Supplementary data

Dissecting the Grey Zone between Follicular Lymphoma and Marginal Zone Lymphoma using Morphological and Genetic Features

Oscar Krijgsman, Patricia Gonzalez, Olga Balagué Ponz, Margaretha GM Roemer, Stefanie Slot, Annegien Broeks, Linde Braaf, Ron M. Kerkhoven, Freek Bot, Krijn van Groningen, Max Beijert, Bauke Ylstra, Daphne de Jong

Extended Methods

Patient selection

From a series of histological formalin-fixed, paraffin-embedded (FFPE) biopsy samples of patients treated in a randomized phase III EORTC trial (E20971) for localized indolent B-cell lymphoma and from the files of the Comprehensive Cancer Center Amsterdam Lymphoma Consult Panel, 4 series of patients for whom adequate arrayCGH results could be obtained were selected for this study. Of the original selection, 5 cases of NMZL, 3 cases of translocation negative FL low stage and 2 cases of translocation positive low stage were omitted, because of technically insufficient quality of the arrayCGH results. Four series of patients for whom arrayCGH results could be obtained were selected for this study:

1) morphologically and immunophenotypical classical NMZL (n=14),

2) indolent B-cell lymphoma with features intermediate between FL and NMZL (translocation negative FL) (n=12),

3) morphological and immunophenotypical classical FL bearing a t(14;18) translocation presenting as stage 1 or limited stage 2 disease (n=16),
4) morphological and immunophenotypical classical FL bearing a t(14;18) translocation presenting as extensive disease (n=14).
All protocols for obtaining and studying human archival tissues and patient data were approved within the local ethical procedures at the Netherlands Cancer Institute and in compliance with the Code for Proper Secondary Use of Human Tissue in The Netherlands (http://www.fmwv.nl and www.federa.org).

Chromosomal copy number aberrations as measured by arrayCGH

Genomic DNA was isolated from tumor FFPE samples using the QIAamps DNA extraction kit (cat. 51306) as described by Beers et al[1]. DNA of 68 samples was hybridized to the NimbleGen Human CGH 12x135K Whole-Genome Tilling v3.0 platform containing 134 937 *in situ* synthesized oligonucleotides (Roche NimbleGen, Madison, USA). Labeling was performed as previously described by Buffart et al. using the Enzo Genomic Labeling kit (Enzo Life Sciences, Farmingdale, NY, USA) with 500 ng of input DNA according to the manufacturers instructions[2]. Images of the arrays were acquired using the Agilent Microarray scanner G2505B (Agilent Technologies) and image analysis was performed using Nimblescan software version 2.5 (Roche NimbleGen).

ArrayCGH Data Analysis

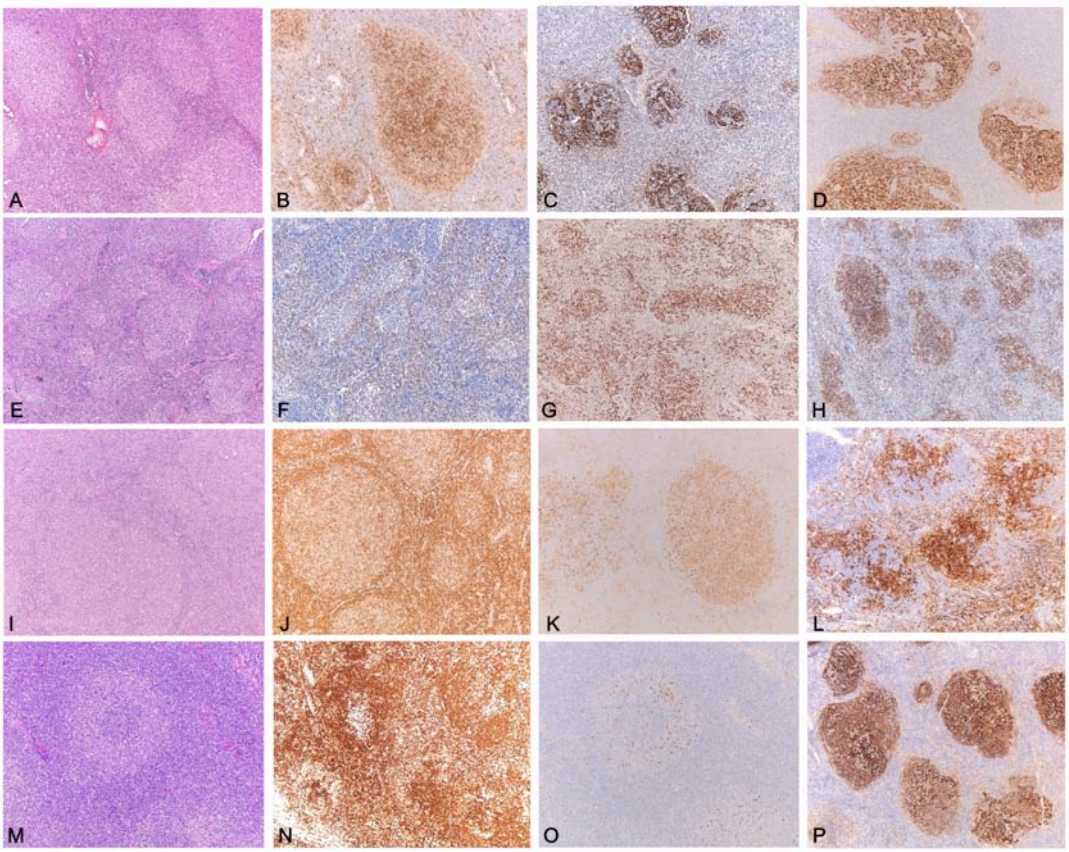
ArrayCGH profiles with a genome-wide median absolute deviation (MAD) >0.20 were excluded from further analysis[2]. Profiles were wave-smoothed using No-waves[3]. Median normalized Log2ratios were segmented with Circular Binary Segmentation (CBS)[4]. Samples with an extreme high number of segments (>1000) were additionally deleted for further analysis. After segmentation samples were mode normalized[3]. Chromosomal copy number aberrations were defined using the R-package CGHcall[3]. CGHregions was used to reduce the dimension of the dataset, accepting maximally 2% information loss (threshold = 0.02)[3]. Hierarchical cluster analysis was performed using the WECCA program, using Kullbach-Leibler divergence for distance measure and 'ward' for linkage.

The high resolution of the arrays allows for the identification of focal chromosomal aberrations, somatic aberrations of 3Mb and smaller[5], which are typically enriched for driver genes[6,7]. To remove common germ-line copy number variants (CNVs) from the dataset, while recognizing lymphoma specific focal chromosomal aberrations, all 494 focal DNA copy number differences smaller than 3 Mb were compared with CNVs in the healthy population as archived in the database of genomic variants (http://projects.tcag.ca/variation/)[8,9]. This removed a total of 263 (53%) CNV regions. The remaining 231 focal regions were furthermore cleaned from germ-line CNVs assuming true somatic aberrations are either gained or lost in a series of tumors, rather than gained and lost as seen for CNVs in a healthy population. This procedure removed another 121 focal DNA copy number differences, such that a total of 110 bonafide focal chromosomal aberrations were carried forward in the down-stream analysis. Chromosome 19 was omitted for tumor characterization because of inconsistent copy number artefacts.

Genes and miRNAs were retrieved using biomaRt (R/Bioconductor) and Ensembl (hg18/ncbi36, ensembl 54). Raw data of all arrays are publicly available in GEO (accession number GSE40641). Analysis was performed and plots were made using the statistical programming language R version 2.11.1 and Bioconductor packages (http://www.r- project.org).

Statistical analysis

To test whether the 110 focal aberrations were significantly enriched for cancer related genes as published in the Cancer Census list[8] enrichment analysis was performed as previously described[15]. Difference between the groups in the percentage of aberrant probes was tested using the non-parametric Mann-Whitney-U test. GAORT (Global Copy Number Aberration Odds Ratio Test) was used to evaluate commonalities. GAORT calculates whether the odds of an aberrant region in one sample group are systematically larger than that of another group[6].

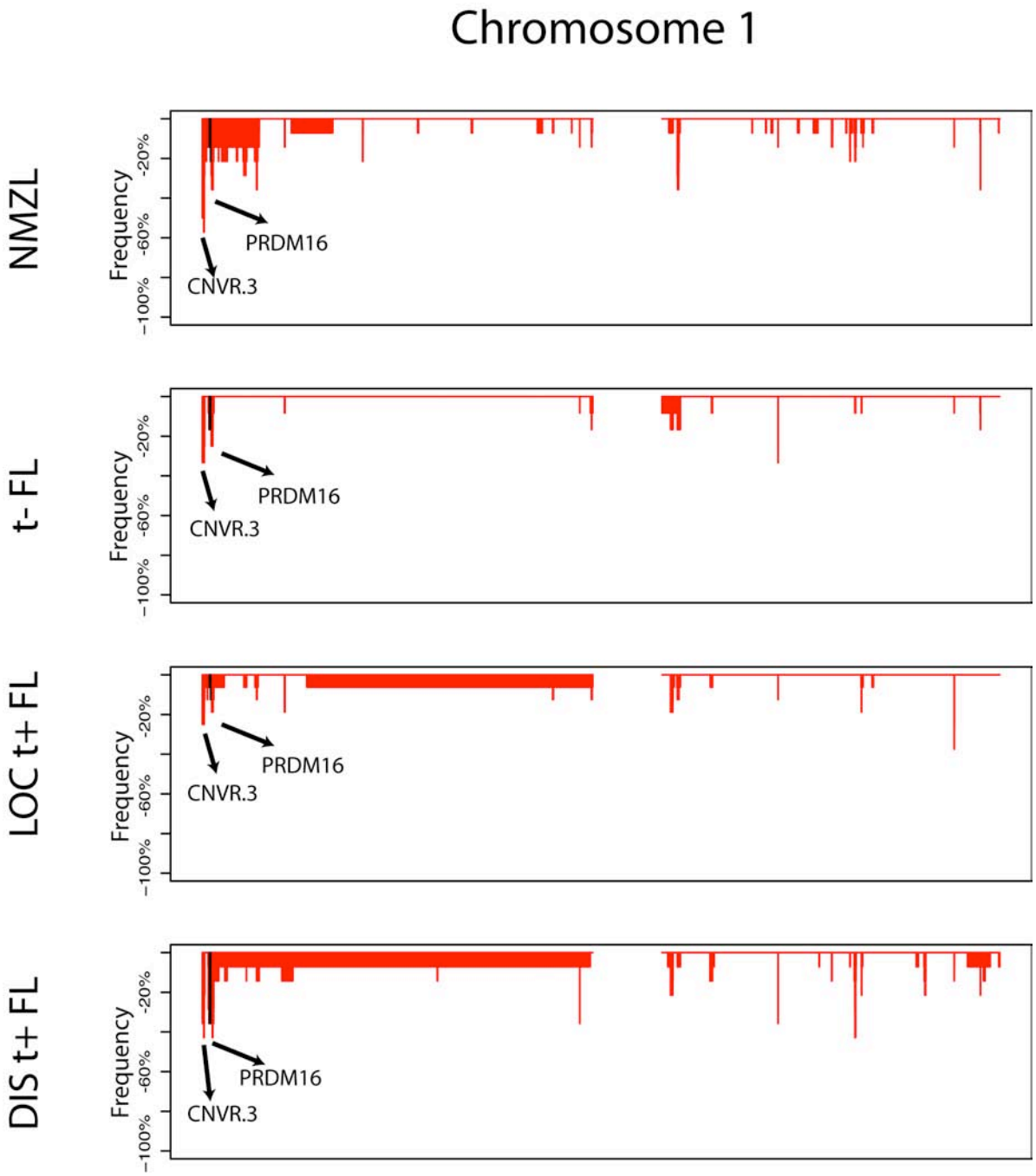


Supplementary figure 1: FL, translocation negative (t- FL case 7: A: morphological features of FL, B: uniform expression of BCL2, C: expression of CD10 in all tumor cells, but with remarkable varying expression, D: CD21 shows well developed FDC meshworks with also a pattern of colonization indistinguishable from those seen in NMZL as in D: NMZL case 1.

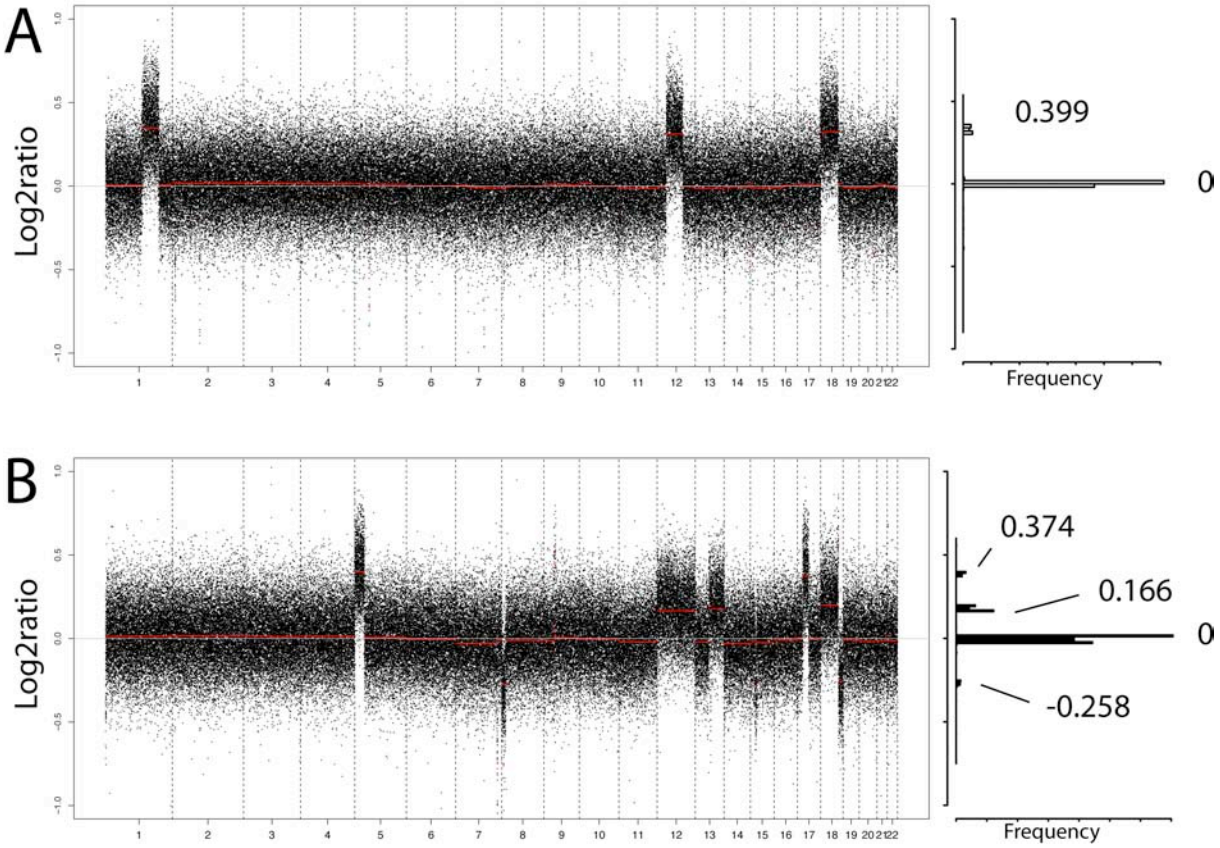
FL, translocation negative (t- FL case 1): E: morphological features of FL with some marginal zone differentiation, F: BCL2 DAKO124 is negative in the malignant population, G: positive for the alternative antibody SantaCruzC2, H: uniformly positive for CD10.

FL, translocation negative (t- FL case 10) I: morphological features of FL J: uniform BCL2 expression, K: heterogeneous CD10 expression, mostly uniform, but occasional nodi with a patchy colonization-type pattern. L: NMZL case 1 shows a patchy pattern for CD10 only very occasionally seen in cases of translocation negative FL.

Prototypical case of NMZL (NMZL case 4) M: marginal zone morphology, N: remnants of reactive GC negative for BCL2, O: CD10 shows an extensive pattern of colonization with minor foci of centrocytes and centroblasts, P: relatively intact CD21 FDC pattern.



Supplementary figure 2: Frequency of loss on chromosome 1. A distinct deletion hotspot was observed in all classes (p36.32). PRDM16 was located on the genomic location most frequently lossed in all subgroups. Black line indicated location of TNFRSF14. CNVR.3 is a well-known germ-line copy number variant (CNV) in the healthy population.



Supplementary figure 3: Example of a sample with low (Top panel) and high (bottom panel) genetic heterogeneity.

Gains of sample 1 on chromosomes 1, 12 and 18 all have a log2ratio of 0.399. Gains of sample 2 on chromosomes 12, 17 and 18 have a log2ratio of 0.166, 0.374 and 0.199 respectively. X-axis represents the probes ordered according to the genomic location, y-axis represents the log2ratio. Red lines represent the CBS segments. Plot on the right side indicate the frequency of the segment values.

Supplementary Table 1. Morphological, immunophenotypical and translocation data of all sample.

Sample type	Sample	H&E morfological pattern	CD10	BCL2	colonization pattern based on all parameters	BCL2 translocation
NMZL	1	MZL pattern	negative, focally moth eaten	negative	yes	no
	2	FL pattern with marginal zone differentiation	negative, focally moth eaten	positive	yes	no
	3	MZL pattern	negative, focally moth eaten	positive, reactive GC, negative foci in GC	yes	no
	4	MZL pattern	negative	positive, reactive GC	no	no
	5	MZL pattern	negative	negative	no	no
	6	MZL pattern	negative	positive, reactive GC	no	no
	7	MZL pattern	negative	positive	no	no
	8	MZL pattern	negative, focally moth eaten	negative	yes	no
	9	MZL pattern	negative	positive, reactive GC	no	no
	10	MZL pattern	negative, focally moth eaten	positive, reactive GC	yes	no
	11	MZL pattern	negative, focally moth eaten	positive, reactive GC, negative foci in GC	yes	no
	12	FL pattern with marginal zone differentiation	negative, focally moth eaten	positive	yes	no
	13	MZL pattern	negative, focally moth eaten	positive, reactive GC, negative foci in GC	yes	no
	14	FL pattern with marginal zone differentiation	negative, focally moth eaten	positive	yes	no
t- FL	1	FL pattern with marginal zone differentiation	positive	DAKO124 negative, Santacruz C2 positive	no	no
	2	FL pattern with marginal zone differentiation	positive	positive	no	no
	3	FL pattern	nodular and focally moth eaten	positive, reactive GC, minor negative foci	dubious	no
	4	FL pattern with marginal zone differentiation	positive	positive	no	no
	5	FL pattern	positive	positive	no	no
	6	FL pattern	positive	positive	no	no

	7	FL pattern	uniform nodular and interfollicular	positive	no	no
	8	FL pattern	nodular and focally moth eaten	positive, minor negative foci	dubious	no
	9	FL pattern	positive nodular and interfollicular	positive	no	no
	10	FL pattern	nodular and focally moth eaten	positive	no	no
	11	FL pattern	uniform nodular	positive	no	no
	12	FL pattern with extensive marginal zone differentiation	uniform nodular and interfollicular	DAKO124 negative, Santacruz C2 positive	no	no*
LOC t+ FL	1	FL pattern	uniform nodular	positive, sporadic reactive GC	no	yes
	2	FL pattern	nodular and focally moth eaten	positive	no	yes
	3	FL pattern with marginal zone differentiation	nodular and focally moth eaten	positive, sporadic reactive GC, minor negative foci	dubious	yes
	4	FL pattern	uniform nodular and interfollicular	positive	no	yes
	5	FL pattern	uniform nodular and interfollicular	positive	no	yes
	6	FL pattern	uniform nodular	positive	no	yes
	7	FL pattern	uniform nodular	DAKO124 partly negative, SC C2 positive	no	yes
	8	FL pattern	uniform nodular	positive	no	yes
	9	FL pattern	uniform nodular	positive	no	yes
	10	FL pattern	uniform nodular	positive	no	yes
	11	FL pattern	uniform nodular	positive	no	yes
	12	FL pattern	uniform nodular	positive	no	yes
	13	FL pattern	uniform nodular	positive	no	yes
	14	FL pattern	uniform nodular	positive	no	yes
	15	FL pattern	uniform nodular	positive	no	yes
	16	FL pattern with a diffuse component	uniform nodular and interfollicular	positive	no	yes

DIS t+ FL	1	FL pattern	positive nodular	positive	no	yes
	2	FL pattern	positive nodular	positive	no	yes
	3	FL pattern	positive and focally moth eaten	positive	no	yes
	4	FL pattern	positive and focally moth eaten	positive	no	yes
	5	FL pattern	positive nodular	positive	no	yes
	6	FL pattern	positive nodular	positive	no	yes
	7	FL pattern	positive nodular	positive	no	yes
	8	FL pattern	positive and focally moth eaten	positive	no	yes
	9	FL pattern	positive nodular	positive	no	yes
	10	FL pattern	positive nodular	positive	no	yes
	11	FL pattern	positive nodular	positive	no	yes
	12	FL pattern	positive nodular	positive	no	yes
	13	FL pattern	positive nodular	positive	no	yes
	14	FL pattern	positive nodular	positive	no	yes

Supplementary table 2: Table with all recurrent focal aberrations and the associated genes.

Chromosome	Start_HFR	End_HFR	MB_HFR	Freq focal_gain	Freq_focal loss	Freq_gain	total_loss	Genes
1	153584094	154451867	0.87	2	0	10	0	SCARNA4, PMF1, BGLAP, SLC25A44, SEMA4A, LMNA, RAB25, MEX3A, ROBLD3, UBQLN4, SSR2, ARHGEF2, RXFP4, KIAA0907, RIT1, SYT11, GON4L, DAP3, YY1AP1, ERVK5, MSTO1, ASH1L, SNORA42
1	156097765	156117348	0.02	12	0	18	0	
1	158432682	158542817	0.11	2	0	9	0	PEX19, COPA, WDR42A, PEA15, CASQ1
1	165561163	165853261	0.29	2	0	10	0	CREG1, CD247, POU2F1
1	173747547	174039070	0.29	2	0	7	0	TNR
1	224942426	224984858	0.04	4	0	6	0	ITPKB
2	30366803	30596856	0.23	4	0	11	0	LCLAT1, LBH
2	60416941	60482548	0.07	6	0	13	0	
2	68813845	68863572	0.05	3	0	11	0	ARHGAP25
2	99695506	100233733	0.54	3	0	7	0	AFF3
2	136596736	136622131	0.03	2	0	6	0	
2	161653601	161796196	0.14	4	0	9	0	TANK
2	196718450	196779274	0.06	8	0	14	0	HECW2, STK17B
2	204253828	204376183	0.12	6	0	12	0	CD28
2	213580351	214267524	0.69	4	0	9	0	SPAG16, IKZF2
2	224930814	225582763	0.65	2	0	8	0	DOCK10, FAM124B, CUL3
3	8744510	8830406	0.09	2	0	6	0	OXTR, CAV3
3	16317781	16531906	0.21	6	0	9	0	RFTN1, OXNAD1
3	31936225	31986803	0.05	2	0	7	0	OSBPL10
3	56692762	56804371	0.11	2	0	6	0	ARHGEF3
3	71231678	71312003	0.08	4	0	7	0	FOXP1
3	71553388	71594056	0.04	4	0	8	0	FOXP1
3	101899670	101976018	0.08	2	0	7	0	ABI3BP, TFG
3	106915256	107059341	0.14	6	0	10	0	CBLB
3	108868468	108936055	0.07	2	0	7	0	BBX
3	126231646	126344329	0.11	2	0	7	0	SLC12A8, HEG1
3	134655281	134680318	0.03	15	0	19	0	BFSP2
3	142550200	142700589	0.15	7	0	10	0	RASA2, ZBTB38
3	144202064	145063075	0.86	2	0	7	0	SLC9A9, CHST2
3	153453862	153645248	0.19	2	0	7	0	TMEM14E, MBNL1
3	189496573	189657311	0.16	4	0	9	0	LPP
4	4520632	4596994	0.08	2	0	3	0	STX18
4	10254686	10618143	0.36	5	0	7	0	CLNK
4	25434953	25504008	0.07	7	0	9	0	
4	26482593	26621391	0.14	3	0	5	0	STIM2
4	39870905	39917026	0.05	7	0	9	0	RHOH
4	114842033	114907643	0.07	3	0	5	0	CAMK2D
4	147027674	147194762	0.17	3	0	5	0	ZNF827
5	66513318	66549719	0.04	2	0	4	0	CD180

5	86312320	86724135	0.41	2	0	3	0	RASA1
5	140720696	140995259	0.27	2	0	3	0	HDAC3, DIAPH1, PCDHGA12, PCDHGC5, PCDHGC4, PCDHGC3, PCDHGA11, PCDHGB7, PCDHGA8, PCDHGA6, PCDHGB3, PCDHGA3, PCDHGA2
5	158154776	158300898	0.15	14	0	15	0	EBF1
5	170279147	170594741	0.32	2	0	3	0	RANBP17
6	5082958	5274217	0.19	4	0	6	0	FARS2, LYRM4
6	34194307	34281616	0.09	2	0	4	0	GRM4
7	22203520	22433468	0.23	3	0	10	0	RAPGEF5
7	30276716	30415475	0.14	9	0	16	0	ZNRF2, MIRN550-1
7	43625391	43729656	0.10	5	0	11	0	C7orf44, STK17A
7	50322454	50385653	0.06	7	0	13	0	IKZF1
7	92123956	92167966	0.04	6	0	11	0	CDK6
8	28798997	28974871	0.18	3	0	9	0	HMBOX1, INTS9
8	29200798	29315589	0.11	4	0	9	0	DUSP4
8	37848532	37882075	0.03	2	0	8	0	RAB11FIP1
8	38250506	38387660	0.14	3	0	9	0	LETM2, WHSC1L1
8	38854976	38970504	0.12	2	0	7	0	TM2D2, HTRA4, PLEKHA2
8	82074390	82221876	0.15	2	0	8	0	PAG1
8	101504486	101604091	0.10	10	0	16	0	ANKRD46
8	129128074	129347941	0.22	3	0	11	0	MIRN1207, MIRN1208
8	134075485	134291792	0.22	3	0	13	0	WISP1, SLA, TG
9	33121502	33285126	0.16	3	0	4	0	NFX1, CHMP5, BAG1, SPINK4, B4GALT1
9	37045315	37285383	0.24	23	0	24	0	ZCCHC7
9	129024037	129082284	0.06	3	0	6	0	GARNL3, RALGPS1
10	3794336	3887363	0.09	2	0	4	0	KLF6
10	5778055	5970500	0.19	4	0	6	0	ANKRD16, GDI2, C10orf18
10	10683449	10726993	0.04	2	0	3	0	
10	30772578	30863840	0.09	3	0	4	0	MAP3K8
11	14267752	14333419	0.07	3	0	5	0	RRAS2
11	35019036	35140815	0.12	3	0	5	0	CD44
11	46049401	46097166	0.05	3	0	5	0	PHF21A
11	47973064	48028222	0.06	2	0	4	0	PTPRJ
11	59822821	60156523	0.33	3	0	5	0	C11orf64, MS4A13, MS4A12, MS4A1, MS4A5, MS4A14, MS4A7, MS4A6E, MS4A4A
11	60880796	60974556	0.09	7	0	8	0	C11orf79, TMEM216, TMEM138, CYBASC3
11	73368586	73406919	0.04	2	0	4	0	UCP3, UCP2
11	82260373	82701089	0.44	2	0	4	0	CCDC90B, ANKRD42, PCF11, RAB30, C11orf82, PRCP
11	94520315	94728163	0.21	2	0	4	0	SESN3
11	118374762	118508084	0.13	3	0	6	0	RPL23AP64, HMBS, H2AFX, DPAGT1, HINFP, C2CD2L, VPS11, HYOU1, SLC37A4, TRAPPC4, RPS25P8, RPS25, CCDC84
11	119698174	119855229	0.16	2	0	5	0	ARHGEF12, TMEM136
11	120787775	120947646	0.16	2	0	5	0	SORL1
12	6756727	6955518	0.20	2	0	8	0	MIRN200C, RPL13P5, EMG1, PHB2, PTPN6, C12orf57, ATN1, ENO2, LRRC23, SPSB2, TPI1, USP5, CDCA3, GNB3, LEPREL2, GPR162, CD4, LAG3, MIRN141
12	55738724	55788870	0.05	3	0	15	0	STAT6, NAB2, TMEM194A
12	91271885	91473227	0.20	8	0	17	0	CLLU1OS, CLLU1
14	23645240	23814519	0.17	2	0	2	0	RABGGTA, TGM1, TINF2, GMPR2, NEDD8, CHMP4A, TSSK4, TM9SF1, IPO4, REC8, IRF9, RNF31, PSME2, FAM158A, PSME1, WDR23, NRL
14	61179264	61218064	0.04	2	0	2	0	

14	67806397	68040627	0.23	24	0	24	0	RAD51L1
14	69139254	69219265	0.08	2	0	2	0	KIAA0247
14	70182478	70199328	0.02	2	0	2	0	TTC9
14	85531791	85623404	0.09	0	2	0	2	
14	87483739	87526857	0.04	3	0	3	0	GALC
14	101268474	101369474	0.10	3	0	3	0	PPP2R5C
15	75704613	75825822	0.12	0	3	0	5	LINGO1
17	4785451	5078597	0.29	4	0	5	0	C17orf87, ZNF594, USP6, ZNF232, ZFP3, GPR172B, KIF1C, CAMTA2, SPAG7, ENO3, PFN1, RNF167
17	7076130	7097166	0.02	3	0	4	0	C17orf81, DULLARD, GABARAP, PHF23, DVL2
17	24069882	24159499	0.09	2	0	4	0	C17orf63, TRAF4, NEK8, TLCD1, RPL23AP65, RPL23AP75, RPL23AP44, RPL23AP43, RPL23AP37, RPL23AP63, RPL23A
17	35173062	35202871	0.03	10	0	12	0	IKZF3
17	52598231	53054955	0.46	10	0	12	0	MSI2
18	19798519	19853000	0.05	3	0	12	0	TTC39C
18	32577911	32784951	0.21	2	0	10	0	KIAA1328, C18orf10, FHOD3
18	51061903	51283893	0.22	4	0	12	0	TCF4
18	54346812	54805154	0.46	6	0	14	0	ZNF532, MALT1, ALPK2
18	59035763	59126593	0.09	12	0	18	0	BCL2
20	18899902	19410394	0.51	0	2	0	2	SLC24A3
20	30270559	30387922	0.12	3	0	3	0	MIRN1825, KIF3B, POFUT1
20	36298669	36349361	0.05	0	2	0	2	KIAA1755
20	36850948	37044675	0.19	2	0	2	0	DHX35, FAM83D, PPP1R16B, NPM1P19
20	45411479	45578391	0.17	8	0	8	0	NCOA3, ZMYND8, RPL35AP
20	49712091	49954450	0.24	0	2	0	3	SALL4, ATP9A
20	56642700	56768776	0.13	0	3	0	4	NPEPL1, STX16
21	42178422	42208432	0.03	4	0	4	0	C2CD2
21	44620380	44859182	0.24	0	2	0	2	KRTAP10-8, KRTAP10-7, KRTAP10-6, KRTAP10-5, KRTAP10-4, KRTAP10-3, KRTAP10-1, KRTAP10-2, C21orf90, C21orf29, LRRC3, TRPM2
22	44202808	44482934	0.28	0	2	0	5	ATXN10, FBLN1, RIBC2

References

1. van Beers EH, Joosse SA, Ligtenberg MJ, Fles R, Hogervorst FB, Verhoef S, Nederlof PM. A multiplex PCR predictor for aCGH success of FFPE samples. Br J Cancer. 2006;94(2):333-7.

2. Buffart TE, Israeli D, Tijssen M, Vosse SJ, Mrsić A, Meijer GA, Ylstra B. Across array comparative genomic hybridization: a strategy to reduce reference channel hybridizations. Genes Chromosomes Cancer. 2008;47(11):994-1004.

3. van de Wiel MA, Picard F, van Wieringen WN, Ylstra B. Preprocessing and downstream analysis of microarray DNA copy number profiles. Brief Bioinform. 2011;12(1):10-21.

4. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. Bioinformatics. 2007;23(6):657-63.

5. Leary RJ, Lin JC, Cummins J, Boca S, Wood LD, Parsons DW, Jones S, Sjöblom T, Park BH, Parsons R, Willis J, Dawson D, Willson JK, Nikolskaya T, Nikolsky Y, Kopelovich L, Papadopoulos N, Pennacchio LA, Wang TL, Markowitz SD, Parmigiani G, Kinzler KW, Vogelstein B, Velculescu VE. Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. Proc Natl Acad Sci U S A. 2008;105(42):16224-9.

6. Bierkens M, Krijgsman O, Wilting SM, Bosch L, Jaspers A, Meijer GA, Meijer CJ, Snijders PJ, Ylstra B, Steenbergen RD. Focal aberrations indicate EYA2 and hsa-

miR-375 as oncogene and tumor suppressor in cervical carcinogenesis. Genes Chromosomes Cancer. 2013;52(1):56-68. doi: 10.1002/gcc.22006.

7. Brosens RP, Haan JC, Carvalho B, Rustenburg F, Grabsch H, Quirke P, Engel AF, Cuesta MA, Maughan N, Flens M, Meijer GA, Ylstra B. Candidate driver genes in focal chromosomal aberrations of stage II colon cancer. J Pathol. 2010;221(4):411-24.

8. Futreal PA, Coin L, Marshall M, Down T, Hubbard T, Wooster R, Rahman N, Stratton MR. A census of human cancer genes. Nat Rev Cancer. 2004;4(3):177-83.

9. Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, Aerts J, Andrews TD, Barnes C, Campbell P, Fitzgerald T, Hu M, Ihm CH, Kristiansson K, Macarthur DG, Macdonald JR, Onyiah I, Pang AW, Robson S, Stirrups K, Valsesia A, Walter K, Wei J; Wellcome Trust Case Control Consortium, Tyler-Smith C, Carter NP, Lee C, Scherer SW, Hurles ME. Origins and functional impact of copy number variation in the human genome. Nature. 2010;464(7289):704-12.