapies. For instance, a 60–70% likelihood of relapse may be sufficient justification to initiate therapy with a novel antibody therapy, or a molecular inhibitor of BCR signaling,[37] either of which likely convey limited risk, as such therapies can be discontinued if undue toxicity is observed, or if therapy is shown to be ineffective at reducing MRD by further IGH–HTS quantification while on therapy. In contrast, a higher degree of certainty in relapse prediction will be desirable prior to treating pre-emptively with unmanipulated donor lymphocyte infusion, which in our experience is associated with subsequent GVHD in more than half of patients, and thus this approach conveys a significantly higher risk, which is not easily reversed.

In 2008, the IW-CLL published recommendations to include MRD studies whenever feasible in CLL clinical trials.[3] The National Cancer Institute's First International Workshop on the Biology, Prevention, and Treatment of Relapse after Allogeneic Hematopoietic Stem Cell Transplantation also cited a need to develop methods for MRD quantification amenable to standardization across institutions.[38] Important efforts have been made to standardize multiparameter flow cytometry methods for MRD quantification in CLL.[5] The sensitive, specific and objective characteristics of the LymphoSIGHT IGH–HTS method, paired with acceptable scalability features, are the significant benefits of this genetic MRD-quantification approach. Comparative studies of state-of-the-art standardized multiparameter flow cytometry and IGH–HTS methods will be needed to determine the relative merits of the two approaches. The development of our IGH–HTS method and its clinical validation with this patient cohort is an important step toward such studies.

AUTHORSHIP

REFERENCES

1 Gribben JG, O'Brien S. Update on therapy of chronic lymphocytic leukemia. *J Clin Oncol* 2011; **29**: 544–550.

2 Zenz T, Gribben JG, Hallek M, Dohner H, Keating MJ, Stilgenbauer S. Risk categories and refractory CLL in the era of chemoimmunotherapy. *Blood* 2012; **119**: 4101–4107.

3 Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner H et al. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood* 2008; **111**: 5446–5456.

4 Rawstron AC, Villamor N, Ritgen M, Bottcher S, Ghia P, Zehnder JL et al. International standardized approach for flow cytometric residual disease monitoring in chronic lymphocytic leukaemia. *Leukemia* 2007; **21**: 956–964.

5 Rawstron AC, Bottcher S, Letestu R, Villamor N, Fazi C, Kartsios H et al. Improving efficiency and sensitivity: European Research Initiative in CLL (ERIC) update on the international harmonised approach for flow cytometric residual disease monitoring in CLL. *Leukemia* 2013; **27**: 142–149.

6 Böttcher S, Ritgen M, Dreger P. Allogeneic stem cell transplantation for chronic lymphocytic leukemia: lessons to be learned from minimal residual disease studies. *Blood Rev* 2011; **25**: 91–96.

7 Böttcher S, Ritgen M, Fischer K, Stilgenbauer S, Busch RM, Fingerle-Rowson G et al. Minimal residual disease quantification is an independent predictor of progression-free and overall survival in chronic lymphocytic leukemia: a multivariate analysis from the randomized GCLLSG CLL8 trial. *J Clin Oncol* 2012; **30**: 980–988.

8 Böttcher S, Ritgen M, Pott C, Bruggemann M, Raff T, Stilgenbauer S et al. Comparative analysis of minimal residual disease detection using four-color flow cytometry, consensus IgH-PCR, and quantitative IgH PCR in CLL after allogeneic and autologous stem cell transplantation. *Leukemia* 2004; **18**: 1637–1645.

9 Böttcher S, Stilgenbauer S, Busch R, Bruggemann M, Raff T, Pott C et al. Standardized MRD flow and ASO IGH RQ-PCR for MRD quantification in CLL patients after rituximab-containing immunochemotherapy: a comparative analysis. *Leukemia* 2009; **23**: 2007–2017.

10 Boyd SD, Marshall EL, Merker JD, Maniar JM, Zhang LN, Sahaf B et al. Measurement and clinical monitoring of human lymphocyte clonality by massively parallel VDJ pyrosequencing. *Sci Transl Med* 2009; **1**: 12ra23.

11 Logan AC, Gao H, Wang C, Sahaf B, Jones CD, Marshall EL et al. High-throughput VDJ sequencing for quantification of minimal residual disease in chronic lymphocytic leukemia and immune reconstitution assessment. *Proc Natl Acad Sci* 2011; **108**: 21194–21199.

12 van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, Lavender FL 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**: 2257–2317.

13 Faham M, Zheng J, Moorhead M, Carlton VE, Stow P, Coustan-Smith E et al. Deep-sequencing approach for minimal residual disease detection in acute lymphoblastic leukemia. *Blood* 2012; **120**: 5173–5180.

14 Lowsky R, Takahashi T, Liu YP, Dejbakhsh-Jones S, Grumet FC, Shizuru JA et al. Protective conditioning for acute graft-versus-host disease. *New Engl J Med* 2005; **353**: 1321–1331.

15 Kohrt HE, Turnbull BB, Heydari K, Shizuru JA, Laport GG, Miklos DB et al. TLI and ATG conditioning with low risk of graft-versus-host disease retains antitumor reactions after allogeneic hematopoietic cell transplantation from related and unrelated donors. *Blood* 2009; **114**: 1099–1109.

16 Arai S, Sahaf B, Narasimhan B, Chen GL, Jones CD, Lowsky R et al. Prophylactic rituximab after allogeneic transplantation decreases B-cell alloimmunity with low chronic GVHD incidence. *Blood* 2012; **119**: 6145–6154.

17 Gawad C, Pepin F, Carlton VE, Klinger M, Logan AC, Miklos DB et al. Massive evolution of the immunoglobulin heavy chain locus in children with B precursor acute lymphoblastic leukemia. *Blood* 2012; **120**: 4407–4417.

18 Ladetto M, Donovan JW, Harig S, Trojan A, Poor C, Schlossnan R et al. Real-time polymerase chain reaction of immunoglobulin rearrangements for quantitative evaluation of minimal residual disease in multiple myeloma. *Biol Blood Marrow Transplant* 2000; **6**: 241–253.

19 Ranheim EA, Jones CD, Zehnder JL. Sensitive detection of clonal immunoglobulin rearrangements in frozen and paraffin embedded tissues by polymerase chain reaction heteroduplex analysis. *Diagn Mol Pathol* 2000; **9**: 177–183.

20 Przepiorka D, Weisdorf D, Martin P, Klingemann HG, Beatty P, Hows J et al. 1994 consensus conference on acute GVHD grading. *Bone marrow transplant* 1995; **15**: 825–828.

21 Filipovich AH, Weisdorf D, Pavletic S, Socie G, Wingard JR, Lee SJ et al. National Institutes of Health consensus development project on criteria for clinical trials in chronic graft-versus-host disease: I. Diagnosis and staging working group report. *Biol Blood Marrow Transplant* 2005; **11**: 945–956.

22 Kaplan EI, Meier P. Nonparametric estimation from incomplete observation. *J Am Statist Assoc* 1958; **53**: 457–481.

23 Team RDC. R: a language and environment for statistical computing. In *R Foundation for Statistical Computing* Vienna, Austria, 2010.

24 Molica S, Alberti A. Prognostic value of the lymphocyte doubling time in chronic lymphocytic leukemia. *Cancer* 1987; **60**: 2712–2716.

25 Voena C, Ladetto M, Astolfi M, Provan D, Gribben JG, Boccadoro M et al. A novel nested-PCR strategy for the detection of rearranged immunoglobulin heavy-chain genes in B cell tumors. *Leukemia* 1997; **11**: 1793–1798.

26 Dreger P, Ritgen M, Bottcher S, Schmitz N, Kneba M. The prognostic impact of minimal residual disease assessment after stem cell transplantation for chronic lymphocytic leukemia: is achievement of molecular remission worthwhile? *Leukemia* 2005; **19**: 1135–1138.

27 Varghese AM, Rawstron AC, Hillmen P. Eradicating minimal residual disease in chronic lymphocytic leukemia: should this be the goal of treatment? *Curr Hematol Malig Rep* 2010; **5**: 35–44.

28 Dreger P, Dohner H, Ritgen M, Bottcher S, Busch R, Dietrich S et al. Allogeneic stem cell transplantation provides durable disease control in poor-risk chronic lymphocytic leukemia: long-term clinical and MRD results of the German CLL Study Group CLL3X trial. *Blood* 2010; **116**: 2438–2447.

29 Farina L, Carniti C, Dodero A, Vendramin A, Raganato A, Spina F et al. Qualitative and quantitative polymerase chain reaction monitoring of minimal residual disease in relapsed chronic lymphocytic leukemia: early assessment can predict long-term outcome after reduced intensity allogeneic transplantation. *Haematologica* 2009; **94**: 654–662.

30 Rawstron AC, Kennedy B, Moreton P, Dickinson AJ, Cullen MJ, Richards SJ et al. Early prediction of outcome and response to alemtuzumab therapy in chronic lymphocytic leukemia. *Blood* 2004; **103**: 2027–2031.

31 Ritgen M, Bottcher S, Stilgenbauer S, Bunjes D, Schubert J, Cohen S et al. Quantitative MRD monitoring identifies distinct GVL response patterns after allogeneic stem cell transplantation for chronic lymphocytic leukemia: results from the GCLLSG CLL3X trial. *Leukemia* 2008; **22**: 1377–1386.

32 Nabhan C, Coutre S, Hillmen P. Minimal residual disease in chronic lymphocytic leukaemia: is it ready for primetime? *Br J Haematol* 2007; **136**: 379–392.

33 Moreton P, Kennedy B, Lucas G, Leach M, Rassam SM, Haynes A et al. Eradication of minimal residual disease in B-cell chronic lymphocytic leukemia after alemtuzumab therapy is associated with prolonged survival. *J Clin Oncol* 2005; **23**: 2971–2979.

34 Sorror ML, Maris MB, Sandmaier BM, Storer BE, Stuart MJ, Hegenbart U et al. Hematopoietic cell transplantation after nonmyeloablative conditioning for advanced chronic lymphocytic leukemia. *J Clin Oncol* 2005; **23**: 3819–3829.

35 Schweighofer CD, Ritgen M, Eichhorst BF, Busch R, Abenhardt W, Kneba M et al. Consolidation with alemtuzumab improves progression-free survival in patients with chronic lymphocytic leukaemia (CLL) in first remission: long-term follow-up of a randomized phase III trial of the German CLL Study Group (GCLLSG). *Br J Haematol* 2009; **144**: 95–98.

36 Moreno C, Villamor N, Colomer D, Esteve J, Gine E, Muntanola A et al. Clinical significance of minimal residual disease, as assessed by different techniques, after stem cell transplantation for chronic lymphocytic leukemia. *Blood* 2006; **107**: 4563–4569.

37 Wiestner A. Emerging role of kinase-targeted strategies in chronic lymphocytic leukemia. *Hematology Am Soc Hematol Educ Program* 2012; **2012**: 88–96.

38 Kroger N, Bacher U, Bader P, Bottcher S, Borowitz MJ, Dreger P et al. NCI first international workshop on the biology, prevention, and treatment of relapse after allogeneic hematopoietic stem cell transplantation: report from the committee on disease-specific methods and strategies for monitoring relapse following allogeneic stem cell transplantation. part II: chronic leukemias, myeloproliferative neoplasms, and lymphoid malignancies. *Biol Blood Marrow Transplant* 2010; **16**: 1325–1346.

Supplementary Information accompanies this paper on the Leukemia website (http://www.nature.com/leu)