2018 JMD CME Program

The Journal of Molecular Diagnostics, Vol. 20, No. 2, March 2018

ELSEVIER

the Journal of Molecular Diagnostics

jmd.amjpathol.org

See related Commentary on page 163

Assessment of Capture and Amplicon-Based Approaches for the Development of a Targeted Next-Generation Sequencing Pipeline to Personalize Lymphoma Management

Stacy S. Hung,* Barbara Meissner,* Elizabeth A. Chavez,* Susana Ben-Neriah,* Daisuke Ennishi,* Martin R. Jones,† Hennady P. Shulha,* Fong Chun Chan,* Merrill Boyle,* Robert Kridel,* Randy D. Gascoyne,*‡ Andrew J. Mungall,† Marco A. Marra,† David W. Scott,* Joseph M. Connors,* and Christian Steidl*‡

From the Centre for Lymphoid Cancer and the Michael Smith Genome Sciences Centre,† BC Cancer Agency, Vancouver; and the Department of Pathology and Laboratory Medicine,‡ University of British Columbia, Vancouver, British Columbia, Canada*

CME Accreditation Statement: This activity ("JMD 2018 CME Program in Molecular Diagnostics") has been planned and implemented in accordance with the accreditation requirements and policies of the Accreditation Council for Continuing Medical Education (ACCME) through the joint providership of the American Society for Clinical Pathology (ASCP) and the American Society for Investigative Pathology (ASIP). ASCP is accredited by the ACCME to provide continuing medical education for physicians.

The ASCP designates this journal-based CME activity ("JMD 2018 CME Program in Molecular Diagnostics") for a maximum of 18.0 AMA PRA Category 1 Credit(s)™. Physicians should claim only credit commensurate with the extent of their participation in the activity.

CME Disclosures: The authors of this article and the planning committee members and staff have no relevant financial relationships with commercial interests to disclose.

Accepted for publication November 3, 2017.

Address correspondence to Christian Steidl, M.D., BC Cancer Research Centre, 675 W 10th Ave, Vancouver, BC V5Z 1L3, Canada. E-mail: csteidl@bccancer.bc.ca.

Targeted next-generation sequencing panels are increasingly used to assess the value of gene mutations for clinical diagnostic purposes. For assay development, amplicon-based methods have been preferentially used on the basis of short preparation time and small DNA input amounts. However, capture sequencing has emerged as an alternative approach because of high testing accuracy. We compared capture hybridization and amplicon sequencing approaches using fresh-frozen and formalin-fixed, paraffin-embedded tumor samples from eight lymphoma patients. Next, we developed a targeted sequencing pipeline using a 32-gene panel for accurate detection of actionable mutations in formalin-fixed, paraffin-embedded tumor samples of the most common lymphocytic malignancies: chronic lymphocytic leukemia, diffuse large B-cell lymphoma, and follicular lymphoma. We show that hybrid capture is superior to amplicon sequencing by providing deep more uniform coverage and yielding higher sensitivity for variant calling. Sanger sequencing of 588 variants identified specificity limits of thresholds for mutation calling, and orthogonal validation on 66 cases indicated 93% concordance with whole-genome sequencing. The developed pipeline and assay identified at least one actionable mutation in 91% of tumors from 219 lymphoma patients and revealed subtype-specific mutation patterns and frequencies consistent with the literature. This pipeline is an accurate and sensitive method for identifying actionable gene mutations in routinely acquired biopsy materials, suggesting further assessment of capture-based assays in the context of personalized lymphoma management. *(J Mol Diagn 2018, 20: 203—214; https://doi.org/10.1016/j.jmoldx.2017.11.010)*

Supported by Genome Canada grant 4108 (J.M.C.), Genome British Columbia grant 141LYM (J.M.C.), the Canadian Institutes of Health Research grant 148393 (J.M.C.), and the Terry Fox Research Institute grant 1061 (C.S.). C.S. is the recipient of a Michael Smith Foundation for Health Research Career Investigator award and a Canadian Institutes of Health Research New Investigator award.

Disclosures: None declared.

Current address of F.C.C., Achilles Therapeutics, London, United Kingdom; of R.K., Princess Margaret Cancer Center, Toronto, Ontario, Canada.

https://doi.org/10.1016/j.jmoldx.2017.11.010

In aggregate, lymphoid cancers are the fourth most common cancer in Canada and the only common cancer with increasing incidence—a robust but unexplained trend that has been documented for >50 years. Lymphoid cancers disproportionately affect younger patients, exaggerating their impact in terms of productive years of life lost. Moreover, lymphoid cancers are the only common cancers that can regularly be cured in a subset of patients, even when widely disseminated, indicating the potential to improve outcomes if currently available treatments are optimized and strategically applied.

Modern diagnosis and molecular assessment of lymphoid cancers currently relies on techniques such as immunohistochemistry, cytogenetics, fluorescence *in situ* hybridization, flow cytometry, and molecular genetics to classify lymphoid neoplasms into >35 distinct entities, each with distinct treatment implications.[1] However, in recent years, these standard procedures and classification schemes have been challenged by the emergence of sequencing techniques that have the potential to add unique diagnostic, prognostic, and predictive value to standard assessment.[2,3]

Recent examples of clinically applicable sequencing-based assays include the development of the M7-FLIPI in follicular lymphoma (FL), clinical trials exploring potential differential efficacy of lenalidomide or ibrutinib to standard chemotherapy for diffuse large B-cell lymphoma (DLBCL) patients with ABC-subtype−specific mutations, and (sub) clonal *TP53* mutations and ibrutinib-resistance associated mutations in chronic lymphocytic leukemia (CLL).[4−6] Thus, genomic profiling has the potential to deliver clinically relevant information that is otherwise missed using current testing. In addition, genomic profiling will generate a rich platform that can be exploited for extensive additional discoveries.

Next-generation sequencing technologies have been instrumental in accelerating discovery in cancer genomics via whole-genome sequencing (WGS), whole-exome sequencing, whole-transcriptome sequencing (RNA sequencing), and deep targeted sequencing. These technologies have been extensively used in disease-specific contexts to identify somatic mutations, understand clonal evolution, and, most recently, advance personalized medicine.[7,8] In contrast to genome-wide applications, targeted cancer sequencing panels, which focus on a select set of genes or gene regions that have known associations with cancer, allow for the rapid detection of a variety of somatic mutations on a single platform.[9−11]

Two methods are commonly used for such targeted approaches: capture hybridization−based sequencing and amplicon-based sequencing, each having its own advantages and disadvantages. A recent study that compared these two types of methods head-to-head indicates that amplicon-based approaches may be preferable for their simplified workflow and smaller amounts of required DNA.[12] However, hybridization-based strategies are less likely to miss mutations and also perform better with respect to sequencing complexity and uniformity of coverage.[12−14]

Herein, to determine the ideal targeted sequencing platform in the context of detecting actionable mutations in lymphoid cancer, we performed a systematic comparison of capture hybridization against amplicon sequencing with the aim of accurately detecting the full spectrum of mutations in the lymphoid cancer gene panel. Specifically, we describe the development and application of a targeted sequencing assay to identify mutations in routinely obtained formalin-fixed, paraffin-embedded (FFPE) lymph node samples (FL and DLBCL) and enriched B lymphocytes of CLL specimens. Our findings demonstrate the feasibility and outline the clinical utility of integrating a lymphoma-specific pipeline into personalized cancer care.

Materials and Methods

Patients and Materials

A cohort of matched tumor and normal DNA specimens was assembled from 229 patients (Supplemental Figure S1), including 30 CLL, 80 DLBCL, and 119 FL cases. FFPE tissue blocks and peripheral blood DNA samples for all FL and DLBCL cases were acquired from the BC Cancer Agency (Vancouver, BC, Canada) lymphoma tumor bank. Tumor DNA was extracted from 1 × 10 μm FFPE sections using the AllPrep DNA/RNA FFPE Kit (Qiagen, Hilden, Germany) or from 15 × 20 μm snap frozen tissue (FF) sections using the AllPrep DNA/RNA Mini Kit (Qiagen). Peripheral blood DNA was extracted using the FlexiGene DNA Kit (Qiagen). Genomic DNA samples were quality checked on agarose gels and quantified using Qubit dsDNA BR Assay (Thermo Fisher Scientific Inc., Waltham, MA).

For CLL, fresh peripheral blood samples were collected and processed within 24 hours to obtain pure tumor (>98% purity) and germline cell fractions by separation methods [CD19-negative RosetteSep (tumor) and CD19 RosetteSep (germline); StemCell Technologies, Vancouver, BC, Canada]. The germline cell fraction was further purified by fluorescence-activated cell sorting [negative sort using CD19-PECy5 (BioLegend, San Diego, CA), CD20-PECy7 (Beckman Coulter, Brea, CA), and DAPI (Sigma, St. Louis, MO)] on a BD ARIA cell sorter (BD Biosciences, San Jose, CA; >97% purity). Purified cell fractions were extracted using the Qiagen AllPrep DNA/RNA Mini Kit standard protocol.

For WGS, specimens for 66 of 229 patients from the original cohort were frozen and embedded in OCT compound for DNA and RNA extractions, as well as frozen sections for histological correlation. Constitutional DNA was extracted from peripheral blood leukocytes. FF surgical tissue resections were divided into sections (50 μm thick), and four sections were added to each tube of 400 μL RLT Plus buffer (Qiagen) containing tris (2-carboxyethyl) phosphine. Simultaneous purification of genomic DNA and total RNA from 3 to 11 tubes, selected for tumor content and cellularity, was performed using the AllPrep DNA/RNA Mini Kit on the automated robotic workstation, QIAcube (Qiagen).

All patient specimens were collected as part of research projects approved by the University of British Columbia–British Columbia Cancer Agency Research Ethics Board (file H05-60103).

Capture Sequencing

Using the Agilent SureDesign software version 5.0.0.112 (Agilent, Santa Clara, CA), a total of 614 regions from the 32 genes were targeted for a final capture size of 263 kb. Adaptor-ligated libraries were constructed following the 100 ng SureSelect XT2 Target Enrichment System for Illumina Multiplexed Sequencing protocol version D.2 (Illumina, San Diego, CA). Indexed libraries from eight tumor-normal pairs were pooled and sequenced in batches on the Illumina MiSeq using 150-bp paired-end reads. Tumor and normal DNA samples were sequenced to an average of 301× coverage (SD, 212×); samples with a mean of <50× were considered failures and were excluded from the analysis. All reads were aligned to the human reference genome (hg19) using bwa-mem version 0.7.5a,[15] and variants were identified using the intersection of calls predicted by VarScan version 2.3.6[16] and Strelka version 1.0.13.[17] Variants were manually reviewed in the Integrated Genome Viewer version 2.3.25.[18]

Amplicon Sequencing

A total of 20 genes were targeted for deep sequencing with an Illumina TruSeq Custom Amplicon (TSCA) assay using a custom mixture of oligonucleotides that generate 474 amplicons covering 39 kb. Samples were prepared following the TSCA protocol using 250 ng of template DNA per reaction. Multiple indexed libraries were pooled and sequenced on the Illumina MiSeq using 150-bp paired-end reads. Tumor and normal DNA samples were sequenced to an average of 1899× coverage (SD, 2628×). Alignment and variant detection was performed using Mutascope version 1.0.0[19] using the human reference genome (hg19). Variants were manually reviewed in Integrated Genome Viewer version 2.3.25.[18]

Sanger Sequencing

Primers were designed using Primer3 version 0.4.0 (*http://bioinfo.ut.ee/primer3-0.4.0*)[20] to amplify a 200- to 280-bp region flanking the variant. Primer sequences are listed in Supplemental Table S1. After amplification, DNA was purified using ExoSAP-IT PCR Product Cleanup Reagent (Thermo Fisher Scientific Inc.) and submitted for conventional Sanger sequencing (GENEWIZ, South Plainfield, NJ).

Whole-Genome Library Construction and Sequencing

To minimize genome library bias and coverage gaps associated with PCR amplification of high GC- or AT-rich regions, an automated version of an in-house–developed PCR-free genome protocol was implemented. Briefly, high-molecular-weight genomic DNA (1 μg) from tumor or blood samples was subjected to shearing to 300 to 600 bp by sonication for 30 seconds (Covaris, Woburn, MA). Sheared DNA was end repaired and size selected using AMPure XP beads targeting a 300- to 400-bp fraction. After 3′ A tailing, full-length TruSeq adapters were ligated. Libraries were purified using AMPure XP beads, and fragment sizes were assessed using an aliquot of PCR-amplified library DNA on an Agilent 2100 Bioanalyzer DNA1000 chip (Agilent) or Caliper GX DNA1000 chip (Caliper, Newton, MA). The PCR-free library concentrations were quantified using a qPCR Library Quantification kit (KK4824; KAPA; Kapa Biosystems, Wilmington, MA). Genome libraries were sequenced on HiSeq2500 or HiSeqX instruments (Illumina) using V3 or V4 chemistry and paired-end 125 base reads.

Somatic Alteration Assessment and Filtering

For the analysis of capture sequencing data, raw sequence read data generated from the Illumina MiSeq instrument were demultiplexed and preprocessed using the MiSeq Control Software version 2.6.2.1 to produce sample-specific FASTQ files. Sequencing adaptors and target-specific primers were trimmed. Paired-end alignments were performed using the bwa-mem aligner version 0.7.5a[15] against the human genome (hg19). Resulting primary BAM files were treated with Picard version 1.126 (Broad Institute, Cambridge, MA; *https://broadinstitute.github.io/picard*) to remove duplicates and SAMtools version 0.1.18[21] for merging, sorting, and indexing mature BAM files. These files were then used to call variants using VarScan version 2.3.6[16] and Strelka version 1.0.13.[17] Database annotation was done by SnpEff version 4.2[22] and further filtered and annotated against dbSNP version 137.[23] The following criteria were defined for the inclusion of variants: i) detected by VarScan and Strelka [or just VarScan for insertions/deletions (indels)], meets or exceeds an allele frequency (AF) of 10% with ≥10 variant reads, and has no strand bias; ii) detected by VarScan and Strelka, falls below an AF of 10% (and/or has <10 variant reads), and is validated by Sanger sequencing; or iii) detected by VarScan only (not in Strelka), meets or exceeds an AF of 10% with ≥10 variant reads, and has been visually inspected to have strong evidence as a real mutation.

For the analysis of TSCA data, the Mutascope version 1.0.0 pipeline was applied for alignment and mutation calls. Final calls were filtered for somatic mutations (on the basis of the Fisher exact test; SS = 2, variant is absent in the normal; GEN[0].GT = 0/0, variant is heterozygous or homozygous in the tumor; GEN[1].GT = 0/1 or 1/1) and containing a PASS (14 filters used by Mutascope based on normal comparison to remove systematic biases attributable to amplicon sequencing). To overcome known PCR-related artifacts associated with amplicon sequencing, additional

Table 1 Comparison of Mutations Determined by Capture or Amplicon Sequencing for Fresh-Frozen and FFPE Samples across 8 Lymphoma Patients

Chromosome	Position	Reference allele	Alternative allele	Gene	Capture		TSCA		Average AF, %
					FF	FFPE	FF	FFPE	
13	41239894	*G*	*C*	*FOXO1*	8.8	**11.8**			10.3
13	41240294	*C*	*T*	*FOXO1*	**11.0**	9.9			10.4
17	63049847	*C*	*T*	*GNA13*	**10.2**	7.8	**14.9**		10.9
1	2491268	*G*	*T*	*TNFRSF14*	2.5	**20.3**			11.4
13	41239891	*G*	*C*	*FOXO1*		**12.4**			12.4
17	7578554	*A*	*T*	*TP53*	**22.6**	**13.4**	**9.6**	**16.4**	15.5
7	148506467	*G*	*C*	*EZH2*	**20.0**	**20.0**	**11.3**	**18.4**	17.4
1	2489781	*G*	*A*	*TNFRSF14*	**18.9**	**23.6**	**19.4**	**28.1**	22.5
17	63052656	*A*	*G*	*GNA13*	**34.2**	**13.1**			23.6
16	3786797	*A*	*G*	*CREBBP*	**26.5**	**29.9**	**26.6**	**24.0**	26.8
3	183217420	*C*	*G*	*KLHL6*	**36.5**	**23.4**	**32.6**	**32.9**	31.3
1	120458264	*G*	*A*	*NOTCH2*	**37.3**	**25.0**	**35.3**	**30.2**	31.9
16	3788617	*C*	*T*	*CREBBP*	**33.7**	**34.6**			34.2
1	2488138	*G*	*A*	*TNFRSF14*	**49.2**	**59.1**			54.2
16	3786704	*A*	*G*	*CREBBP*	**66.5**	**58.7**	**64.4**	**65.1**	63.7
16	3786704	*A*	*C*	*CREBBP*	**59.4**	**78.0**	**51.1**	**68.4**	64.2

AF is shown for all technologies and is bold where the mutation met defined calling thresholds (*Materials and Methods*). All mutations have been Sanger validated.

AF, allele frequency; FF, fresh frozen; FFPE, formalin fixed, paraffin embedded; TSCA, TruSeq Custom Amplicon.

constraint was added for mutations to occur in pooled FFPE tissue (ie, present in duplicate samples).

For the comparative analysis of TSCA versus capture data, amplicon data were analyzed as before and capture sequencing reads were aligned using bwa-mem version 0.5.7a and variants called using the intersection of calls made by VarScan version 2.3.6 and Strelka version 1.0.13. VarScan predictions were filtered for somatic mutations with a minimum depth of 20×, having at least 10 variant reads, and an AF $\geq$10%. Default filter parameters were applied for Strelka predictions. For both TSCA and capture methods, in cases in which mutations were associated with multiple annotations, only the highest impact or most relevant mutation was kept, and only mutations with a nonsynonymous, stop gained, or splice site effect were retained. Final mutations were inspected with Integrated Genome Viewer version 2.3.25 and omitted if they were predicted in regions outside of the target space or were located in poorly aligned regions. Single-nucleotide variant calling accuracy was assessed for eight patients in whom constitutional, FF, and FFPE samples could be obtained. FFPE samples were analyzed in replicate pairs for amplicon sequencing. Concordances were compared between capture on FF, capture on FFPE, TSCA on FF, and TSCA on FFPE.

Sequence reads from the whole-genome libraries were aligned to the human reference genome (hg19) using Burrows-Wheeler Alignment tool version 0.5.7.[24] The tumor's genomic sequence was compared with that of patient's constitutive DNA to identify somatic alterations. Single-nucleotide mutations were identified using a probabilistic joint variant calling approach using SAMtools version 0.1.17,[21] MutationSeq version 1.0.2,[25] and Strelka version 1.0.6[17]; small insertions and deletions were identified using Strelka version 1.0.6.[17] Variants were annotated to genes using the Ensembl database version 69.[26]

Results

Hybrid Capture Provides Superior Assay Sensitivity Compared with Amplicon Sequencing

A next-generation sequencing–targeted sequencing assay was developed and optimized for the detection of potentially actionable and clinically relevant mutations in routinely acquired diagnostic lymphoma specimens. An overview of the workflow is shown in Supplemental Figure S2. To establish feasibility of FFPE sequencing and to guide platform selection, the performance of two targeted next-generation sequencing technologies was evaluated using constitutional DNA and matching DNA for both FF and FFPE tissue from eight lymphoma patients. Samples were sequenced using a panel of 20 lymphoma-specific genes via capture hybridization (Sure Select; Agilent) and amplicon sequencing (TruSeq Custom Amplicon; Illumina). In capture hybridization, genomic DNA is sheared to produce small fragments. After library preparation with sequencer-specific adaptors and indexes, the sample is hybridized with biotinylated RNA library baits. Targeted regions are then pulled out using magnetic streptavidin beads, amplified, and loaded onto the sequencer. On the other hand, in amplicon sequencing, custom oligo capture probes are first designed to flank the regions of interest. Custom probes are then hybridized to unfragmented genomic DNA, where extension/ligation takes place between the custom probes and across the

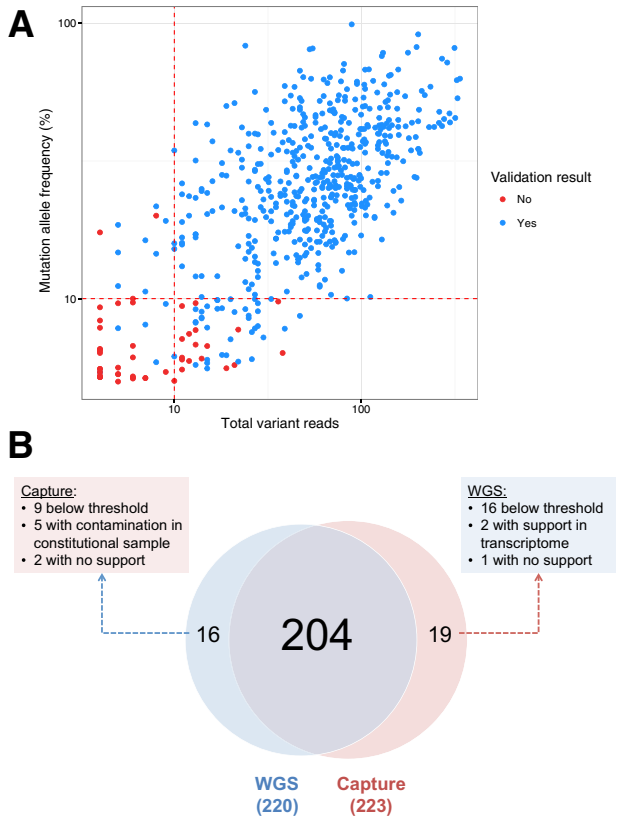

Figure 1 Validation of capture with Sanger and whole-genome sequencing (WGS). **A:** Sanger sequencing validates 100% of somatic mutations called by capture-sequenced lymphoma cases having an allele frequency of ≥10% and with variant read support of at least 10 reads (**dashed lines**). This experiment provided rationale to use a threshold of 10/10 (at least 10 variant reads and 10% allele frequency) for reporting mutations. **B:** Overlap of single-nucleotide variant and insertion/deletion mutation calls between capture and WGS across 66 patients and within the 32-gene panel. $n = 588$ (**A**).

regions of interest. PCR is performed to add indexes and sequencing primers. Finally, the uniquely tagged amplicon library is ready for cluster generation and sequencing. A general comparison of the technologies is provided in Supplemental Table S2. Single-nucleotide variants and indels detected by an in-house bioinformatics pipeline were then compared between the two technologies. Variant detection accuracy was assessed across the possible combinations of technologies and tissue types, herein named as follows: i) capture FF, ii) capture FFPE, iii) TSCA FF, and iv) TSCA FFPE. Only mutations that fell within the target region were included in the concordance analysis using strict variant calling thresholds (amplicon: AF $\geq$ 5% and variant reads $\geq$ 3 in duplicate samples; capture: AF $\geq$ 10% and variant reads $\geq$ 10). A summary of the results is shown in Table 1. A total of 16 variants were detected that were all successfully validated by Sanger sequencing, demonstrating high specificity in all sample type/method combinations. Eight of the 16 variants (50%) were predicted by all four approaches (Supplemental Figure S3). Seven variants (44%) that were predicted by capture in either FF or FFPE tissue samples were not detected in equivalent TSCA samples, demonstrating limited sensitivity of amplicon sequencing even with the use of duplicates. These mutations occurred within amplicon regions that have minimal to zero coverage (dropout amplicons) and represent areas of the genome that were difficult to amplify because of high GC content[27] or were affected by suboptimal PCR amplification conditions.[28] Coverage for the equivalent regions was adequately high in samples sequenced through capture (Supplemental Table S3). The remaining discrepancies between methods and tissue types (Table 1), in particular between FF and FFPE samples, were related to the chosen AF thresholds, and they likely reflect malignant clone size differences between tissue samples that were subjected to the respective sequencing pipelines. In summary, hybrid capture sequencing displayed superior sensitivity compared with amplicon sequencing, while maintaining 100% specificity using Sanger sequencing as the validation method (Figure 1A).

Amplicon sequencing and hybrid capture sequencing methods were also comparatively assessed in FFPE samples of a cohort of 41 lymphoma patients across a target panel of 20 genes (Figure 2A). Coverage from amplicon sequencing was found to be extremely deep (approximately 1900×) but also highly variable (SD, approximately 2600×), with a significant number of dropout regions ($n = 29$). In comparison, capture provided sufficiently deep coverage (approximately 300×) with greater uniformity across the target regions of interest ($P < 0.01$) (Figure 2B and Supplemental Figure S4). Of the dropout amplicon regions observed with amplicon sequencing, 28 of 29 were adequately covered with capture sequencing (Supplemental Table S3). On the basis of the combined evidence of higher sensitivity and more uniform coverage in the comparative analyses, hybrid capture sequencing was selected for subsequent assay development.

Design and Evolution of the 32-Gene Panel

The chosen targeted hybrid capture sequencing assay was further developed by adding target genes of interest (Figure 2A) in accordance with the evolving literature. Specifically, a six-person working group of hematopathologists, clinician-oncologists, and translational scientists (C.S., D.W.S., R.D.G., D.E., R.K., and B.M.) was assembled to curate a comprehensive list of genes from the published literature that are recurrently mutated and could be associated with potential changes in clinical management in CLL, DLBCL, and FL (actionable mutations). Criteria for inclusion of a target gene into the panel were as follows: i) reported outcome correlation using standard of care, ii) improved diagnostic accuracy, and iii) biological rationale for targeted therapy or mutation-associated drug target (Supplemental Table S4).[29–60] The panel, which has passed through several iterations (capture coverage is homogeneous across these versions) (Figure 2B), includes exons of 32 genes spanning a total of 263 kb of genome sequence

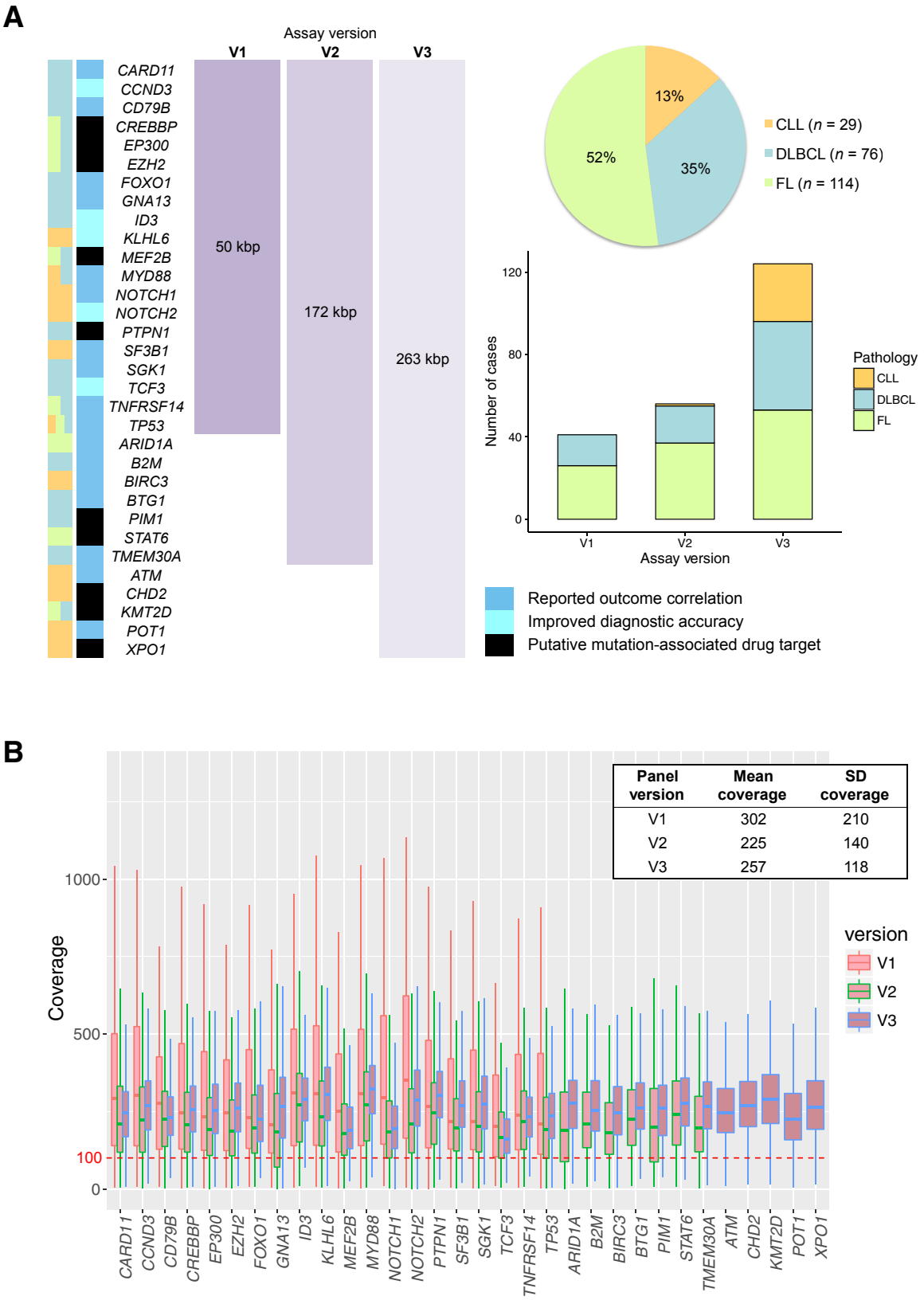

Figure 2 **A:** Design and evolution of the lymphoma-specific 32-gene targeted sequencing panel. **B:** Coverage is homogeneous across the target space and across panel versions. All panel versions have a mean coverage that achieves the desired minimum target coverage of 100× (**dashed line**). Samples having a mean coverage <50× were removed from the analysis. CLL, chronic lymphocytic leukemia; DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; V, version.

representing the latest assay version (version 3). During each iteration, genes were only added, not removed. Assay reports were automatically generated to integrate data from next-generation sequencing with data obtained by other molecular techniques (ie, fluorescence *in situ* hybridization, immunoglobulin sequencing, immunohistochemistry, and gene expression profiling). A summary of the panel design and associated cohort sizes can be found in Figure 2A. The final cohort that was analyzed consisted of 29 CLL, 76 DLBCL, and 114 FL cases (Supplemental Table S5). The target average coverage of the sequencing assay was 200× to detect allele frequencies of 10% supported by 10 reads with high sensitivity; the assay returned an average coverage across the capture space of 250×. To maintain adequate sensitivity for individual samples, a strict threshold of 50× was defined, below which samples were failed and excluded from the analysis. Samples with an average coverage falling between 50× and 100× were repeated to achieve a desired coverage of >100×. The failure analysis indicated that 96% (219/229) of samples analyzed passed the 50× cutoff.

The bioinformatics pipeline and thresholds were also optimized to increase assay accuracy of variant detection and calling. Thresholds for variant calling were adjusted on the basis of Sanger sequencing, and mutations were also validated against a whole-genome data set of 66 patients (7 CLL, 21 DLBCL, and 38 FL cases). Analysis of 588 mutation calls using Sanger sequencing in 169 patients revealed that mutations with an allele frequency of ≥10% and ≥10 supporting variant reads were validated by Sanger sequencing in all cases (100% specificity) (Figure 1A). Within the predefined capture target space, the WGS data yielded 220 single-nucleotide variants/indels, of which capture matched 204 (93% concordance) (Figure 1B). Most mutations found only by WGS (8 of 16) were instances in which variant read support was present, but below our 10% AF/10 supporting variant read threshold. Such missed mutations were typically associated with FFPE samples having poor-quality DNA. For instance, in WGS data, where high-quality DNA samples from GE0285 and GE0361 were sequenced, single indels were detected; however, in capture sequencing, the corresponding samples had degraded DNA and did not contain any supporting reads for the indel mutations. Less commonly, contamination of the normal sample occurred, such as in GE0096, where variant support for three mutations was present in both tumor and constitutional DNA. A complete listing of mutations that were detected in WGS data, but not capture data, is summarized in Supplemental Table S6. Conversely, mutations detected through capture and not WGS were almost exclusively the result of insufficient variant read support in the WGS data set to meet calling thresholds (Supplemental Table S7). In summary, variant calling thresholds were established with 100% specificity that maintained high concordance with a WGS pipeline.

Subtype-Specific Mutational Profiles Are Highly Representative of Published Profiles in FL, DLBCL, and CLL

Finally, the capture sequencing pipeline was applied to call gene mutations (Supplemental Table S8) in FL, DLBCL, and CLL, with which mutation patterns and frequencies were compared with previously published studies (Figure 3 and Supplemental Figures S5 and S6).[31,32,60–65] In total, 114 FL, 76 DLBCL, and 29 CLL cases, which passed the threshold of 50× average coverage, were analyzed. Consistent with previous studies,[32,33,61,62] *KMT2D* was detected as the most commonly mutated gene in both FL (58%) and DLBCL (40%), and other known driver genes, including *PIM1* (28%), *MYD88* (28%), *B2M* (23%), and *TP53* (21%) in DLBCL and *CREBBP* (57%), *TNFRSF14* (43%), and *EZH2* (27%) in FL, were also found to be frequently mutated in this study (Supplemental Figure S5, B and C).[32,62,63] Analyzing mutation types, truncating mutations were predominantly found in known tumor-suppressor genes, such as *KMT2D*, *CREBBP*, *B2M*, *GNA13*, *TP53*, *ARID1A*, and *TNFRSF14*, closely resembling findings in the literature (Figure 3, B and C).[63,64] *MYD88*, *EZH2*, *MYC*, and *BCL2* mutations were found to be exclusively missense mutations in DLBCL, as previously described.[33,65] Trends from the CLL data, including recurrent mutations in *ATM* and *SF3B1* (Figure 3A and Supplemental Figure S5A), are consistent with the previous landscape studies[66]; however, comparisons requiring precise frequencies are limited because of the small cohort size.

In addition, recurrent *NOTCH2* truncating mutations were detected in three FL cases (2.6%) and six DLBCL cases (7.6%). Interestingly, mutations in *NOTCH2* have been reported as the most common lesion in splenic marginal zone lymphoma,[51,67] but gene mutations are not well characterized in other B-cell lymphomas. A recent study has detected three DLBCL cases (4.3%) with a *NOTCH2* nonsense mutation affecting the PEST domain, which leads to cell proliferation and NF-κB signaling activation.[68] All truncating mutations in this study were located in exon 34 in the PEST domain, suggesting its potential pathogenetic role in B-cell lymphomas. Taken together, these results indicate that the sequencing approach with prospectively collected samples can robustly detect a wide range of mutations in lymphoma-related driver genes.

Discussion

This study focusing on the development of a sequencing pipeline to personalize lymphoma management establishes feasibility of capture sequencing of routinely acquired FFPE tissue (DLBCL) and fresh cell preparations (CLL) in a single assay. It demonstrates superiority of hybrid-capture sequencing over amplicon-based method with respect to assay sensitivity in an approach intended to cover a large

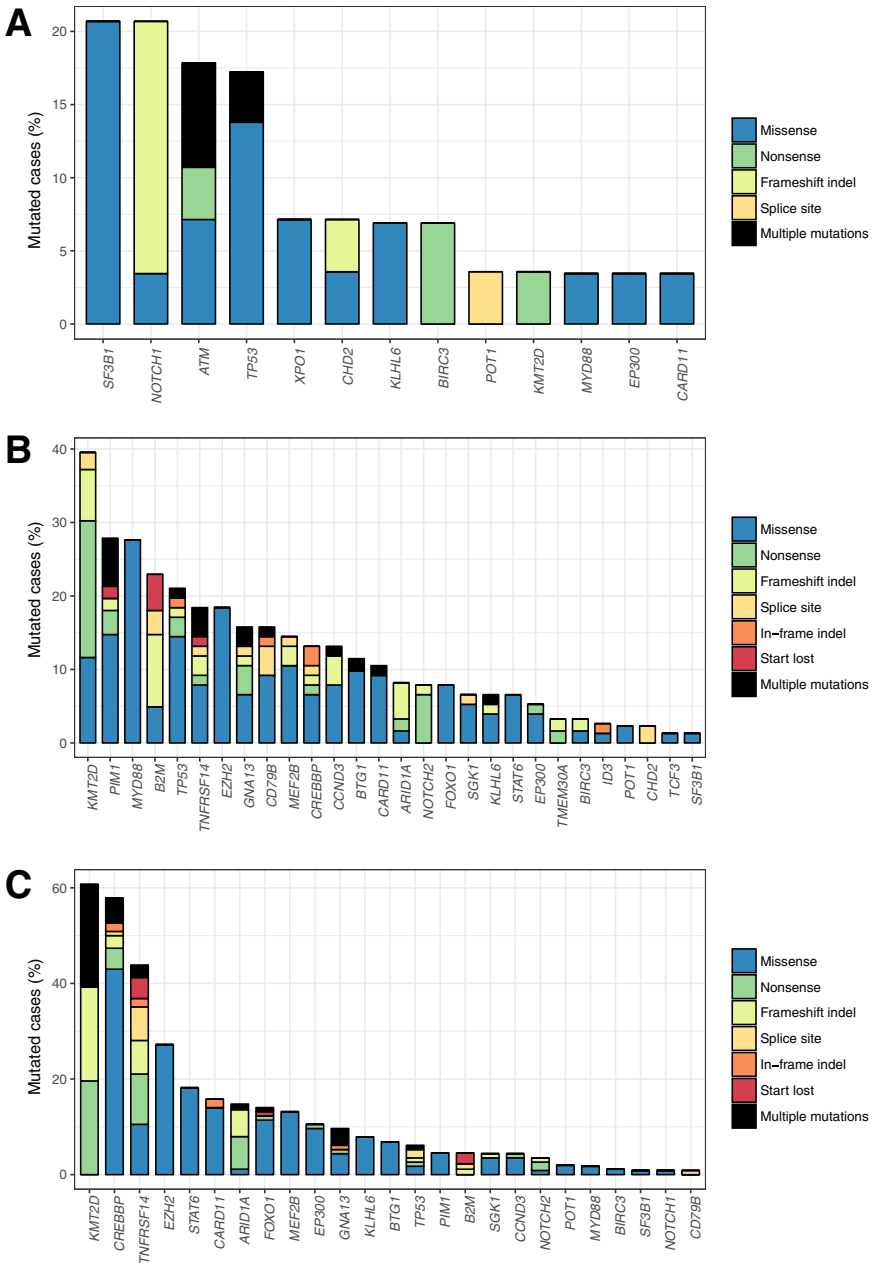

Figure 3 Frequency of mutations across all mutated genes in chronic lymphocytic leukemia (**A**), diffuse large B-cell lymphoma (**B**), and follicular lymphoma (**C**), divided into the classes frameshift insertion/deletion (indel), in-frame indel, missense, nonsense, splice site, and start lost counts. Genes carrying more than one class of mutation are categorized as multiple mutations.

capture space across the coding region of 32 genes. Finally, using an iterative process of assay development and pipeline optimization, the comparability of assay results was shown to the published literature in FL, DLBCL, and CLL. As evidenced by the version progression and evolution of the assays over a 3-year time period, the modularity of the assay platform that allows for incorporation of additional gene targets as the knowledge base around actionable gene mutations evolves in lymphoid cancers was also demonstrated.

The sequencing design was tightly linked to the specific needs and parameters for sequencing lymphoid cancers. These parameters include the need for assays applicable to FFPE material routinely produced in diagnostic workflows, the capture space across entire gene bodies harboring potentially actionable gene mutations (hotspot mutations and scattered inactivating gene mutations), and the application range across the most common lymphoid cancers (DLBCL, FL, and CLL). Given this framework, the 32-gene capture sequencing design proved to be ideal and outperformed amplicon-sequencing–based approaches. However, amplicon-based methods, which are designed to simplify preparation and use smaller DNA inputs, have also been successfully used in other contexts; in particular, they have been used in applications probing for specific gene mutations and constitutional variants.[54,69,70] The application of this assay to FFPE proved to be a major driver for platform selection. The use of hybridization capture in highly degraded and damaged DNA led to more homogeneous

coverage and fewer library construction–induced artifacts. More important, capture sequencing enabled bioinformatics alignment based on nonduplicate reads avoiding amplification of PCR-associated noise. Varying GC content, small and variable DNA fragment sizes and the presence of repetitive sequences are well-recognized challenges for the PCR-based method, resulting in biased amplification. These shortcomings in the amplicon-based TSCA pipeline manifested predominantly in limited sensitivity because of dropout amplicons and low-coverage regions.

Through comparisons with Sanger sequencing ($n = 588$) and WGS ($n = 66$) in a large subset of cases, thresholds for optimal assay accuracy, which was estimated at 93% sensitivity and 100% specificity, were adjusted. An approach that would maximize specificity to avoid false-positive actionable mutations that may lead to projected management changes deviating from standard of care was chosen. Although minimum read and AF filters are objective and standardized approaches for mutation calling, a manual curation step based on individual inspection of variants using read alignment visualization tools (ie, Integrated Genome Viewer browser[18]) was also maintained. Failure criteria were also established for individual cases to maintain adequate coverage across the capture space. Deep coverage improves the sensitivity of detecting mutations, including subclonal mutations, with more recent studies suggesting that >100× coverage is preferable for sensitive single-nucleotide variant detection in exomes.[71,72] Notably, although allele frequency distributions varied considerably among the lymphoma entities because of differences in tumor purity, sequencing coverage was consistent, emphasizing the importance of setting a strict coverage threshold for maintaining high-quality sequencing data (Supplemental Figure S7). Herein, a strict threshold of 50× was implemented for failing samples (ie, exclude from the analysis), while setting a soft threshold of 100× for inclusion in the study (Supplemental Figure S8). These thresholds proved to be practical (only 4% failure rate, even in highly degraded FFPE tissue), while maintaining sensitivity to reliably detect variants down to the lower bound of 10% AF with 10 supporting reads.

This sequencing assay has been developed for the ultimate purpose of guiding treatment decisions in the era of personalized medicine in lymphoid cancers to capture interpatient heterogeneity with specific treatment implications. The genes included in this assay have been selected on the basis of the criteria of reported outcome correlation, improved diagnostic accuracy, biological rationale for targeted therapy, and mutation-associated drug targets. Although the projected utility of targeted sequencing in future clinical trials and standard of care might be compelling, the field has yet to demonstrate clear clinical utility of sequencing in lymphoid cancers, a demonstration that should be aided by robust assays, such as we describe. Other challenges remain that relate to the integration with other biomarkers, standardization of annotation and reporting of variants, and interpretation of individual variants that are currently of unknown significance.

In conclusion, this work highlights the value of a comprehensive approach to targeted capture sequencing that integrates preanalytical, analytical, and postanalytical quality measures and analyses and offers reliable detection of gene mutations from FFPE and fresh specimens of FL, DLBCL, and CLL. The described strategy and similar approaches with the demonstrated potential for assay evolution are the first steps to expanding diagnostic and predictive biomarker testing based on sequencing in lymphoid cancers. Our results warrant further development of capture sequencing assays in regulated diagnostic pathology environments, to demonstrate clinical utility and socioeconomic benefit.

Acknowledgments

We thank the Research Oversight Committee of a Genome Canada Large-Scale Applied Research Project in Genomics and Personalized Health for their scientific input and fruitful discussions.

Supplemental Data

Supplementary data related to this article can be found at *https://doi.org/10.1016/j.jmoldx.2017.11.010*.

References

1. Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, Advani R, Ghielmini M, Salles GA, Zelenetz AD, Jaffe ES: The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood 2016, 127:2375–2391
2. Diamandis M, White NMA, Yousef GM: Personalized medicine: marking a new epoch in cancer patient management. Mol Cancer Res 2010, 8:1175–1187
3. Schrijver I, Aziz N, Farkas DH, Furtado M, Gonzalez AF, Greiner TC, Grody WW, Hambuch T, Kalman L, Kant JA, Klein RD, Leonard DGB, Lubin IM, Mao R, Nagan N, Pratt VM, Sobel ME, Voelkerding KV, Gibson JS: Opportunities and challenges associated with clinical diagnostic genome sequencing: a report of the Association for Molecular Pathology. J Mol Diagn 2012, 14:525–540
4. Pastore A, Jurinovic V, Kridel R, Hoster E, Staiger AM, Szczepanowski M, Pott C, Kopp N, Murakami M, Horn H, Leich E, Moccia AA, Mottok A, Sunkavalli A, Van Hummelen P, Ducar M, Ennishi D, Shulha HP, Hother C, Connors JM, Sehn LH, Dreyling M, Neuberg D, Moller P, Feller AC, Hansmann ML, Stein H, Rosenwald A, Ott G, Klapper W, Unterhalt M, Hiddemann W, Gascoyne RD, Weinstock DM, Weigert O: Integration of gene mutations in risk prognostication for patients receiving first-line immunochemotherapy for follicular lymphoma: a retrospective analysis of a prospective clinical trial and validation in a population-based registry. Lancet Oncol 2015, 16:1111–1122
5. Wilson WH, Young RM, Schmitz R, Yang Y, Pittaluga S, Wright G, Lih C-J, Williams PM, Shaffer AL, Gerecitano J, de Vos S, Goy A, Kenkre VP, Barr PM, Blum KA, Shustov A, Advani R, Fowler NH, Vose JM, Elstrom RL, Habermann TM, Barrientos JC, McGreivy J, Fardis M, Chang BY, Clow F, Munneke B, Moussa D, Beaupre DM, Staudt LM: Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. Nat Med 2015, 21:922–926
6. Nadeu F, Delgado J, Royo C, Baumann T, Stankovic T, Pinyol M, Jares P, Navarro A, Martin-Garcia D, Bea S, Salaverria I, Oldreive C,

Aymerich M, Suarez-Cisneros H, Rozman M, Villamor N, Colomer D, Lopez-Guillermo A, Gonzalez M, Alcoceba M, Terol MJ, Colado E, Puente XS, Lopez-Otin C, Enjuanes A, Campo E: Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia. Blood 2016, 127:2122–2130

7. Mwenifumbo JC, Marra MA: Cancer genome-sequencing study design. Nat Rev Genet 2013, 14:321–332
8. Lim MS, Elenitoba-Johnson KSJ: Precision medicine for diffuse large B-cell lymphoma. Clin Cancer Res 2016, 22:2829–2831
9. Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J, Turner DJ: Target-enrichment strategies for next-generation sequencing. Nat Methods 2010, 7: 111–118
10. Altmuller J, Budde BS, Nurnberg P: Enrichment of target sequences for next-generation sequencing applications in research and diagnostics. Biol Chem 2014, 395:231–237
11. Choi M, Scholl UI, Ji W, Liu T, Tikhonova IR, Zumbo P, Nayir A, Bakkaloglu A, Ozen S, Sanjad S, Nelson-Williams C, Farhi A, Mane S, Lifton RP: Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. Proc Natl Acad Sci U S A 2009, 106:19096–19101
12. Samorodnitsky E, Jewell BM, Hagopian R, Miya J, Wing MR, Lyon E, Damodaran S, Bhatt D, Reeser JW, Datta J, Roychowdhury S: Evaluation of hybridization capture versus amplicon-based methods for whole-exome sequencing. Hum Mutat 2015, 36:903–914
13. Wang J, Yu H, Zhang VW, Tian X, Feng Y, Wang G, Gorman E, Wang H, Lutz RE, Schmitt ES, Peacock S, Wong L-J: Capture-based high-coverage NGS: a powerful tool to uncover a wide spectrum of mutation types. Genet Med 2016, 18:513–521
14. Wong SQ, Li J, Salemi R, Sheppard KE, Do H, Tothill RW, McArthur GA, Dobrovic A: Targeted-capture massively-parallel sequencing enables robust detection of clinically informative mutations from formalin-fixed tumours. Sci Rep 2013, 3:3494
15. Li H: Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv.org 2013, 1303.3997:1–3
16. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK: VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 2012, 22:568–576
17. Saunders CT, Wong WSW, Swamy S, Becq J, Murray LJ, Cheetham RK: Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. Bioinformatics 2012, 28: 1811–1817
18. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP: Integrative genomics viewer. Nat Biotechnol 2011, 29:24–26
19. Yost SE, Alakus H, Matsui H, Schwab RB, Jepsen K, Frazer KA, Harismendy O: Mutascope: sensitive detection of somatic mutations from deep amplicon sequencing. Bioinformatics 2013, 29: 1908–1909
20. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG: Primer3—new capabilities and interfaces. Nucleic Acids Res 2012, 40:e115
21. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R: The sequence alignment/map format and SAMtools. Bioinformatics 2009, 25:2078–2079
22. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM: A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) 2012, 6:80–92
23. Sherry ST: dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 2001, 29:308–311
24. Li H, Durbin R: Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 2010, 26:589–595
25. Ding J, Bashashati A, Roth A, Oloumi A, Tse K, Zeng T, Haffari G, Hirst M, Marra MA, Condon A, Aparicio S, Shah SP: Feature-based classifiers for somatic mutation detection in tumour-normal paired sequencing data. Bioinformatics 2012, 28:167–175
26. Flicek P, Amode MR, Barrell D, Beal K, Billis K, Brent S, et al: Ensembl 2014. Nucleic Acids Res 2014, 42:D749–D755
27. Mamedov TG, Pienaar E, Whitney SE, TerMaat JR, Carvill G, Goliath R, Subramanian A, Viljoen HJ: A fundamental study of the PCR amplification of GC-rich DNA templates. Comput Biol Chem 2008, 32:452–457
28. Bustin SA, Nolan T: Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. J Biomol Tech 2004, 15:155–166
29. Wang L, Lawrence MS, Wan Y, Stojanov P, Sougnez C, Stevenson K, Werner L, Sivachenko A, DeLuca DS, Zhang L, Zhang W, Vartanov AR, Fernandes SM, Goldstein NR, Folco EG, Cibulskis K, Tesar B, Sievers QL, Shefler E, Gabriel S, Hacohen N, Reed R, Meyerson M, Golub TR, Lander ES, Neuberg D, Brown JR, Getz G, Wu CJ: SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. N Engl J Med 2011, 365:2497–2506
30. Guieze R, Robbe P, Clifford R, de Guibert S, Pereira B, Timbs A, Dilhuydy M-S, Cabes M, Ysebaert L, Burns A, Nguyen-Khac F, Davi F, Veronese L, Combes P, Le Garff-Tavernier M, Leblond V, Merle-Beral H, Alsolami R, Hamblin A, Mason J, Pettitt A, Hillmen P, Taylor J, Knight SJL, Tournilhac O, Schuh A: Presence of multiple recurrent mutations confers poor trial outcome of relapsed/refractory CLL. Blood 2015, 126:2110–2117
31. Chiaretti S, Marinelli M, Del Giudice I, Bonina S, Piciocchi A, Messina M, Vignetti M, Rossi D, Di Maio V, Mauro FR, Guarini A, Gaidano G, Foa R: NOTCH1, SF3B1, BIRC3 and TP53 mutations in patients with chronic lymphocytic leukemia undergoing first-line treatment: correlation with biological parameters and response to treatment. Leuk Lymphoma 2014, 55:2785–2792
32. Ennishi D, Hoffer C, Shulha H, Mottok A, Farinha P, Chan FC, Meissner B, Boyle M, Ben-Neriah S, Morin RD, Marra MA, Savage KJ, Sehn LH, Connors JM, Steidl C, Scott DW, Gascoyne RD: Clinical significance of genetic aberrations in diffuse large B cell lymphoma. Blood 2014, 124:703
33. Dubois S, Viailly P-J, Mareschal S, Bohers E, Bertrand P, Ruminy P, Maingonnat C, Jais J-P, Peyrouze P, Figeac M, Molina TJ, Desmots F, Fest T, Haioun C, Lamy T, Copie-Bergman C, Briere J, Petrella T, Canioni D, Fabiani B, Coiffier B, Delarue R, Peyrade F, Bosly A, Andre M, Ketterer N, Salles G, Tilly H, Leroy K, Jardin F: Next-generation sequencing in diffuse large B-cell lymphoma highlights molecular divergence and therapeutic opportunities: a LYSA Study. Clin Cancer Res 2016, 22:2919–2928
34. Schmitz R, Young RM, Ceribelli M, Jhavar S, Xiao W, Zhang M, et al: Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. Nature 2012, 490:116–120
35. Rodriguez D, Bretones G, Quesada V, Villamor N, Arango JR, Lopez-Guillermo A, Ramsay AJ, Baumann T, Quiros PM, Navarro A, Royo C, Martin-Subero JI, Campo E, Lopez-Otin C: Mutations in CHD2 cause defective association with active chromatin in chronic lymphocytic leukemia. Blood 2015, 126:195–202
36. Pasqualucci L, Dominguez-Sola D, Chiarenza A, Fabbri G, Grunn A, Trifonov V, Kasper LH, Lerach S, Tang H, Ma J, Rossi D, Chadburn A, Murty VV, Mullighan CG, Gaidano G, Rabadan R, Brindle PK, Dalla-Favera R: Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature 2011, 471:189–195
37. Jiang Y, Ortega-Molina A, Geng H, Ying H-Y, Hatzi K, Parsa S, McNally D, Wang L, Doane AS, Agirre X, Teater M, Meydan C, Li Z, Poloway D, Wang S, Ennishi D, Scott DW, Stengel KR, Kranz JE, Holson E, Sharma S, Young JW, Chu C-S, Roeder RG, Shaknovich R, Hiebert SW, Gascoyne RD, Tam W, Elemento O, Wendel H-G, Melnick AM: CREBBP inactivation promotes the development of HDAC3-dependent lymphomas. Cancer Discov 2017, 7:38–53
38. Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, et al: Somatic mutations altering EZH2 (Tyr641) in follicular and

diffuse large B-cell lymphomas of germinal-center origin. Nat Genet 2010, 42:181−185
39. Morin RD, Assouline S, Alcaide M, Mohajeri A, Johnston RL, Chong L, Grewal J, Yu S, Fornika D, Bushell K, Nielsen TH, Petrogiannis-Haliotis T, Crump M, Tosikyan A, Grande BM, MacDonald D, Rousseau C, Bayat M, Sesques P, Froment R, Albuquerque M, Monczak Y, Oros KK, Greenwood C, Riazalhosseini Y, Arseneault M, Camlioglu E, Constantin A, Qiang PH, Peng R, Mann KK, Johnson NA: Genetic landscapes of relapsed and refractory diffuse large B-cell lymphomas. Clin Cancer Res 2016, 22:2290−2300
40. Trinh DL, Scott DW, Morin RD, Mendez-Lago M, An J, Jones SJM, Mungall AJ, Zhao Y, Schein J, Steidl C, Connors JM, Gascoyne RD, Marra MA: Analysis of FOXO1 mutations in diffuse large B-cell lymphoma. Blood 2013, 121:3666−3674
41. Muppidi JR, Schmitz R, Green JA, Xiao W, Larsen AB, Braun SE, An J, Xu Y, Rosenwald A, Ott G, Gascoyne RD, Rimsza LM, Campo E, Jaffe ES, Delabie J, Smeland EB, Braziel RM, Tubbs RR, Cook JR, Weisenburger DD, Chan WC, Vaidehi N, Staudt LM, Cyster JG: Loss of signalling via Galpha13 in germinal centre B-cell-derived lymphoma. Nature 2014, 516:254−258
42. Gebauer N, Bernard V, Feller AC, Merz H: ID3 mutations are recurrent events in double-hit B-cell lymphomas. Anticancer Res 2013, 33:4771−4778
43. Puente XS, Pinyol M, Quesada V, Conde L, Ordóñez GR, Villamor N, et al: Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature 2011, 475:101−105
44. Zhang J, Grubor V, Love CL, Banerjee A, Richards KL, Mieczkowski PA, et al: Genetic heterogeneity of diffuse large B-cell lymphoma. Proc Natl Acad Sci U S A 2013, 110:1398−1403
45. Ying CY, Dominguez-Sola D, Fabi M, Lorenz IC, Hussein S, Bansal M, Califano A, Pasqualucci L, Basso K, Dalla-Favera R: MEF2B mutations lead to deregulated expression of the oncogene BCL6 in diffuse large B cell lymphoma. Nat Immunol 2013, 14: 1084−1092
46. Pon JR, Wong J, Saberi S, Alder O, Moksa M, Grace Cheng S-W, Morin GB, Hoodless PA, Hirst M, Marra MA: MEF2B mutations in non-Hodgkin lymphoma dysregulate cell migration by decreasing MEF2B target gene activation. Nat Commun 2015, 6:7953
47. Zhang J, Dominguez-Sola D, Hussein S, Lee J-E, Holmes AB, Bansal M, Vlasevska S, Mo T, Tang H, Basso K, Ge K, Dalla-Favera R, Pasqualucci L: Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. Nat Med 2015, 21:1190−1198
48. Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim K-H, Kohlhammer H, Xu W, Yang Y, Zhao H, Shaffer AL, Romesser P, Wright G, Powell J, Rosenwald A, Muller-Hermelink HK, Ott G, Gascoyne RD, Connors JM, Rimsza LM, Campo E, Jaffe ES, Delabie J, Smeland EB, Fisher RI, Braziel RM, Tubbs RR, Cook JR, Weisenburger DD, Chan WC, Staudt LM: Oncogenically active MYD88 mutations in human lymphoma. Nature 2011, 470: 115−119
49. Fabbri G, Rasi S, Rossi D, Trifonov V, Khiabanian H, Ma J, Grunn A, Fangazio M, Capello D, Monti S, Cresta S, Gargiulo E, Forconi F, Guarini A, Arcaini L, Paulli M, Laurenti L, Larocca LM, Marasca R, Gattei V, Oscier D, Bertoni F, Mullighan CG, Foá R, Pasqualucci L, Rabadan R, Dalla-Favera R, Gaidano G: Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. J Exp Med 2011, 208:1389−1401
50. Shukla V, Ma S, Joshi S, Lu R: Notch2 is critical for CLL development in the IRF4 -/- Vh11 mice. Blood 2014, 124:891
51. Kiel MJ, Velusamy T, Betz BL, Zhao L, Weigelin HG, Chiang MY, Huebner-Chan DR, Bailey NG, Yang DT, Bhagat G, Miranda RN, Bahler DW, Medeiros LJ, Lim MS, Elenitoba-Johnson KSJ: Whole-genome sequencing identifies recurrent somatic NOTCH2 mutations in splenic marginal zone lymphoma. J Exp Med 2012, 209: 1553−1565
52. Peters TL, Li L, Tula-Sanchez AA, Pongtornpipat P, Schatz JH: Control of translational activation by PIM kinase in activated B-cell diffuse large B-cell lymphoma confers sensitivity to inhibition by PIM447. Oncotarget 2016, 7:63362−63373
53. Amin NA, Seymour E, Saiya-Cork K, Parkin B, Shedden K, Malek SN: A quantitative analysis of subclonal and clonal gene mutations before and after therapy in chronic lymphocytic leukemia. Clin Cancer Res 2016, 22:4525−4535
54. Gunawardana J, Chan FC, Telenius A, Woolcock B, Kridel R, Tan KL, Ben-Neriah S, Mottok A, Lim RS, Boyle M, Rogic S, Rimsza LM, Guiter C, Leroy K, Gaulard P, Haioun C, Marra MA, Savage KJ, Connors JM, Shah SP, Gascoyne RD, Steidl C: Recurrent somatic mutations of PTPN1 in primary mediastinal B cell lymphoma and Hodgkin lymphoma. Nat Genet 2014, 46:329−335
55. Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD, et al: Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature 2011, 476:298−303
56. Yildiz M, Li H, Bernard D, Amin NA, Ouillette P, Jones S, Saiya-Cork K, Parkin B, Jacobi K, Shedden K, Wang S, Chang AE, Kaminski MS, Malek SN: Activating STAT6 mutations in follicular lymphoma. Blood 2015, 125:668−679
57. Tyryshkin K, Li YD, Good D, Shepherd LE, Baetz T, Rauh MJ, LeBrun DP: Differential expression of TCF3 target genes defines subclasses of diffuse large B-cell lymphoma with striking differences in clinical outcome following R-CHOP therapy. Blood 2016, 128:3037
58. Cheung K-JJ, Johnson NA, Affleck JG, Severson T, Steidl C, Ben-Neriah S, Schein J, Morin RD, Moore R, Shah SP, Qian H, Paul JE, Telenius A, Relander T, Lam W, Savage K, Connors JM, Brown C, Marra MA, Gascoyne RD, Horsman DE: Acquired TNFRSF14 mutations in follicular lymphoma are associated with worse prognosis. Cancer Res 2010, 70:9166−9174
59. Lohr JG, Stojanov P, Lawrence MS, Auclair D, Chapuy B, Sougnez C, Cruz-Gordillo P, Knoechel B, Asmann YW, Slager SL, Novak AJ, Dogan A, Ansell SM, Link BK, Zou L, Gould J, Saksena G, Stransky N, Rangel-Escareño C, Fernandez-Lopez JC, Hidalgo-Miranda A, Melendez-Zajgla J, Hernández-Lemus E, Schwarz-Cruz y Celis A, Imaz-Rosshandler I, Ojesina AI, Jung J, Pedamallu CS, Lander ES, Habermann TM, Cerhan JR, Shipp MA, Getz G, Golub TR: Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proc Natl Acad Sci U S A 2012, 109:3879−3884
60. O'Shea D, O'Riain C, Taylor C, Waters R, Carlotti E, MacDougall F, Gribben J, Rosenwald A, Ott G, Rimsza LM, Smeland EB, Johnson N, Campo E, Greiner TC, Chan WC, Gascoyne RD, Wright G, Staudt LM, Lister TA, Fitzgibbon J: The presence of TP53 mutation at diagnosis of follicular lymphoma identifies a high-risk group of patients with shortened time to disease progression and poorer overall survival. Blood 2008, 112:3126−3129
61. Pasqualucci L, Trifonov V, Fabbri G, Ma J, Rossi D, Chiarenza A, Wells VA, Grunn A, Messina M, Elliot O, Chan J, Bhagat G, Chadburn A, Gaidano G, Mullighan CG, Rabadan R, Dalla-Favera R: Analysis of the coding genome of diffuse large B-cell lymphoma. Nat Genet 2011, 43:830−837
62. Kridel R, Chan FC, Mottok A, Boyle M, Farinha P, Tan K, Meissner B, Bashashati A, McPherson A, Roth A, Shumansky K, Yap D, Ben-Neriah S, Rosner J, Smith MA, Nielsen C, Giné E, Telenius A, Ennishi D, Mungall A, Moore R, Morin RD, Johnson NA, Sehn LH, Tousseyn T, Dogan A, Connors JM, Scott DW, Steidl C, Marra MA, Gascoyne RD, Shah SP: Histological transformation and progression in follicular lymphoma: a Clonal Evolution Study. PLoS Med 2016, 13:e1002197
63. Pasqualucci L, Dalla-Favera R: The genetic landscape of diffuse large B-cell lymphoma. Semin Hematol 2015, 52:67−76
64. Okosun J, Montoto S, Fitzgibbon J: The routes for transformation of follicular lymphoma. Curr Opin Hematol 2016, 23:385−391
65. Ennishi D, Mottok A, Ben-Neriah S, Shulha HP, Farinha P, Chan FC, Meissner B, Boyle M, Hother C, Kridel R, Lai D, Saberi S,

Bashashati A, Shah SP, Morin RD, Marra MA, Savage KJ, Sehn LH, Steidl C, Connors JM, Gascoyne RD, Scott DW: Genetic profiling of MYC and BCL2 in diffuse large B-cell lymphoma determines cell-of-origin-specific clinical impact. Blood 2017, 129:2760−2770

66. Landau DA, Tausch E, Taylor-Weiner AN, Stewart C, Reiter JG, Bahlo J, Kluth S, Bozic I, Lawrence M, Bottcher S, Carter SL, Cibulskis K, Mertens D, Sougnez CL, Rosenberg M, Hess JM, Edelmann J, Kless S, Kneba M, Ritgen M, Fink A, Fischer K, Gabriel S, Lander ES, Nowak MA, Dohner H, Hallek M, Neuberg D, Getz G, Stilgenbauer S, Wu CJ: Mutations driving CLL and their evolution in progression and relapse. Nature 2015, 526:525−530
67. Rossi D, Trifonov V, Fangazio M, Bruscaggin A, Rasi S, Spina V, et al: The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. J Exp Med 2012, 209:1537−1551
68. Zhang X, Shi Y, Weng Y, Lai Q, Luo T, Zhao J, Ren G, Li W, Pan H, Ke Y, Zhang W, He Q, Wang Q, Zhou R: The truncate mutation of Notch2 enhances cell proliferation through activating the NF-κB signal pathway in the diffuse large B-cell lymphomas. PLoS One 2014, 9:e108747
69. Bosdet IE, Docking TR, Butterfield YS, Mungall AJ, Zeng T, Coope RJ, Yorida E, Chow K, Bala M, Young SS, Hirst M, Birol I, Moore RA, Jones SJ, Marra MA, Holt R, Karsan A: A clinically validated diagnostic second-generation sequencing assay for detection of hereditary BRCA1 and BRCA2 mutations. J Mol Diagn 2013, 15:796−809
70. Rohr J, Guo S, Huo J, Bouska A, Lachel C, Li Y, Simone PD, Zhang W, Gong Q, Wang C, Cannon A, Heavican T, Mottok A, Hung S, Rosenwald A, Gascoyne R, Fu K, Greiner TC, Weisenburger DD, Vose JM, Staudt LM, Xiao W, Borgstahl GEO, Davis S, Steidl C, McKeithan T, Iqbal J, Chan WC: Recurrent activating mutations of CD28 in peripheral T-cell lymphomas. Leukemia 2016, 30:1062−1070
71. Clark MJ, Chen R, Lam HYK, Karczewski KJ, Chen R, Euskirchen G, Butte AJ, Snyder M: Performance comparison of exome DNA sequencing technologies. Nat Biotechnol 2011, 29:908−914
72. Meynert AM, Bicknell LS, Hurles ME, Jackson AP, Taylor MS: Quantifying single nucleotide variant detection sensitivity in exome sequencing. BMC Bioinformatics 2013, 14:195