Comprehensive translocation and clonality detection in lymphoproliferative disorders by next-generation sequencing

Detection and characterization of clonal immunoglobulin (IG)/T-cell receptor (TR) rearrangements and translocations in lymphoproliferative neoplasms provide critical information in the diagnostic pathway and are valuable tools for addressing research questions involving B- and T-cell lymphoproliferative disorders (LPD).[1,2] These include ascertaining the clonal nature of lymphoid proliferations,[3,4] characterization of translocations in lymphomas and leukemias,[4] characterization of CDR3 regions for minimal residual disease target identification[5] and stereotyping analysis.[6] Until recently, collecting this information required a combination of different methodologies, such as gene-scanning/heteroduplex analysis, fluorescence *in situ* hybridization (FISH) and Sanger sequencing. The incorporation of next-generation sequencing (NGS) in clinical laboratories opens up new possibilities since an integrated NGS approach can provide data on sequence and structural variation in a single assay, including translocations and IG/TR rearrangements, and has been shown to be successful for the characterisation of IG translocations in myeloma and lymphomas.[7,8]

Within the EuroClonality-NGS consortium, we have designed a capture-based protocol covering the coding V, D and J genes of the IG/TR loci, as well as switch regions in the IGH locus. This design allows the identification of D-J and V-(D)-J rearrangements as well as chromosomal translocations involving IG/TR genes by sequencing through the breakpoint regions in genomic DNA. We piloted this approach using a sample cohort (n=24) consisting of three B-cell precursor acute lymphoblastic leukemias (BCP-ALL), four Burkitt lymphomas, eight chronic lymphocytic leukemias (CLL), two splenic marginal zone lymphomas (SMZL), two diffuse large B-cell lymphomas (DLBCL), two follicular lymphomas, two T-cell acute lymphoblastic leukemias (T-ALL) and one T-cell non-Hodgkin lymphoma (T-NHL). Twenty-one cases were known to carry a translocation arising within the IG/TR loci with the remaining three cases being included for their well characterized D-J or V-(D)-J gene rearrangements. Libraries were constructed from 1 μg of genomic DNA which was fragmented to an average of 200 bp using an E220 Focused-ultrasonicator (Covaris, Woburn, MA, USA). Fragmented DNA was processed using the TruSeq DNA LT sample preparation kit (Illumina, Cambridge, UK). Libraries were hybridized to a custom-designed EZ SeqCap gene panel (Roche-Nimblegen, Madison, MI, USA) which encompassed 180 kb containing the V, D, J and constant regions of the IG and TR loci as well as the switch regions of the IGH locus. Enriched samples were sequenced on a MiSeq (Illumina) using 75 bp or 120 bp paired-end reads. Reads were aligned to the reference genome (hg19) with translocations and variants called using a previously described bioinformatics pipeline.[9] The average depth of sequencing in the 21 samples with successful NGS results was 322x. IG/TR gene rearrangements were also determined by polymerase chain reaction (PCR) analysis and Sanger sequencing using the BIOMED-2 protocol in 14 cases.[1] For the detailed characterization of D-J/V(D)J gene rearrangements the IMGT V-Quest software was used.[5] IG/TR

Table 1. Translocations detected by karyotyping or FISH and the EuroClonality NGS panel.

| Sample | Diagnosis | Karyotyping or FISH results | NGS Capture results | | | | | | | |
| | | | der(IG/TR) | | | | der(partner chromosome) | | | |
			Break 1	Location	Break 2	Location	Break 1	Location	Break 2	Location
EC20	BCP-ALL	t(X;14)(p22;q32)	IGHJ6	chr14:106329465	*CRLF2*	chr.X/Y:1358468	IGHD3-9	chr14:106370544	*CRLF2*	chrX/Y:1358472
EC19	BCP-ALL	t(Y;14)	IGHJ5	chr14:106330070	*CRLF2*	chr.X/Y:1333757	IGHD6-19	chr14:106357574	*CRLF2*	chrX/Y:1333695
EC21	Burkitt lymphoma	t(8;14)	ND	ND	ND	ND	IGHA1	chr14:106177928	*MYC*	chr8:128749373
EC23	Burkitt lymphoma	t(8;14)	IGHG1 and IGHD2-2	chr14: complex	*MYC*	chr8:128746892	IGHG2/3*	chr14	*MYC*	chr8:128746903
EC18	Burkitt lymphoma	t(8;14)	IGHJ4	chr14:106330462	*MYC*	chr8:128746504	ND	ND	ND	ND
EC17	Burkitt lymphoma	t(8;14)	IGHA1	chr14:106177015	*MYC*	chr8:128749477	IGHA1	chr14:106177496	*MYC*	chr8:128749448
EC30	CLL	t(2;14)(p13;32)	IGHM	chr14:106327487	*BCL11A*	chr2:60781261	IGHM	chr14:106327500	*BCL11A*	chr2:60781273
EC31	CLL	t(14;18)(q32;q21)	IGHJ6	chr14:106329459	*BCL2*	chr18:60793474	ND	ND	ND	ND
EC29	CLL	t(14;18)(q32;q21)	IGHJ5	chr14:106330071	*BCL2*	chr18:60769465	IGHD2-15	chr14: 106363819	*BCL2*	chr18:60769465
EC5	DLBCL	t(14;18)	IGHJ5	chr14:106330072	*BCL2*	chr18:60793494	IGHV3-21	chr14:106691673	*BCL2*	chr18:60793498
EC34	SMZL	t(6;14)(p21;q32)	IGHA2*	chr14	*CCND3*	chr6:41942244	ND	ND	ND	ND
EC3	DLBCL	IGH break‡	IGHM	chr14:106326706	*IRF4*	chr6:472682	IGHG4*	chr14	*EXOC2*	chr6
EC25	CLL	t(14;16)(q32;q22)	ND	ND	ND	ND	IGHD2-2	chr14: 106382689	unknown†	chr16:69479932
EC24	CLL	t(1;14)(q32;q32)	IGHM or IGHA1*	chr14	unknown†	chr1:206286226	IGHM or IGHA1*	chr14	unknown†	chr1:206286210
EC33	SMZL	t(5;14)(q13;q32)	IGHM	chr14:106326138	unknown†	chr5:88608990	IGHM	chr14:106326162	unknown†	chr5:88608986
EC11	T-ALL	inv14/t(14;14)	TRDD3	chr14:22918106	*BCL11B*	chr14:99689173	ND	ND	ND	ND
EC12	T-ALL	t(7;10)	TRBJb2.5	chr7:142494805	*TLX1*	chr10:102902431	ND	ND	ND	ND
EC14	T-NHL	inv7	TRGV8	ND	TRBJb2.7	ND	ND	ND	ND	ND

*Alignment equivocal due to sequence homology. †No known lymphoma/leukemia-associated gene in proximity. ‡demonstrated by FISH IGH break-apart probe. ND: exact breakpoint could not be determined due to insufficient or lack of aligned reads. BCP-ALL: B-cell precursor acute lymphoblastic leukemia; CLL: chronic lymphocytic leukemia; DLBCL: diffuse large B-cell lymphoma; SMZL: splenic marginal zone lymphoma; T-ALL: T-cell acute lymphoblastic leukemia; T-NHL: T-cell non-Hodgkin lymphoma.

translocations had been previously determined by routine FISH, karyotyping and/or Sanger sequencing in the referring laboratories.

In 18 out of 21 samples with a known translocation, identified by either FISH or karyotyping, the breakpoints were identified by the NGS capture panel (Table 1). Of the three samples that failed to produce results, two were fresh frozen samples from lymphomas with degraded and low-quantity DNA (<500 ng total DNA and <1 kb median fragment size by TapeStation analysis) and one was a case of BCP-ALL in which a technical error occurred due to evaporation during the hybridization step.

Of the 18 samples that yielded results, 15 (83%) were from patients with B-cell LPD, of which 11 had well-

Table 2. IGH and IGK rearrangements detected by Sanger sequencing and the EuroClonality NGS panel.

Sample	Diagnosis	Results	IGH Allele 1 IGHV	IGHD	IGHJ	IGH Allele 2 IGHV	IGHD	IGHJ	IGK Allele 1 IGKV	IGKJ	IGK Allele 2 IGKV	IGKJ
EC21	Burkitt lymphoma	Sanger Seq.	3-15	3-22	4	-	3-16	4				
		NGS	3-15	3-22	4	-	3-16	4	1-9	2	D1-13	2
EC23	Burkitt lymphoma	Sanger Seq.	3-23	-	4	-	-	-	4-1	3	-	-
		NGS	3-23	4-23	4	-	-	-	4-1	3	-	-
EC18	Burkitt lymphoma	Sanger Seq.	3-22	6-13	4	-	-	-	4-1	2	intron RSS-KDE*	
		NGS	3-22	6-13	4	-	-	-	4-1	2	intron RSS/VK1-8	
EC30	CLL	Sanger Seq.	4-39	6-25	4	-	-	-				
		NGS	4-39	6-25	4	-	-	-	4-1	1	-	-
EC31	CLL	Sanger Seq.	3-30	2-2	4	-	-	-				
		NGS	3-30	-	4	-	-	-	4-1	3	KDE-VK1-8	
EC29	CLL	Sanger Seq.	5-51	4-17	4	-	-	-				
		NGS	5-51	4-17	4	-	-	-	1-16	4	-	-
EC25	CLL	Sanger Seq.	4-34	5-18	6	-	-	-				
		NGS	4-34	5-18	6	-	2-8	4	4-1	2	-	-
EC24	CLL	Sanger Seq.	4-61	6-19	5	5-51	5-12	4				
		NGS	4-61	6-19	5	5-51	5-12	4	2-30	2	4-1	3
EC33	SMZL	Sanger Seq.	3-23	3-22	3	-	-	-				
		NGS	3-23	3-22	3	-	-	-	1-5	1	-	-
EC32	SMZL	Sanger Seq.	2-5	6-19	2	-	-	-				
		NGS	2-5	6-19	2	-	-	-	2-28	2	-	-
EC34	SMZL	Sanger Seq.	5-51	4-4	6	-	-	-				
		NGS	5-51	4-4	6	-	-	-	-	-	-	-
EC27	CLL	Sanger Seq.	4-39	6-13	5	-	-	-				
		NGS	4-39	6-13	5	-	5-18	6	1-39 4-1†	1 2	1-17	1
EC26	CLL	Sanger Seq.	4-39	6-13	5	-	-	-				
		NGS	4-39	6-13	5	-	2-2	5	1-39 5-2†	1 1	1D-8	2
EC28	CLL	Sanger Seq.	4-39	6-13	5	-	-	-				
		NGS	4-39	6-13	5	-	3-16	4	1-39	4	1-17	1
EC20	BCP-ALL	Sanger Seq.										
		NGS	3-30	2-2	5	-	-	-	KDE-VK2-30 and IGKV7-3-IGKV2-28			
EC22	BCP-ALL	Sanger Seq.										
		NGS	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL	FAIL
EC19	BCP-ALL	Sanger Seq.										
		NGS	3-30	3-9	5	-	-	-	3-20	1	KDE-IGKJ2 and KDE-IGKV3-20	
EC17	Burkitt lymphoma	Sanger Seq.										
		NGS	3-72	6-13	4	-	3-22	4	1-5	4	KDE-VK3-11	
EC3	DLBCL	Sanger Seq.										
		NGS	1-46	6-6	4	-	-	-	3-7	4	-	-
EC5	DLBCL	Sanger Seq.										
		NGS	3-48	-	4	-	-	-	1-6	3	-	-

Blank cells: No Sanger sequencing data available. *intron RSS-Kde not detectable by NGS as no probes against this region were included in the original panel design. †More than two rearrangements no detected. CLL: chronic lymphocytic leukemia; SMZL: splenic marginal zone lymphoma; BCP-ALL: B-cell precursor acute lymphoblastic leukemia; DLBCL: diffuse large B-cell lymphoma; Sanger Seq.: Sanger sequencing; NGS: next-generation sequencing.

known translocations partners: *CRLF2* (2 BCP-ALL), *MYC* (4 Burkitt), *BCL11A* (1 CLL), *BCL2* (2 CLL and 1 DLBCL) and *CCND3* (1 SMZL). In these samples, the exact location of different breakpoints could be delineated from the sequencing reads. In addition, analysis of the reads mapping to the IGH locus showed that in seven cases (47%) the break involved an IGHJ and/or IGHD gene whereas in eight cases (53%) the break lay within the switch regions. The NGS capture approach thus provided additional information about the timing and aberrant cellular processes giving rise to the translocation, either occurring at the time of the D-J or V-DJ recombination or during class-switch recombination.[7] Identification of the specific breakpoint was possible in the majority of cases, although due to the high homology between the different switch regions it was not always possible to specifically map reads unequivocally.

In three of the 15 B-cell LPD samples (20%) only karyo-typing results were available and the location of the breakpoint on the partner chromosome as defined by NGS was mapped to chromosomes 5q14.2 (SMZL), 1q32.1 (CLL) or 16q22.1 (CLL). However, no known lymphoma-associated genes lay in proximity to these breakpoints, highlighting the potential of this approach to identify novel translocation partners and/or mechanisms of disease. Alternatively, these chromosomal alterations may represent "passenger" recombination events, where genomic rearrangements are present in the IGH loci due to a by-product of the processing of double-stranded DNA breaks by the enzymatic machinery, not resulting in an oncogenic translocation.

In the remaining B-cell LPD sample - a case with DLBCL - NGS analysis was able to detect a translocation between IGH and chromosomal location 6p25. The break occurred downstream of *EXOC2* and thus in the vicinity of *IRF4*, a constellation similar to the activating *IRF4* translocations described in germinal-center-derived LPD.[10] FISH analysis of the material using an IGH breakapart probe (Abbott Molecular, Maidenhead, UK) showed the presence of a translocation with no evidence of a *MYC, BCL2* or *BCL6* rearrangement, supporting the NGS findings and highlighting the strength of NGS capture approaches to identify novel or uncommon translocation partners and breakpoints in one single analysis.

The three T-cell LPD included in the cohort consisted of two T-ALL and one T-NHL case. The translocations in each of the cases arose from either the TRD, TRG or the TRB locus emphasizing the benefit of the NGS capture panel of being able to interrogate all IG/TR loci in samples from different diseases at the same time. In all three samples the NGS approach identified the same rearrangement as demonstrated previously by karyotyping: t(7;10)(q34;q24) involving TRBJ2-5 and *TLX1*, inv(14)(q11;q32) involving TRDD3 and *BCL11B*[11] and an inversion on chromosome 7 involving TRGV8 (7p14.2) and TRBJ2-7 (7q34).[12]

In all 14 samples with well-characterized V-(D)-J and/or D-J IG gene rearrangements by PCR and Sanger sequencing, NGS was able to detect the same IGH and IGK/IGL gene rearrangements (Table 2), including four CLL cases with hypermutated VDJ rearrangements (90% to 96% homology). All cases included in the study were diagnostic material with tumor infiltration >60% and in the remaining six samples clonal V-(D)-J rearrangement(s) were also identified by NGS, consistent with the clonal nature of the disorder; however, PCR and Sanger sequencing data were not available for these six cases for comparison due to insufficient DNA. This version of the EuroClonality-NGS panel did not include probes for the intron RSS or KDE sequence, explaining why the intron RSS-KDE rearrangement found by Sanger sequencing in a case of Burkitt lymphoma was not detected by the NGS approach. In two CLL cases, a total of three IGK locus gene rearrangements were detected by NGS in each case, raising the possibility of more than one clonal population being present, as previously described in chronic B-cell LPD.[13] Additionally, aberrant clonal rearrangements were seen that were not detected with conventional PCR-based approaches (e.g. IGKV to IGK intron), which warrant further analysis. In the three T-cell LPD cases, Sanger sequencing identified TRDV1-TRDJ1 [T-ALL with inv(14)], a TRBV5-TRBJ1-6 [T-ALL with t(7;10)] and a TRBV5-1-TRBJ2-5 (T-NHL) rearrangement, all of which were also identified in the NGS analysis. In addition, NGS reads demonstrated functional rearrangements in TRG in both T-ALL cases (TRGV11-TRGJ1 and TRGV2-TRGJ1/J2), and a non-functional TRDV1-TRDJ1 rearrangement in one of the T-ALL. No confirmatory Sanger sequencing analysis was feasible for these latter rearrangements due to insufficient DNA.

In summary, this pilot study demonstrates the ability of the EuroClonality-NGS capture approach to simultaneously detect IG/TR translocations and V-(D)-J rearrangements in diagnostic clinical specimens from a range of malignant LPD, including cases with hypermutated VDJ rearrangements. By using capture probes against the V, D and J gene regions of the TR and IG loci (with additional switch regions for IGH), clonal rearrangements and chromosomal translocations arising from these loci can be detected and at the same time the genomic breakpoint sequence involved in the rearrangements and translocations can be identified without the need for additional tests. Other technologies such as target locus amplification have also recently demostrated the ability to detect structural variants and translocations in cancer.[14] An important advantage of these approaches lies in the fact that no prior knowledge of the translocation partner is needed and, therefore, novel or rare chromosomal rearrangements can also be identified by this method, improving their diagnostic value. Sequencing of the V-(D)-J gene rearrangements in any of the IG/TR loci can be used not only to assess clonality and enable a more in-depth analysis of clonal relationships and clonal evolution, but also to identify targets for minimal residual disease monitoring and analysis of the IG/TR repertoire of diverse lymphoid populations. Additional information, for example the somatic hypermutation status of the IGHV-IGHD-IGHJ gene rearrangements, relevant for prognosis in CLL[6] can also be obtained. A new version of this EuroClonality-NGS panel is being designed to include common non-IG/TR translocations as well as genes relevant for diagnosis and prognosis in LPD and a clinical multicenter validation study is now underway within the EuroClonality-NGS consortium.

Dörte Wren,[1]* Brian A. Walker,[1,2]* Monika Brüggemann,[3] Mark A. Catherwood,[4] Christiane Pott,[3] Kostas Stamatopoulos,[5] Anton W. Langerak[6]† and David Gonzalez[1,7]† (on behalf of the EuroClonality-NGS consortium)

[1]*The Centre for Molecular Pathology, The Royal Marsden NHS Foundation Trust, London, UK;* [2]*Myeloma Institute, University of Arkansas for Medical Sciences, Little Rock, AR, USA;* [3]*Second Medical Department, University Hospital Schleswig-Holstein,*

Campus Kiel, Germany; ⁴Department of Haematology, Belfast City Hospital, UK; ⁵Institute of Applied Biosciences, CERTH, Thessaloniki, Greece; ⁶Deptartment of Immunology, Laboratory for Medical Immunology, Erasmus MC, University Medical Center Rotterdam, the Netherlands and ⁷Centre for Cancer Research and Cell Biology, Queen's University Belfast, UK

Correspondence: d.gonzalezdecastro@qub.ac.uk
doi:10.3324/haematol.2016.155424

These authors contributed equally to this work. †These authors contributed equally to this work and share senior authorship of this article.

References

1. van Dongen JJ, Langerak AW, Bruggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. Leukemia. 2003;17(12):2257-2317.
2. Langerak AW, Groenen PJ, Bruggemann M, et al. EuroClonality/BIOMED-2 guidelines for interpretation and reporting of Ig/TCR clonality testing in suspected lymphoproliferations. Leukemia. 2012;26(10):2159-2171.
3. Bruggemann M, White H, Gaulard P, et al. Powerful strategy for polymerase chain reaction-based clonality assessment in T-cell malignancies. Report of the BIOMED-2 Concerted Action BHM4 CT98-3936. Leukemia. 2007;21(2):215-221.
4. Evans PA, Pott C, Groenen PJ, et al. Significantly improved PCR-based clonality testing in B-cell malignancies by use of multiple immunoglobulin gene targets. Report of the BIOMED-2 Concerted Action BHM4-CT98-3936. Leukemia. 2007;21(2):207-214.
5. Bruggemann M, Schrauder A, Raff T, et al. Standardized MRD quantification in European ALL trials: proceedings of the Second International Symposium on MRD assessment in Kiel, Germany, 18-20 September 2008. Leukemia. 2010;24(3):521-535.
6. Vardi A, Agathangelidis A, Sutton LA, Ghia P, Rosenquist R, Stamatopoulos K. Immunogenetic studies of chronic lymphocytic leukemia: revelations and speculations about ontogeny and clinical evolution. Cancer Res. 2014;74(16):4211-4216.
7. Walker BA, Wardell CP, Johnson DC, et al. Characterization of IGH locus breakpoints in multiple myeloma indicates a subset of translocations appear to occur in pregerminal center B cells. Blood. 2013;121(17):3413-3419.
8. Bouamar H, Abbas S, Lin AP, et al. A capture-sequencing strategy identifies IRF8, EBF1, and APRIL as novel IGH fusion partners in B-cell lymphoma. Blood. 2013;122(5):726-733.
9. Walker BA, Boyle EM, Wardell CP, et al. Mutational spectrum, copy number changes, and outcome: results of a sequencing study of patients with newly diagnosed myeloma. 2015;33(33):3911-3920.
10. Salaverria I, Philipp C, Oschlies I, et al. Translocations activating IRF4 identify a subtype of germinal center-derived B-cell lymphoma affecting predominantly children and young adults. Blood. 2011;118(1):139-147.
11. Przybylski GK, Dik WA, Wanzeck J, et al. Disruption of the BCL11B gene through inv(14)(q11.2q32.31) results in the expression of BCL11B-TRDC fusion transcripts and is associated with the absence of wild-type BCL11B transcripts in T-ALL. Leukemia. 2005;19(2):201-208.
12. Stern MH, Lipkowitz S, Aurias A, Griscelli C, Thomas G, Kirsch IR. Inversion of chromosome 7 in ataxia telangiectasia is generated by a rearrangement between T-cell receptor beta and T-cell receptor gamma genes. Blood. 1989;74(6):2076-2080.
13. Sanchez ML, Almeida J, Gonzalez D, et al. Incidence and clinicobiologic characteristics of leukemic B-cell chronic lymphoproliferative disorders with more than one B-cell clone. Blood. 2003;102(8):2994-3002.
14. de Vree PJ, de Wit E, Yilmaz M, et al. Targeted sequencing by proximity ligation for comprehensive variant detection and local haplotyping. Nat Biotechnol. 2014;32(10):1019-1025.