Genetic evolution of in situ follicular neoplasia to aggressive B-cell lymphoma of germinal center subtype

Antonio Vogelsberg,1 Julia Steinhilber,1 Barbara Mankel,1 Birgit Federmann,1 Janine Schmidt,1 Ivonne A. Montes-Mojarrio,1 Katrin Hüttl,2 Maria Rodríguez-Pinilla,3 Praveen Baskaran,4 Sven Nahnsen,4 Miguel A. Piris,3 German Ott,2 Leticia Quintanilla-Martínez,1 Irina Bonzheim,1# and Falko Fend1#

1Institute of Pathology and Neuropathology, University Hospital and Comprehensive Cancer Center Tübingen, Tübingen, Germany; 2Department of Clinical Pathology, Robert-Bosch-Krankenhaus, and Dr. Margarete Fischer-Bosch Institute of Clinical Pharmacology, Stuttgart, Germany; 3Department of Pathology, Fundación Jiménez Díaz, Madrid, Spain and 4Quantitative Biology Center, University of Tübingen, Tübingen, Germany

#IB and FF contributed equally as co-senior authors.

©2021 Ferrata Storti Foundation. This is an open-access paper. doi:10.3324/haematol.2020.254854

Received: April 8, 2020.
Accepted: August 12, 2020.
Pre-published: August 27, 2020.

Correspondence: FALKO FEND - falko.fend@med.uni-tuebingen.de
IRINA BONZHEIM - irina.bonzheim@med.uni-tuebingen.de
Supplementary methods

Diagnosis of in situ follicular neoplasia (ISFN)

The diagnosis of ISFN was based on the criteria published in the update of the 4th Edition of the WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Specifically, a diagnosis of ISFN was made when (1) the lymph node architecture was completely preserved, with normally sized follicles, and routine hematoxylin and eosin (H&E) stains gave no evidence of FL involvement, (2) all germinal centers involved by ISFN, as evidenced by strongly BCL2+ and CD10+ centrocytes, had a clearly preserved and well-delineated mantle zone, and (3) BCL2 and CD10 stains failed to show any indication of extrafollicular spread of ISFN cells.

Microdissection and DNA isolation

Laser microdissection of ISFN samples was performed from 10 to 20 serial H&E sections with the first and every sixth slide stained for BCL2 to localize the ISFN lesions. After microdissection of between 13 and 65 germinal centers per slide, the tissue was pooled and digested with proteinase K (Merck, Darmstadt, Germany) and DNA extracted applying standard phenol/chloroform purification procedures. If macrodissection of paraffin sections was performed, DNA was extracted using the Maxwell 16 MDx Instrument (Promega, Mannheim, Germany) according to the manufacturer’s instructions.

Immunohistochemistry and fluorescence in situ hybridization (FISH)

Immunohistochemistry was performed on an automated immunostainer (Ventana Medical Systems, Tucson, AZ, USA). All ISFN samples were stained for BCL2, CD20, MIB1 (DAKO,
Hamburg, Germany) and CD10 (Novocastra, Wetzlar, Germany). High-grade B-cell lymphoma (HGBL), diffuse large B-cell lymphoma (DLBCL) and follicular lymphoma (FL) samples were additionally stained for CD3 (DCS, Hamburg, Germany), BCL6 (Zytomed, Berlin, Germany), MUM1 (DAKO), MYC (Roche, Penzberg, Germany) and P53 (Novacastra). All aggressive B-cell lymphomas (BCLs) were sub-classified according to the Hans algorithm. Images were taken with the Axioskop 2 Plus microscope (Zeiss, Oberkochen, Germany) and the Jenoptik ProgRes C10 Plus camera and software (Jenoptik, Jena, Germany). Objectives used were Plan-Neofluar x1.25/0.035, x2.5/0.075, x10/0.30, x20/0.50 and x40/0.75 (Zeiss). FISH analysis was performed using Vysis LSI BCL2, LSI BCL6 and LSI MYC Dual Color Break Apart Rearrangement Probes (Abbott Molecular, Wiesbaden, Germany) for the detection of BCL2, MYC and BCL6 translocations, respectively. Case 5 was also analyzed with the Vysis LSI IGH/BCL2 Dual Color Dual Fusion Translocation Probe (Abbott Molecular) and the ZytoLight SPEC IGH Dual Color Break Apart Probe (ZytoVision, Bremershaven, Germany). For ISFN lesions, FISH for BCL6 and MYC was performed only if the respective rearrangement had been detected in the paired aggressive BCL. Samples that carried a TP53 mutation were analyzed with the Vysis LSI TP53 SpectrumOrange/CEP 17 SpectrumGreen Probe (Abbott Molecular) to investigate a loss of the second TP53 allele.

**PCR and Sanger sequencing of the t(14;18) breakpoint region**

Forward primers used were 5’ TTAGAGAGTTGGCTTTACGTGGCCTG 3’ for the major breakpoint region (MBR)\(^4\), 5’ TCGTTTCTCAGTAAGTGAGAGTGC 3’ for the intermediate cluster region (ICR)\(^5\) and 5’ CGTGCTGGTACCACCTCCTG 3’ for the minor cluster region (MCR)\(^6\) as well as eight additional primers that cover a region of about 1 kilobase downstream of the MCR primer. The joining region consensus primer 5’ CTTACCTGAGGAGACGGTGACC 3’ was used as the reverse primer.\(^7\) PCR was performed with 100 ng of purified DNA in a final volume of 25 µl using 0.4 mM dNTPs, 1.5 mM MgCl\(_2\), 0.4 µM of each primer and 1.25 U Taq polymerase (AmpliTaq Gold DNA Polymerase; Applied Biosystems, Foster City, CA, USA). Cycling involved an initial denaturation at 95°C for 5 min followed by 45 cycles of denaturation (95°C for 45 s), annealing (60 s at 57°C for
MBR and MCR, 60 s at 56°C for the ICR) and elongation (72°C for 60 s), with a final elongation at 72°C for 10 min. To increase the detection rate, we additionally used the IdentiClone BCL2/JH Translocation Assay, which was performed according to the manufacturer’s instructions (Invivoscribe, San Diego, CA, USA). PCR products were purified using Agencourt AMPure XP magnetic beads (Beckman Coulter, Brea, CA, USA) and mixed with 1 µl of the respective primer (10 μM) and 2 µl of the GenomeLab DTCS-Quick Start Kit (Beckman Coulter) to a final volume of 10 µl for the sequencing reaction according to the manufacturer’s protocol. Sequencing reactions were purified (CleanSEQ; Beckman Coulter), analyzed in a GenomeLab GeXP Genetic Analysis System (Beckman Coulter) and evaluated by the GenomeLab GeXP software 11.0 (Beckman Coulter) to investigate the t(14;18) breakpoint sequence.

For the ISFN samples of cases 3 and 9, primers specific to the breakpoint of the corresponding aggressive BCL were designed using the Primer3web software 4.1.0 (http://primer3.ut.ee/), with primers binding to the respective BCL2 and t(14;18) de novo sequences.8 Forward BCL2 primers used were 5’ AACACAGACCCACCCAGAG 3’ (Case 3) and 5’ GCTTTCTCATGGCTGTCCTT 3’ (Case 9). Reverse de novo sequence primers used were 5’ ATACCGTACGTCCGAAAGCA 3’ (Case 3) and 5’ GGGACCACATCGAGAAGC 3’ (Case 9). PCR was performed with 100 ng of genomic DNA and modified annealing temperatures (54°C and 53°C). A successful amplification in the respective ISFN lesion was seen as evidence of the same t(14;18) breakpoint. Primer specificity was ensured using clonally unrelated t(14;18)+ samples as negative controls. Additionally, all PCR products were sequenced as stated above.

Clonality analysis

PCRs for the detection of immunoglobulin gene rearrangements were performed in duplicate with two different concentrations of genomic DNA using 1 U AmpliTaq Gold DNA polymerase (Applied Biosystems) and BIOMED-2 FR2, FR3, JH, and Vk, Jk, IntronRSS and Kde primers.7 Modified amplification conditions were carried out with an initial denaturation step of 95°C (7 min), 40 cycles (95°C for 30 s, 60°C for 45 s, 72°C for 45 s) and a final step of 72°C
for 4 min. The JH, JK1-4, JK5 and Kde primers were modified with D4 fluorescent dyes (Sigma-Aldrich, St. Louis, MO, USA). For cases 6 and 10, clone-specific primers were designed based on the respective DLBCL framework region and complementarity-determining region 3 using the Primer3web software 4.1.0 (http://primer3.ut.ee/). Forward primers used were 5’ GAATATGCTGCGTCGGTGAA 3’ (Case 6) and 5’ ATGGAGTTGAGGAGGCTGAC 3’ (Case 10). Reverse primers used were 5’ TGTGGCTACGGACCTCTCTA 3’ (Case 6) and 5’ GCCCCAGACGTCCATAACAT 3’ (Case 10). Reverse primers were modified with D4 fluorescent dyes (Sigma-Aldrich) and PCR was performed with 100 ng of genomic DNA and modified annealing temperatures (54°C and 53°C). For GeneScan analysis 1 µl of the PCR products were mixed with sample loading solution containing 30 µl DNA Size Standard 400 (Beckman Coulter). The products were separated by capillary electrophoresis on the GenomeLab GeXP Genetic Analysis System and analyzed by the GenomeLab GeXP software 11.0 (Beckman Coulter).

**Immunoglobulin sequence analysis**

Next generation sequencing (NGS) of the immunoglobulin genes was performed with the LymphoTrack Dx IGH FR1, FR2 and FR3 Assay – PGM (Invivoscribe) according to the manufacturer’s instructions. Libraries were purified and quantified applying Agencourt AMPure XP (Beckman Coulter) magnetic beads and the Ion Library Quantitation Kit (Thermo Fisher Scientific, Waltham, MA, USA) on the LightCycler 480 real-time PCR system (Roche Molecular Systems, Pleasanton, CA, USA). Generated libraries were run on the Ion Torrent Personal Genome Machine (PGM; Thermo Fisher Scientific). NGS data were analyzed with the LymphoTrack Dx Software – PGM (Invivoscribe) and interpreted according to the manufacturer’s protocol, which allows the detection of clonal immunoglobulin rearrangements with variable and joining gene usage and sequence information. Clonal sequences were submitted to IMGT/V-QUEST (http://www.imgt.org/ IMGT_vquest/vquest) and IgBlast (https://www.ncbi.nlm.nih.gov/igblast/) for analysis. N-glycosylation motifs were identified by the consensus sequence Asn-X-Ser/Thr, where X can be any amino acid except proline. To investigate intraclonal heterogeneity, the ten most prevalent clone-
specific sequences (i.e. subclones) of each sample were identified through the alignment of their sequence with that of the respective dominant rearrangement. Clear-cut sequencing artifacts, i.e. insertions/deletions (InDels) in homopolymer regions, InDels at the beginning of a sequence and changes of the first nucleotide were manually corrected to the sequence of the major clone. To calculate the share of each subclone, the sequence count was divided by the total number of clone-specific reads.

**Phylogenetic tree construction**

Phylogenetic trees for cases 1, 2, 4, 7, and 9 were built using the ten most prevalent subclones of the respective samples and the corresponding unmutated VDJ germline sequence, which was determined with IMGT/V-QUEST ([http://www.imgt.org/IMGT_vquest/vquest](http://www.imgt.org/IMGT_vquest/vquest)) and IgBlast ([https://www.ncbi.nlm.nih.gov/igblast/](https://www.ncbi.nlm.nih.gov/igblast/)). Multiple sequence alignments were generated using MAFFT (Version 7.4) with localpair alignment mode and max iteration of 1000. jModelTest (Version 2.1) was used to find the best-fit substitution model for each multiple sequence alignment based on Bayesian information criteria strategy. In summary, JC69 (Case 2), K80 + I (Case 4), and K80 (Cases 1, 7, and 9) were determined as most suitable. The construction of phylogenetic trees was done using the maximum likelihood method implemented in RAxML (Version 8.2). The corresponding plots were generated in R (Version 3.4) ([http://www.R-project.org/](http://www.R-project.org/)) using the “ape” and “phytools” packages.

**Library preparation and sequencing**

Amplicon library preparation and semiconductor sequencing were performed according to the manufacturer’s instructions (Thermo Fisher Scientific). For each reaction, 10 ng of DNA were mixed with AmpliSeq HiFi Mix (Thermo Fisher Scientific) and the respective primer pool to amplify the target regions. Subsequently, primer end sequences were partially digested using FuPa reagent (Thermo Fisher Scientific), followed by the ligation of barcoded sequencing adapters (Ion Xpress Barcode Adapters; Thermo Fisher Scientific). The final libraries were purified and quantified as described in “Immunoglobulin sequence analysis”.
Libraries were diluted to 100 pM each and pooled. In the next step, DNA fragments were attached to Ion Sphere Particles (ISPs) and clonally amplified using the Ion PGM Hi-Q OT2 Kit (Thermo Fisher Scientific) and the Ion OneTouch Instrument (Thermo Fisher Scientific). The amount of template-positive ISPs was determined with the Qubit 3.0 Fluorometer (Life Technologies, Darmstadt, Germany) and the Ion Sphere Quality Control Kit (Thermo Fisher Scientific). Afterwards, the Ion OneTouch ES (Thermo Fisher Scientific) was used to enrich template-positive ISPs. In a last step, sequencing primers were attached to the DNA fragments bound to the ISPs, which were subsequently loaded on a semiconductor chip (Ion 318 Chip Kit; Thermo Fisher Scientific). Finally, sequencing was performed using the Ion PGM Hi-Q Sequencing Kit and the Ion Torrent PGM platform (Thermo Fisher Scientific).

**Targeted NGS data analysis**

Detection of variants in comparison to the human reference sequence (hg19) was performed using the Torrent Suite (Version 5.6.0) and the Ion Torrent Variant Caller (5.8.0.19) (Thermo Fisher Scientific). Detection thresholds were set at an allele frequency of 5%. Variants were annotated and filtered against the dbSNP and COSMIC databases using the Annotate variants single sample workflow of the Ion Reporter Software (Version 5.6) (Thermo Fisher Scientific). The Integrative Genomics Viewer (Version 2.3.94) (Broad Institute, Cambridge, MA, USA) software was used to inspect each detected variant to exclude possible artifacts. Variants considered to be artifacts were those only detected in one sequencing direction and InDels at sites of homopolymer regions. Caution was also exercised when variants occurred in regions with low coverage, especially concerning CG>TA transitions and/or alterations with VAFs <10%. All sequences that harbored an alteration in at least one sample of a case were specifically reviewed in paired samples, even when not called by the Ion Reporter Software. If the mutation could not be detected in a paired sample, but the coverage was low (<100 reads), the location was reevaluated with bidirectional single amplicon sequencing to avoid a false negative result. Prediction of the deleteriousness of variants was done using the Combined Annotation Dependent Depletion (CADD) predictor (http://cadd.gs.washington.edu/home). For the construction of clonal evolution patterns,
synonymous and 5’ untranslated region (5’UTR) mutations of BCL2 were taken into account as additional markers (Supplementary Table S5). Sequencing data are deposited in the European Nucleotide Archive (Accession number PRJEB34446).

Variant validation and single amplicon sequencing

To further exclude sequencing artifacts, the majority of variants were validated (see Supplementary Table S4). If TP53 was mutated in an aggressive BCL, single amplicon sequencing was used to investigate if the mutation could be detected in paired ISFN and FL samples. Single amplicons were prepared following the Ion Amplicon Library Preparation Fusion Method protocol (Thermo Fisher Scientific). Primers were designed using the primer3 software 4.1.0 (http://primer3.ut.ee). The primers were composed of either the A adapter or the trP1 adapter, the barcode sequence and barcode adapter sequence, and the target primer sequence (Supplementary Table S2). Each gene region was amplified using two primer pairs (A Forward and trP1 Reverse or A Reverse and trP1 Forward) to enable bidirectional sequencing. Library preparation was done according to the manufacture’s protocol (Thermo Fisher Scientific).

References

1. Swerdlow SH, Campo E, Harris NL, et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Revised 4th ed. Lyon: IARC Press, 2017.
2. Schmidt J, Salaverria I, Haake A, et al. Increasing genomic and epigenomic complexity in the clonal evolution from in situ to manifest t(14;18)-positive follicular lymphoma. Leukemia. 2014;28(5):1103-1112.
3. Hans CP, Weisenburger DD, Greiner TC, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood. 2004;103(1):275-282.
4. Stetler-Stevenson M, Raffeld M, Cohen P, Cossman J. Detection of occult follicular lymphoma by specific DNA amplification. Blood. 1988;72(5):1822-1825.
5. Albinger-Hegyi A, Hochreutener B, Abdou MT, et al. High frequency of t(14;18)-translocation breakpoints outside of major breakpoint and minor cluster regions in follicular lymphomas: improved polymerase chain reaction protocols for their detection. Am J Pathol. 2002;160(3):823-832.
6. Gribben JG, Freedman A, Woo SD, et al. All advanced stage non-Hodgkin's lymphomas with a polymerase chain reaction amplifiable breakpoint of bcl-2 have residual cells containing the bcl-2 rearrangement at evaluation and after treatment. Blood. 1991;78(12):3275-3280.

7. 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.

8. Untergasser A, Cutcutache I, Koressaar T, et al. Primer3--new capabilities and interfaces. Nucleic Acids Res. 2012;40(15):e115.

9. Brochet X, Lefranc MP, Giudicelli V. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. Nucleic Acids Res. 2008;36(Web Server issue):W503-508.

10. Ye J, Ma N, Madden TL, Ostell JM. IgBLAST: an immunoglobulin variable domain sequence analysis tool. Nucleic Acids Res. 2013;41(Web Server issue):W34-40.

11. Berg DT, Grinnell BW. Pro to Gly (P219G) in a silent glycosylation site results in complete glycosylation in tissue plasminogen activator. Protein Sci. 1993;2(1):126-127.

12. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772-780.

13. Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: more models, new heuristics and parallel computing. Nature methods. 2012;9(8):772.

14. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30(9):1312-1313.

15. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

16. Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics. 2019;35(3):526-528.

17. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol Evol. 2012;3(2):217-223.

18. Robinson JT, Thorvaldsdottir H, Winckler W, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24-26.

19. Mardis ER. Next-generation sequencing platforms. Annu Rev Anal Chem (Palo Alto Calif). 2013;6:287-303.

20. Wong SQ, Li J, Tan AY, et al. Sequence artefacts in a prospective series of formalin-fixed tumours tested for mutations in hotspot regions by massively parallel sequencing. BMC Med Genomics. 2014;7:23.
Supplementary tables

Supplementary Table S1. Genes analyzed with AmpliSeq Custom Panels.

| Gene   | Position (GRCh37/hg19) | Exon(s) | Amplicons |
|--------|------------------------|---------|-----------|
| BCL2   | chr18: 60,795,858 - 60,985,965 | CDS | 9         |
| BCL6   | chr3: 187,440,246 - 187,451,481 | CDS | 27        |
| BTG1   | chr12: 92,537,856 - 92,539,311 | CDS | 7         |
| BTG2   | chr1: 203,274,735 - 203,276,566 | CDS | 6         |
| CARD11 | chr7: 2,946,272 - 2,998,140 | CDS | 54        |
| CD79B  | chr17: 62,006,586 - 62,009,621 | CDS | 11        |
| CREBBP | chr16: 3,777,719 - 3,929,917 | CDS | 96        |
| EP300  | chr22: 41,489,009 - 41,574,960 | CDS | 63        |
| EZH2   | chr7: 148,508,712 - 148,508,817 | CDS | 16        |
| FOXO1  | chr13: 41,133,660 - 41,240,349 | CDS | 10        |
| GNA13  | chr17: 63,010,375 - 63,052,711 | CDS | 8         |
| HIST1H1B | chr6: 27,834,627 - 27,835,307 | CDS | 6         |
| HIST1H1C | chr6: 26,056,015 - 26,056,656 | CDS | 6         |
| HIST1H1D | chr6: 26,234,496 - 26,235,161 | CDS | 6         |
| HIST1H1E | chr6: 26,156,619 - 26,157,278 | CDS | 5         |
| IL7R   | chr22: 23,230,234 - 23,237,874 | CDS | 8         |
| KMT2D  | chr12: 49,415,563 - 49,449,107 | CDS | 120       |
| IRF4   | chr6: 393,153 - 407,598 | CDS | 18        |
| MEF2B  | chr19: 19,256,503 - 19,261,544 | CDS | 11        |
| MYD88  | chr3: 38,181,350 - 38,182,777 | CDS | 2-5       |
| PIM1   | chr6: 37,138,079 - 37,141,867 | CDS | 16        |
| PRDM1  | chr6: 106,534,429 - 106,555,361 | CDS | 29        |
| TBL1XR1 | chr3: 176,743,286 - 176,782,765 | CDS | 32        |
| TNFAIP3 | chr3: 138,192,365 - 138,202,456 | CDS | 29        |
| TNFRSF14 | chr1: 2,488,104 - 2,494,712 | CDS | 11        |

CDS, coding sequence.

Supplementary Table S2. Primer sequences for targeted resequencing including the sequences of the A or trP1 adapter and the barcodes.

| Primer          | Sequence 5’-3’ |
|-----------------|---------------|
| GNA13 Ex4 326 BC50 AF | CCAATCATCTCCGCGTGTCCTCGACTCACGGCGCAATGCGGATCCCACCTGCTTAAGAGACG |
| GNA13 Ex4 326 BP1F | CCAATCATGGCCAGTCTCGGATCCCACCTGCTTAAGAGACG |
| GNA13 Ex4 326 AR1F | CCAATCATCCGCGTGTCCTCGACTCACGGCGCAATGCGGATCCCACCTGCTTAAGAGACG |
| GNA13 Ex4 326 AR1R | CCAATCATGGCCAGTCTCGGATCCCACCTGCTTAAGAGACG |
| TP53 Ex8 273 BC51 AF | CCAATCATCTCCGCGTGTCCTCGACTCACGGCGCAATGCGGATCCCACCTGCTTAAGAGACG |
| TP53 Ex8 273 AR1F | CCAATCATGGCCAGTCTCGGATCCCACCTGCTTAAGAGACG |
| TP53 Ex8 273 AR1R | CCAATCATCTCCGCGTGTCCTCGACTCACGGCGCAATGCGGATCCCACCTGCTTAAGAGACG |
| CREBBP Ex7 551 BC52 AF | CCAATCATCTCCGCGTGTCCTCGACTCACGGCGCAATGCGGATCCCACCTGCTTAAGAGACG |
| CREBBP Ex7 551 AR1F | CCAATCATCTCCGCGTGTCCTCGACTCACGGCGCAATGCGGATCCCACCTGCTTAAGAGACG |
| CREBBP Ex7 551 AR1R | CCAATCATCTCCGCGTGTCCTCGACTCACGGCGCAATGCGGATCCCACCTGCTTAAGAGACG |
| CREBBP Ex30 1680 BC60 AF | CCAATCATCTCCGCGTGTCCTCGACTCACGGCGCAATGCGGATCCCACCTGCTTAAGAGACG |
| CREBBP Ex30 1680 AR1F | CCAATCATCTCCGCGTGTCCTCGACTCACGGCGCAATGCGGATCCCACCTGCTTAAGAGACG |
| CREBBP Ex30 1680 AR1R | CCAATCATCTCCGCGTGTCCTCGACTCACGGCGCAATGCGGATCCCACCTGCTTAAGAGACG |
| KMT2D Ex31 2823 BC61 AF | CCAATCATCTCCGCGTGTCCTCGACTCACGGCGCAATGCGGATCCCACCTGCTTAAGAGACG |
Supplementary Table S3. Immunohistochemical findings of FL and aggressive BCL samples.

| Case | Diagnosis     | CD10 | BCL6 | MUM1 | BCL2 | MIB-1 (%) | CD20 | P53* | MYC* (%) |
|------|---------------|------|------|------|------|-----------|------|------|----------|
|      |               |      |      |      |      |           |      |      |          |
| De novo aggressive B-cell lymphoma |               |      |      |      |      |           |      |      |          |
| 1    | HGBL-TH       | +    | +    | -    | +    | 70        | +    | +    | 40 (w)   |
| 2    | DLBCL         | +    | +    | -    | +    | 40        | +    | -    |          |
| 3    | DLBCL         | +    | +    | -    | +    | 70        | +    | +    |          |
| 4    | DLBCL         | +    | +    | -    | +    | 70        | +    | -    | 30 (h)   |
| 5    | DLBCL         | +    | +    | -    | +    | 40        | +    | -    | 5 (s)    |
| 6    | DLBCL         | +    | +    | -    | +    | 30        | +    | -    |          |
| Transformed FL |               |      |      |      |      |           |      |      |          |
| 7    | FL            | +    | +    | -    | +    | 10        | +    | -    |          |
|      | DLBCL         | +    | +    | +    | +    | 90        | +    | -    | 30 (s)   |
| 8    | HGBL-DH       | +    | +    | -    | +    | 80        | +    | -    | 60 (s)   |
| 9    | FL            | +    | +    | -    | +    | 5         | +    | -    |          |
|      | HGBL-DH       | +    | +    | -    | +    | 70        | +    | -    | 40 (h)   |
| 10   | FL            | +    | +    | -    | +    | N/A       | +    | -    |          |
|      | DLBCL         | +    | +    | -    | +    | 50        | +    | -    | 15 (s)   |

DH, Double-hit; N/A, Not available; TH, Triple-hit. *Only samples with a strong staining of ≥ 20% of neoplastic cells were considered positive. #Percentages represent the share of positive lymphoma cells with strong (s), heterogeneous (h) or weak (w) staining. °Complete loss in the neoplastic cells.
## Supplementary Table S4. Overview of non-synonymous and splice site mutations.

| Case | Diagnosis | Gene    | Transcript | Predicted protein change | cDNA change | VAF (%) | Coverage | Validation  | CADD Score' |
|------|-----------|---------|------------|--------------------------|-------------|---------|----------|-------------|-------------|
| 1    | ISFN      | BCL2    | NM_00633   | p.D31N                   | c.91G>A     | 15      | 3210     | Confirmed   | 20.0        |
|      | HGBL-TH   | BCL2    | NM_00633   | p.T166S                  | c.166A>T    | 42      | 2427     | ND          | 10.65       |
|      | BCL2      | NM_00633 | p.L369S    | c.369delinsTC            | ND          | 45      | 2426     | ND          | 21.6        |
|      | TP53      | NM_000546| p.R273C    | c.817C>T                 | 52          | 8697    | Confirmed | 25.3        |
|      | GNA13     | NM_006572| p.L320R    | c.977T>G                 | 8           | 2322    | Confirmed | 32          |
| 2    | ISFN      | BCL2    | NM_00633   | p.Q33R                   | 8           | 3266    | Confirmed | 15.30       |
|      | BCL2      | NM_00633 | p.G101A    | c.302G>C                 | 7           | 7474    | Confirmed | 25.1        |
|      | CREBBP    | NM_004380| p.S1680del | c.5039_5041del           | 28          | 588     | Confirmed | 22.7        |
|      | DLBCL     | BCL2    | NM_00633   | p.R65G                   | 20          | 2283    | Confirmed | 21.2        |
|      | BCL2      | NM_00633 | p.G33R     | c.164A>G                 | 21          | 3097    | Confirmed | 15.30       |
|      | BCL2      | NM_00633 | p.A113G    | c.338C>G                 | 22          | 6631    | Confirmed | 13.61       |
|      | CREBBP    | NM_004380| p.S1680del | c.5039_5041del           | 67          | 565     | Confirmed | 22.7        |
|      | TP53      | NM_000546| p.T150fs   | c.447_450del             | 52          | 9399    | ND        | 28.7        |
|      | CARD11    | NM_032415| p.Q249P    | c.746A>C                 | 46          | 7624    | Confirmed | 27.1        |
| 3    | ISFN      | CREBBP  | NM_004380  | p.Y1503D                 | c.4507T>G   | 16      | 1673     | Confirmed   | 29.4        |
|      | BCL2      | NM_00633 | p.Y1503D   | c.4507T>G                | 15          | 1171    | ND        | 26.3        |
|      | BCL2      | NM_00633 | p.I545fs   | c.16365_1637del          | 51          | 32774   | ND        | 32          |
|      | CREBBP    | NM_004380| p.Y1503D   | c.4507T>G                | 26          | 8807    | ND        | 36          |
|      | TP53      | NM_000546| p.H179N    | c.4767del                | 45          | 1983    | ND        | 26.3        |
|      | BCL2      | NM_00633 | p.W214C    | c.642G>T                 | 60          | 14102   | ND        | 32          |
|      | KMT2D     | NM_003482| p.I455fs   | c.16365_1637del          | 47          | 4421    | ND        | 36          |
|      | CREBBP    | NM_004380| p.Y1503D   | c.4507T>G                | 21          | 1385    | Confirmed | 29.4        |
|      | BCL2      | NM_00633 | p.Y1503D   | c.4507T>G                | 42          | 1187    | ND        | 26.3        |
|      | CREBBP    | NM_004380| p.N1589fs  | c.4767del                | 84          | 3049    | ND        | 28.2        |
|      | TP53      | NM_000546| p.H179N    | c.535C>A                 | 83          | 3576    | ND        | 28.2        |
| 4    | ISFN      | BCL2    | NM_00633   | p.A76D                   | c.227C>A    | 44      | 1015     | Confirmed   | 13.25       |
|      | TNFRSF14  | NM_003820| p.W12*     | c.35G>A                  | 25          | 7259    | Confirmed | 35          |
|      | HIST1H1D  | NM_005320| p.N77K     | c.231C>G                 | 31          | 6107    | Confirmed | 24.8        |
|      | EP300     | NM_001429| p.L415P    | c.1244T>C                | 24          | 841     | Confirmed | 24.8        |
|      | BCL2      | NM_00633 | p.P95S     | c.175C>T                 | 24          | 5580    | ND        | 13.71       |
|      | KMT2D     | NM_003482| p.A76D     | c.227C>A                 | 23          | 5649    | Confirmed | 13.25       |
|      | HIST1H1D  | NM_005320| p.N77K     | c.231C>G                 | 13          | 8530    | Confirmed | 24.8        |
|      | EP300     | NM_001429| p.L415P    | c.1244T>C                | 19          | 1200    | Confirmed | 24.8        |
|      | BCL2      | NM_00633 | p.P95S     | c.175C>T                 | 24          | 5580    | ND        | 13.71       |
|      | KMT2D     | NM_003482| p.A76D     | c.227C>A                 | 23          | 5649    | Confirmed | 13.25       |
|      | TNFRSF14  | NM_003820| p.W12*     | c.35G>A                  | 14          | 182     | Confirmed | 35          |
|      | HIST1H1D  | NM_005320| p.N77K     | c.231C>G                 | 13          | 8530    | Confirmed | 24.8        |
|      | EP300     | NM_001429| p.L415P    | c.1244T>C                | 19          | 1200    | Confirmed | 24.8        |
| 5    | ISFN      | BCL2    | NM_00633   | p.P53A                   | c.157C>G    | 28      | 2009     | 10.80       |
|      | BCL2      | NM_00633 | p.R129C    | c.385C>T                 | 13          | 7370    | Confirmed | 24.3        |
|      | BCL2      | NM_00633 | p.F153S    | c.457T>C                 | 8           | 8140    | Confirmed | 32          |
|      | KMT2D     | NM_003482| p.Q4473*   | c.13417C>T               | 24          | 702     | Confirmed | 43          |
|      | CREBBP    | NM_004380| p.V1371D   | c.4112T>A                | 30          | 956     | ND        | 29.4        |
|      | IGLL5     | NM_001250296| p.C35     | c.8G>C                   | 25          | 1331    | ND        | 0.018       |
|   |   |   |   |   |   |
|---|---|---|---|---|---|
| 5 | DLBCL |   |   |   |   |
|   | BCL2 | NM_000633 | p.P598S | c.175C>T | 51 | 2893 | ND | 13.71 |
|   | BCL2 | NM_000633 | p.A282T | c.244G>A | 50 | 3013 | ND | 12.65 |
|   | BCL2 | NM_000633 | p.D102G | c.307G>T | 40 | 8806 | ND | 26.2 |
|   | KMT2D | NM_003482 | p.Q4473* | c.13417C>T | 30 | 4234 | Confirmed | 43 |
|   | EZH2 | NM_004456 | p.Y646F | c.1927A>T &   | 27 | 10985 | ND | 25.3 |
|   | IGLL5 | NM_001256296 | p.G132S | c.396G>T | 54 | 1168 | ND | 0.18 |
|   | GNA13 | NM_006572 | p.Q476R | c.1421T>C | 16 | 6861 | Confirmed | 29.4 |
|   | GNA13 | NM_006572 | p.D155A | c.464A>C | 35 | 9386 | ND | 27.7 |
|   | GNA13 | NM_006572 | p.T203A | c.609G>T | 38 | 118 | Confirmed | 23.8 |
|   | HIST1H1D | NM_005320 | p.T93S | c.278G>C | 45 | 21464 | ND | 23.3 |
|   | MEF2B | NM_001145785 | p.R38M | c.115G>T | 36 | 13251 | ND | 26.1 |
|   | CD79B | NM_000626 | p.Y196H | c.589A>C | 38 | 5105 | Confirmed | 24.8 |

|   |   |   |   |   |   |
| 6 | ISFN |   |   |   |   |
|   | BCL2 | NM_000633 | p.G5V | c.14G>T | 14 | 3104 | Confirmed | 25.5 |
|   | BCL2 | NM_000633 | p.A24V | c.125C>T | 15 | 8584 | Confirmed | 15.15 |
|   | BCL2 | NM_000633 | p.S97R | c.261C>A | 26 | 1338 | Confirmed | 21.2 |
|   | KMT2D | NM_003482 | p.S469* | c.1403C>A | 13 | 9012 | Confirmed | 35 |
|   | KMT2D | NM_003482 | p.S477P | c.1429T>C | 13 | 13514 | Confirmed | 14.21 |
|   | CREBBP | NM_004380 | p.Y1503D | c.4507T>G | 8 | 6861 | Confirmed | 29.4 |
|   | IGLL5 | NM_001256296 | p.I91S | c.273C>T | 16 | 694 | ND | 10.87 |
|   | IGLL5 | NM_001256296 | p.A30V | c.90C>T | 26 | 702 | ND | 5.331 |
|   | GNA13 | NM_006572 | p.L54* | c.159_161delinsCTA | 15 | 15333 | Confirmed | 35 |
|   | GNA13 | NM_006572 | p.D222N | c.664G>A | 17 | 695 | Confirmed | 32 |
|   | MEF2B | NM_001145785 | p.E77A | c.230A>C | 14 | 1248 | Confirmed | 27.9 |
|   | TBL1X1R1 | NM_0024656 | p.L198* | c.592_609delinsT | 34 | 16045 | ND | 35 |
|   | PIM1 | NM_002648 | p.M11 | c.35G>A | 13 | 2241 | Confirmed | 24.1 |

|   |   |   |   |   |   |
| 7 | ISFN |   |   |   |   |
|   | BCL2 | NM_000633 | p.D26V | c.765G>G | 12 | 1902 | Confirmed | 13.04 |
|   | BCL2 | NM_000633 | p.R90S | c.268C>T | 23 | 1911 | Confirmed | 23.9 |
|   | BCL2 | NM_000633 | p.F153L | c.457T>C | 12 | 1481 | ND | 32 |
|   | BCL2 | NM_000633 | p.V162D | c.487G>A | 10 | 5631 | ND | 26.5 |
|   | EZH2 | NM_004456 | p.Y646C | c.1937A>G | 27 | 2218 | ND | 25.7 |
|   | CREBBP | NM_004380 | p.Y1482S | c.4445A>C | 15 | 6079 | Confirmed | 28.4 |
|   | MEF2B | NM_001145785 | p.D83V | c.248A>T | 22 | 1813 | ND | 26.5 |
|   | TBL1X1R1 | NM_0024656 | p.L198* | c.592_609delinsT | 23 | 10682 | ND | 35 |

|   |   |   |   |   |   |
| 8 | ISFN |   |   |   |   |
|   | EZH2 | NM_004456 | p.Y1492C | c.4496T>C | 32 | 715 | Confirmed | 32 |
|   | CREBBP | NM_004380 | p.L1499P | c.4496T>C | 32 | 715 | Confirmed | 32 |
|   | CARD11 | NM_003245 | p.S50P | c.150G>A | 23 | 9399 | Confirmed | 23.9 |

|   |   |   |   |   |   |
| 9 | HGBL-DH |   |   |   |   |
|   | BCL2 | NM_000633 | p.D100A | c.300G>A | 37 | 4171 | Confirmed | 25.0 |
|   | BCL2 | NM_000633 | p.N11D | c.331G>A | 32 | 4766 | Confirmed | 23.1 |
|   | EZH2 | NM_004456 | p.Y646N | c.1937T>A | 56 | 21249 | Confirmed | 24.5 |
|   | CREBBP | NM_004380 | p.L1499P | c.4496T>C | 32 | 715 | Confirmed | 32 |
|   | TBL1X1R1 | NM_0024656 | p.S171C | c.512G>C | 68 | 5035 | ND | 24.2 |
|    | ISFN | BCL2 | NM_000633 | p.L86F | c.256C>T | 15  | 110 | Confirmed | 19.85 |
|----|------|------|-----------|--------|----------|-----|-----|-----------|-------|
|    | FL   | BCL2 | NM_000633 | p.L86F | c.256C>T | 25  | 301 | Confirmed | 19.85 |
|    |      | BCL2 | NM_000633 | p.E135D| c.405G>C | 8   | 801 | Confirmed | 18.34 |
|    |      |      |           |        |          |     |     |           |       |
|    | HGBL-DH | BCL2 | NM_000633 | p.L86F | c.256C>T | 55  | 2690| Confirmed | 19.85 |
|    |      | EZH2 | NM_004456 | p.Y646F| c.1937A>T| 23  | 4420| Confirmed | 25.3  |
|    |      | HIST1H1B | NM_005322 | p.S107C| c.320C>G | 32  | 1371| Confirmed | 32    |
|    | ISFN | EZH2 | NM_004456 | p.Y646F| c.1937A>T| 9   | 357 | Confirmed | 25.3  |
|    | FL   | EZH2 | NM_004456 | p.Y646F| c.1937A>T| 52  | 409 | Confirmed | 25.3  |
|    |      | KMT2D | NM_003482 | p.S831*| c.2492C>A| 13  | 505 | Confirmed | 22.2  |
|    | DLBCL | EZH2 | NM_004456 | p.Y646F| c.1937A>T| 40  | 1559| Confirmed | 25.3  |
|    |      | KMT2D | NM_003482 | p.S831*| c.2492C>A| 21  | 1131| Confirmed | 22.2  |

Bold letters indicate that mutations are shared between ISFN and FL and/or aggressive BCL. ND, Not done. Mutations with a CADD algorithm score >15 were considered deleterious. *Bidirectional single amplicon sequencing.
Supplementary Table S5. Overview of synonymous and 5'UTR mutations of **BCL2**.

| Case | Diagnosis | Protein level | cDNA change | VAF (%) | Coverage |
|------|-----------|---------------|-------------|---------|----------|
| De novo aggressive B-cell lymphoma |
| 1 ISFN | HGBL-TH | 5'UTR | c.1-18G>A | 42 | 7049 |
| | | 5'UTR | c.1-1G>A | 42 | 7057 |
| | | Synonymous | c.207C>T | 44 | 2475 |
| | | Synonymous | c.381G>A | 30 | 27929 |
| 2 ISFN | DLBCL | 5'UTR | c.1-2G>C | 12 | 1516 |
| | | 5'UTR | c.1-17C>G | 20 | 2472 |
| | | 5'UTR | c.1-2G>C | 20 | 2478 |
| 3 ISFN | DLBCL | Synonymous | c.67C>T | 15 | 6895 |
| | | 5'UTR | c.1-17C>G | 55 | 2601 |
| | | Synonymous | c.67C>T | 55 | 3771 |
| | | Synonymous | c.588T>C | 58 | 8330 |
| 4 ISFN | DLBCL | 5'UTR | c.1-49G>C | 9 | 1085 |
| | | 5'UTR | c.1-1G>A | 15 | 8031 |
| 5 ISFN | DLBCL | Synonymous | c.186C>T | 51 | 2892 |
| | | Synonymous | c.261C>T | 52 | 3191 |
| 6 ISFN | DLBCL | 5'UTR | c.1-1G>C | 21 | 3096 |
| | | Synonymous | c.67C>T | 21 | 9335 |
| | | Synonymous | c.291C>G | 16 | 26582 |
| | | Synonymous | c.355C>T | 12 | 26894 |
| | | Synonymous | c.438G>A | 17 | 12968 |
| | | Synonymous | c.456G>A | 9 | 15227 |
| | | 5'UTR | c.1-17C>T | 26 | 5333 |
| Transformed FL |
| 7 ISFN | DLBCL | 5'UTR | c.1-17C>A | 11 | 1391 |
| | | Synonymous | c.24G>A | 12 | 1398 |
| | | Synonymous | c.67C>T | 11 | 1981 |
| | | 5'UTR | c.1-17C>A | 22 | 1609 |
| | | Synonymous | c.24G>A | 23 | 1630 |
| | | Synonymous | c.67C>T | 24 | 2201 |
| 8 ISFN | DLBCL | 5'UTR | c.1-17C>A | 50 | 2574 |
| | | Synonymous | c.67C>T | 56 | 3402 |
| | | Synonymous | c.354G>A | 25 | 8459 |
| | | Synonymous | c.408G>A | 55 | 7506 |
| | | Synonymous | c.447C>G | 64 | 7523 |
| 9 ISFN | HGBL-DH | 5'UTR | c.1-49G>C | 29 | 4229 |
| | | 5'UTR | c.1-17C>G | 14 | 4215 |
| | | Synonymous | c.426G>A | 43 | 8014 |
| 10 ISFN | HGBL-DH | Synonymous | c.66G>A | 10 | 389 |
| | | Synonymous | c.258C>T | 16 | 109 |
| | | Synonymous | c.357G>A | 9 | 1637 |
| | FL | Synonymous | c.66G>A | 23 | 618 |
| | | Synonymous | c.258C>T | 23 | 295 |
| | | Synonymous | c.357G>A | 13 | 1832 |
| | 5'UTR | Synonymous | c.66G>A | 46 | 5918 |
| | | Synonymous | c.93T>C | 48 | 5722 |
| | | Synonymous | c.258C>T | 56 | 2682 |

All mutations refer to the NM_000633 transcript of the **BCL2** gene. Bold letters indicate that mutations are shared between ISFN and FL and/or aggressive BCL. 5'UTR, 5' untranslated region.
Supplementary figures

Supplementary Figure S1. Branched evolution illustrated by phylogenetic trees. The trees were constructed using the ten most prevalent subclones of every sample and rooted to the corresponding VDJ germline sequence. Aggressive BCL is represented in blue, FL in green and ISFN in red. The bar graphs show the share of each subclonal sequence out of the total number of clone-specific reads of the respective sample.
Supplementary Figure S2. Patterns of clonal evolution based on the distribution of private and shared mutations. The respective evolutionary pattern is indicated in parentheses. All variants are depicted at protein level. Mutations highlighted in red were gained during the evolution. Synonymous and 5'UTR variants of BCL2 are not shown, but were also taken into account for the construction. The existence of “Progenitor clones” was assumed based on the distribution of mutations.