Supplementary material

Schmidt et al.

Materials and Methods

DNA isolation

DNA of samples was isolated with the QIAamp DNA FFPE Tissue Kit (QIAGEN GmbH, Hilden, Germany) according to the manufacture's protocol. BCL2 positive germinal center (GC) cells of ISFN samples were microdissected as previously described [1]. DNA concentration of mFL samples was determined with the Qubit dsDNA BR Assay Kit (Life Technologies GmbH, Darmstadt, Germany) according to the manufacture's protocol.

Sanger sequencing

For validation of the NGS data primers were designed with Primer3 software (version 4.0.0; primer3.ut.ee/) and purchased over Sigma-Aldrich (Steinheim, Germany) (see Table 1). PCRs were performed using 100 ng of DNA in a final volume of 25 µl with 200 µM dNTPs (Thermo Scientific, Schwerte, Germany), 0.2 µM per Primer, and 0.02 U/µl Phusion Hot Start High Fidelity Polymerase (Thermo Scientific). Purification of PCR products and subsequent analysis of sequencing reactions was performed as described elsewhere [2].

Targeted resequencing on the Illumina MiSeq system

Variants that could not be detected by Sanger sequencing due to low allelic frequencies were validated by targeted resequencing on the Illumina MiSeq system using the Access Array Barcode Library (Fluidigm). For amplification of the regions of interest we took the same primers as above or designed new ones (see Table 2). The first PCR was performed using 50 ng of DNA in a final volume of 25 µl with 200 µM dNTPs (Thermo Scientific), 0.2 µM per Primer, and 0.02 U/µl Phusion Hot Start High Fidelity Polymerase (Thermo Scientific). PCR products were purified with AMPure Beads. In a second PCR the Fluidigm Barcodes were ligated to the fragments of the first PCR using 1 µl of PCR product, 2 µl Barcode primers (2 µM), 0.05 U/µl Fast Start High Fidelity Enzyme Blend, 0.4 µl DMSO and 0.2 mM dNTPs in a final volume of 20 µl. The products were again purified with AMPure Beads and were pooled in a final volume of 300 µl and a final DNA concentration of 8 nM. The Illumina MiSeq run was carried out by the Department of Medical Genetics of the University Hospital Tübingen.

Raw Next generation sequencing (NGS) data analysis

The reads were aligned to the human reference sequence build 38 (hg19) using the TMAP aligner implemented in the Torrent Suite software. Detection of single base pair variants and insertion-deletion polymorphisms (InDels) compared to the human reference sequence was performed using Ion Torrent Variant Caller (5.0.3.5). Variants were annotated and filtered against the dbSNP and COSMIC databases using the Annotate Variants Single Sample

Workflow of the Ion Reporter Software (version 4.2). Each variant was also inspected with the Integrative Genomics Viewer (IGV, Broad Institute, version 2.3) software to exclude artifacts.

Prediction of mutation effect

Prediction of the deleterious effect of variants was done using SIFT (http://sift.jcvi.org/) and Polyphen-2 (Polymorphism Phenotyping-2) (http://genetics.bwh.harvard.edu/pph2/), both sequence homology-based tools that predict the possible functional consequences of an amino acid substitution on the structure and function of a human protein [3,4]; and CADD (Combined Annotation Dependent Depletion) (http://cadd.gs.washington.edu/home) predictor for scoring the deleteriousness of single nucleotide variants as well as insertion/deletion variants in the human genome [5]. CADD analysis was performed over VCF files generated.

NGS-based B-cell clonality analysis and determination of N-glycosylation sites

B-cell clonality analysis was performed in eight ISFN (cases 4-11) and three mFL cases (cases 4-6) with the LymphoTrack Dx *IGH* FR2 Assay – PGM (Invivoscribe, San Diego, CA) according to the manufacturer's instructions. Libraries were purified and quantified applying Agencourt AMPure XP (Beckman Coulter, Brea, CA, USA) magnetic beads and the Ion Library Quantitation Kit (Thermo Fisher Scientifc, 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 PGM platform. Sequencing data were analyzed and visualized with the associated LymphoTrack Dx Software – PGM and interpreted according to the manufacturer's protocol. This analysis allows the detection of IGH clonal rearrangements and provides the frequency distribution of IGH *VH* - *JH* gene rearrangements and sequence information. The sequences of the clonal rearrangements were submitted to IMGT/V-QUEST, an integrated alignment tool for nucleotide sequences of immunoglobulins and T-cell receptors [6], which reports the translated amino acid sequence of the V-region including framework 2 (FR2), complementarity-determining region 2 (CDR2), FR3 and CDR3. N-glycosylation sites in the V-region are based on a motif with the consensus sequence Asn-Xaa-Ser/Thr, whereupon the presence of proline between Asn and Ser/Thr will inhibit N-glycosylation [7].

Supplemental References

1. 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, 28:1103-12
2. Adam P, Baumann R, Schmidt J, et al, The BCL2 E17 and SP66 antibodies discriminate 2 immunophenotypically and genetically distinct subfroups of conventionally BCL" "negative" grade1/2 follicular lymphomas. Human Pathol, 2013, 44:1817-26
3. Kumar, P., Henikoff, S., & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature Protocols 2009, 47:1073–1081.
4. Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Sunyaev, S. R. A method and server for predicting damaging missense mutations. Nature Methods 2010
5. Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., & Shendure, J. A general framework for estimating the relative pathogenicity of human genetic variants. Nature Genetics 2014, 46(3), 310–315.
6. Brochet, X., Lefranc, M.P., 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, 1(36), 503-508
7. Berg, D.T., Grinnell, B.W. Pro to Gly (P219G) in a silent glycosylation site results in complete glycosylation in tissue plasminogen activator. Protein Sci 1993, 2(1):126-127.

Supplemental Table 1. Primer sequences for mutation verification with Sanger Sequencing.

Primer	Sequence
TNFRSF14_AMPL7153185521_F	5‘-TCTTTCTCTTCTGGCCCACA-3‘
TNFRSF14_AMPL7153185521_R	5‘-AGCCTCAAGACGTCGGTTTT-3‘
CREBBP_AMPL7153985750_F	5‘-AGAAGCACCTTGTCAGCAAC-3‘
CREBBP_AMPL7153985750_R	5‘-ACATCCACGCCGTCAATTTC-3‘
EP300_AMPL7153092224_F	5‘-CAAAGAGCCTGGGAGAGTGA-3‘
EP300_AMPL7153092224_R	5‘-ACTGCAGTCCTCAAGCATTT-3‘
KMT2D_AMPL7153087589_F	5‘-TACCCCGAATGCAAGCAGA-3‘
KMT2D_AMPL7153087589_R	5‘-TCACCCTGGCTCAGATTAGAGATC-3‘
CREBBP_AMPL7153985800_F	5‘-CCTGAGTTAAACATGTGCCTCC-3‘
CREBBP_AMPL7153985800_R	5‘-CAAACTCACCCTAATTTCTTCACATACTCT-3‘
KMT2D_AMPL7153087920_F	5‘-CAGGGTCAGGGCCTTATGG-3‘
KMT2D_AMPL7153087920_R	5‘-CTGCTGAAGCTGCTGTAAAGAG-3‘
KMT2D_AMPL7153086220_F	5‘-GAATTGCCCACTTCCCCG-3‘
KMT2D_AMPL7153086220_R	5‘-GGGAACAGACGAGATGCC-3‘
TNFRSF14_AMPL7153185647_F	5‘-CTCCCTGAGGCTGAGTGAAC-3‘
TNFRSF14_AMPL7153185647_R	5‘-TGTGGAGCAAACAATGACGA-3‘
KMT2D_AMPL7153088071_F	5‘-TGGGCATCAGATGTAGGACA-3‘
KMT2D_AMPL7153088071_R	5‘-AGACAGGACACAGTAACCCC-3‘
TNFRSF14_AMPL7153185584_F	5‘-CGGAAAAGACAGGAGGCAGA-3‘
TNFRSF14_AMPL7153185584_R	5‘-CAGGTATCTCTGGCGTCGG-3‘

Supplemental Table 2. Primer sequences for mutation verification with the Access Array Barcode Library on the Illumina MiSeq.

Primer	Sequence
TNFRSF14_AMPL-185584_Flui-F	5‘-CGGAAAAGACAGGAGGCAGA-3‘
TNFRSF14_AMPL-185584_Flui-R	5‘-CAGGTATCTCTGGCGTCGG-3‘
CREBBP_AMPL-985750_Flui-F	5‘-AGAAGCACCTTGTCAGCAAC-3‘
CREBBP_AMPL-985750_Flui-R	5‘-ACATCCACGCCGTCAATTTC-3‘
EP300_AMPL-092224_Flui-F	5‘-CAAAGAGCCTGGGAGAGTGA-3‘
EP300_AMPL-092224_Flui-R	5‘-ACTGCAGTCCTCAAGCATTT-3‘
TNFRSF14_AMPL-185647_Flui-F	5‘-CTCCCTGAGGCTGAGTGAAC-3‘
TNFRSF14_AMPL-185647_Flui-R	5‘-TGTGGAGCAAACAATGACGA-3‘
CREBBP_AMPL-985800_Flui-F	5‘-CCTGAGTTAAACATGTGCCTCC-3‘
CREBBP_AMPL-985800_Flui-R	5‘-CAAACTCACCCTAATTTCTTCACATACTCT-3‘
KMT2D_AMPL-087589_Flui-F	5‘-TACCCCGAATGCAAGCAGA-3‘
KMT2D_AMPL-087589_Flui-R	5‘-TCACCCTGGCTCAGATTAGAGATC-3‘
CREBBP_Exon27A_Flui-F	5‘-CACTGGCACGTTCATCTGAC-3‘
CREBBP_Exon27A_Flui-R	5‘-TTGATCAGGTGGGTGGCAAT-3‘
CREBBP_Exon27B_Flui-F	5‘-CCTGTCCTCCAAGTGAAGGA-3‘
CREBBP_Exon27B_Flui-R	5‘-CAAACGCCTTGTCCAGCAT-3‘
CREBBP_Exon26_Flui-F	5‘-TGAGGGATACCCTGAGTTAAACA-3‘
CREBBP_Exon26_Flui-R	5‘-CCAAACTCACCCTAATTTCTTCA-3‘
EZH2_Flui_F	5‘-TATTGCTGGCACCATCTGAC-3‘
EZH2_Flui_R	5‘-TCCAATCAAACCCACAGACTT-3‘

Supplemental Table 3. Clinical data of 11 patients with ISFN-only and ISFN/mFL

Patient	Age (years)	Sex	Site	Stage	Additional information	Follow-up
ISFN/mFL pairs						
ISFN-1	76	M	Supraclavicular LN		Synchronous	
mFL-1			Supraclavicular LN	IIIA	BCL2 negative due to BCL2 mutation	CR – 36 months (CT)
ISFN-2	48	F	Cervical LN		ISFN was diagnosed 1 month after mFL	
mFL-2			Skin	IEA		CR – 134 months (CT)
ISFN-3	55	F	Submandibular LN		ISFN was diagnosed 24 months before mFL	
mFL-3			Cervical LN	IA		CR – 110 months (RT)
ISFN-4	74	F	Mesentierial LN		Synchronous	
mFL-4			Mesenterial LN	IVB		DwD – 30 months (CT)§
ISFN-5	57	M	Paragrastric LN		ISFN incidental diagnosis. Gastric carcinoma	
mFL-5			Cervical LN	IIIB	ISFN was diagnosed 48 months after mFL	DwD – 60 months (RT+CT)*
ISFN-6	71	F	Axillary LN		Synchronous	
mFL-6			Axillary LN	IIIA	Breast carcinoma, FL incidental diagnosis	12 months (W&W)
ISFN-only cases						
ISFN-7	80	F	Jugular LN	-	Classic Hodgkin Lymphoma, synchronous	LFU
ISFN-8	45	M	Cervical LN	-	NLPHL, synchronous	LFU
ISFN-9	89	F	LN	-	DLBCL in the tonsil 20 years before, in CR	LFU
ISFN-10	69	M	Mediastinal LN	-	Lung carcinoma	NED – 46 months
ISFN-11	72	M	Mesenterial LN	-	Pancreas carcinoma	DwD – 30 months*

ISFN: in situ follicular neoplasia; mFL: manifest follicular lymphoma; F:female; M:male; LN:lymph node; NLPHL: nodular lymphocyte predominant Hodgkin lymphoma; DLBCL: Diffuse large B-cell lymphoma; CR: complete remission; CT: chemotherapy; RT: radiotherapy; W&W; watch and wait, LFU: lost to follow up; NED: not evidence of disease
*Died secondary to carcinoma without evidence of follicular lymphoma; §died without evidence of follicular lymphoma; DwD: dead without disease

Supplemental Table 4. Genetic analyses in ISFN-only and in pairs ISFN/mFL

Patient	*BCL2*-BAP FISH	MBR /MCR/ICR	Clonality	Glycosylation site Location motif	IGH V/J Usage	% of clonal sequence
ISFN/mFL pairs						
ISFN-1 (1)	+ (80%)	MBR-JH	monoclonal	NA	-	27.3
mFL-1	+	MBR-JH	monoclonal	NA		43.10
ISFN-2	+	MBR-JH	monoclonal	NA	-	28.30
mFL-2	NA	MBR-JH	monoclonal	NA		34.3
ISFN-3	+	neg	monoclonal	NA	-	14.4%
mFL-3	+	neg	monoclonal	NA		17.7%
ISFN-4	+	neg	Monoclonal§	Clone not detected	-	-
mFL-4	+	neg	Monoclonal§	Clone not detected		
ISFN-5	+	MBR-JH	monoclonal	FR2/CDR2 NIS	V3-11/J4	13.47
mFL-5	+	MBR-JH	monoclonal	FR2/CDR2 NIS	V3-11/J4	30.25
ISFN-6	+	MBR-JH	Monoclonal	FR2 PCR ND	V3-48/J6	28.00
mFL-6	+	MBR-JH	Monoclonal	FR2 PCR ND	V3-48/J6	14.91
ISFN-only cases						
ISFN-7	NA	ICR-JH	monoclonal	FR2/CDR2 NIS	V3-46/J6	48.13
ISFN-8	+ (20%)	neg*	monoclonal	CDR3 NGT	V3-30/J5	30.03
ISFN-9	+ (80%)	MBR-JH	poly#	Clone not detected	-	-
ISFN-10 (2)	+ (80%)	neg	monoclonal	CDR3 NCS	V2-23/J4	48.76
ISFN-11	+ (60%)	neg	Poly#	Clone not detected	-	-

BAP: Break apart probe; FISH: Fluorescence in situ hybridization; ISFN: in situ follicular neoplasia; mFL: manifest follicular lymphoma; MBR: major breakpoint region; ICR: intermediate cluster region;JH: joining region; NA: not available,
ND: not detected; * only MBR was performed. # FR2/FR3 only
Pairs ISFN/mFL 1 to 3: the % of clonal sequence was determined by GeneScan analysis.
§ monoclonal only with IGK, kappa deleted element: IgKappa A: 150 (Vk1f/6/Vk7-Jk) IgKappa B: 230 (Vk1f/6/Vk7-Kde).

(1) Bonzheim I, Salaverria I, Haake A, et al, Blood 2011;118(12):3442-4
(2) Kosmidis P, Bonzheim I, Dufke C, et al PLoS One 2017; 97(9):1095-1102

Supplemental Table 5. Prediction of amino acid changes that affect protein function using SIFT, Polyphen and CADD algorithms.

Case	Gene	Protein level	DNA level	SIFT	Polyphen	CADD score*
mFL-1	*CREBBP*	p.L1499Q	c.4496T>A	Affect protein function (0)	Probably damaging (1)	---
mFL-2	*EZH2*	p.Y646N	c.1936T>A	Affect protein function (0)	Probably damaging (0.98)	32
	CREBBP	p.I1471T	c.4412T>C	Affect protein function (0)	Probably damaging (0.998)	---
mFL-3	*TNFRSF14*	p.C185S	c.553T>A	Affect protein function (0)	Probably damaging (0.994)	13.29
	CREBBP	p.L1434P	c.4301T>C	Affect protein function (0)	Probably damaging (1)	24.8
	CREBBP	p.R1446H	c.4337G>A	Affect protein function (0)	Probably damaging (1)	29.1
	EZH2	p.Y646H	c.1936T>C	Affect protein function (0)	Probably damaging (0.98)	32
	KMTD2	p.K1840fs	c.5519_5529delAAGCCGATACA	---	---	35
mFL-4	*TNFRSF14*	p.S171C	c.512C>G	Affect protein function (0.01)	Probably damaging (1)	24.2
	EZH2	p.Y646N	c.1936T>A	Affect protein function (0)	Probably damaging (0.98)	32
	CREBBP	p.L1499P	c.4496T>C	Affect protein function (0)	Probably damaging (1)	24.5
mFL-5	*CREBBP*	p.S1382fs	c.4145delA	---	---	35
	EP300	p.D1399Y	c.4195G>T	Affect protein function (0)	Probably damaging (1)	35
mFL-6	*CREBBP*	p.R1446C	c.4336C>T	Affect protein function (0)	Probably damaging (1)	26.9
	EZH2	p.Y646F	c.1937A>T	Affect protein function (0.02)	Possibly damaging (0.817)	32
ISFN-7	*KMT2D*	p.C5227Ter	c.15681C>A	---	---	56
	KMT2D	p.Q3518Ter	c.10552C>T	---	---	43
ISFN-8	*EP300*	p.T1332P	c.3994A>C	Affect protein function (0.01)	Probably damaging (0.996)	26.7
	TNFRSF14	Splice site	c.179-2A>T	---	---	17.25
ISFN-9	*TNFRSF14*	p.C121Ter	c.363C>A	---	---	32
	CREBBP	p.H1487Y	c.4459C>T	Affect protein function (0)	Probably damaging (0.998)	24
ISFN-10	*EZH2*	p.Y646F	c.1937A>T	Affect protein function (0.02)	Possibly damaging (0.817)	32

CADD: combined annotation-dependent depletion; *CADD algorithm score cut-off for deleteriousness is > 15.

Supplemental Figure 1. Chromosome 1 view (band 1p36) of pairs ISFN/mFL-3, ISFN/mFL-4 and ISFN/mFL-5 including *TNFRSF14* mutational status.