Supplemental Material and Methods

Clustering

Unsupervised hierarchical clustering was done in R using Euclidean distance and Ward's clustering method. Samples were clustered based on the segmented and corrected log2 ratios in all focal and broad regions identified by GISTIC.

HRS density

HRS cells per mm^2 were automatically detected by Qupath using the following parameters: background radius 8 µm, median filter radius 0 µm, sigma 1.2 µm, minimum area 90 µm2, maximum area 400 µm², CD30 staining intensity threshold 0.2, max background intensity 2, cell expansion 2 µm, score compartment: Cell DAB OD mean. From the two-dimensional HRS density in biopsy slides, we calculated the three-dimensional HRS density in the MTV. To simplify this conversion, tumors were assumed spherical with a volume defined by the MTV, allowing the radius to be inferred ($\sqrt[3]{\frac{3}{4}\frac{MTV}{\pi}}$). Using the radius, the area of the spherical tumor was calculated ($4\pi R^2$) and multiplied by the two-dimensional HRS density to generate the number of HRS cells in that area. As the volume of a sphere can be inferred from the area (area * $\frac{1}{3}R$), this number was multiplied accordingly to generate the HRS density in the MTV.

Supplemental tables

Supplemental Table 1. Characteristics of CN studies in cHL to date.

Ref	Study	Material	Cases	Technique	Scope
16	Küppers *et al* (2001)	single HRS cell suspensions	6	FISH	*MDM2*
10	Joos *et al* (2002)	microdissected HRS cells	41	aCGH	genome-wide
11	Martin-Subero *et al* (2002)	cytogenetic suspensions	31	FICTION	*BCL11A* and *REL*
12	Chui *et al* (2003)	microdissected HRS cells	20 + 4 cell lines	aCGH	genome-wide
13	Giefing *et al* (2008)	HL cell lines	4 cell lines	aCGH + PCR	small homozygous deletions
14	Hartmann *et al* (2008)	microdissected HRS cells	10	aCGH	genome-wide
14	Hartmann *et al* (2008)	cytogenetic suspensions	28	FISH	*NOTCH1* (9q34), *STAT6* (12q13) and *JUNB* (18p13)
18	Green *et al* (2010)	microdissected HRS cells	23	qPCR	*PD-L1*
5	Steidl *et al* (2010)	microdissected HRS cells	53	aCGH	genome-wide
15	Slovak *et al* (2011)	microdissected HRS cells	27	aCGH	genome-wide
25	Nomoto *et al* (2012)	FFPE tumor biopsy	44	FICTION	*TNFAIP3*
24	Twa *et al* (2014)	FFPE tumor biopsy	20	FISH	*PD-L1* and *PD-L2*
17	Ansell *et al* (2015)	FFPE tumor biopsy	10	FISH	*PD-L1* and *PD-L2*
4	Reichel *et al* (2015)	FACS sorted HRS cells	10 + 2 cell lines	WES	genome-wide

22	Roemer *et al* (2016)	FFPE	108	FISH	*PD-L1* and *PD-L2*
21	Van Roosbroeck *et al* (2016)	FFPE tumor biopsy	200	FISH	*JAK2*
20	Juskevicius *et al* (2018)[a]	FFPE + FACS + WGA	10	aCGH	genome-wide
19	Tanaka *et al* (2018)	FFPE tumor biopsy	20	FISH	*PD-L1/PD-L2* and *CIITA*
23	Wienand *et al* (2019)	FACS sorted HRS cells	23	WES	genome-wide

[a] although this was a genome-wide study, no raw data were available; frequencies could not be accurately determined, but were estimated based on the accompanying figure.

Supplemental Table 2. FISH validation of CNA at low estimated clonal fraction.

Case	Estimated clonal fraction (%)	Observation Ichor CNA	Probes [a] (fluorochrome)(location)	# of HRS cells analyzed	FISH pattern [# of HRS cells][b]	Interpretation
HL19	2,4	gain 2p	LSI **NMYC** (SG)(2p24) / **CEP2** (SO) / CEP12 (SA)	5	5G/4R/3A[3]; 5G/4R/2A[1]; 5G/4R/6A[1]	Gain **NMYC** relative to CEP12 in 4/5 cells
		loss 6q	**MYB** (DC,BA)(6q23) [c]	13	2DC [13]	Loss of **MYB** relative to CEP12 control in 4/5 cells in paired experiment
HL26	3,4	gain 5p	**D5S23,D5S721** (SG)(5p15) / LSI *EGR1* (SO)(5q31)	4	5G/3R [1]; 4G/3R [3]	Gain of **D5S23,D5S721** relative to *EGR1* in 4/4 cells
		loss 6q	**MYB** (DC,BA)(6q23) [c]	4	2DC [2]; 1DC [2]	Loss of **MYB** in at least 2/4 cells, in 4/4 cells relative to 5p
HL15	3,6	gain 2p	LSI **ALK** (DC BA)(2p23) / CEP2 (SA)	9	3DC/2A [7]; 3DC/3A [2]	Gain of **ALK** relative to CEP2 in 7/9 cells
HL44	3,9	loss 1p	**CDKN2C** (G) (1p32) / CKS1B (TR)(1q21) /	8	4G/3R [1]; 3G/5R [2]; 4G/5R [2]; 3G/6R[1]; 4G/6R [1]; 5G/7R [1]	Loss of **CDKN2C** relative to 1q in 7/8 cells
		loss 6q	**MYB** (DC,BA)(6q23) / CEP8 (SA)[d]	5	3DC/5A [4];3DC/6A [1]	Loss of **MYB** relative to CEP8 in 5/5 cells
		gain 9p	**JAK2** (DC,BA)(9p24) / CEP8 (SA)[d]	5	2DC/2A [2]; 8-10DC/5A [1]; 11DC/5A [1]; 10-11DC/4-5A [1]	Gain of **JAK2** relative to CEP8 in 3/5 cells

HL58	4,4	gain 2p	LSI **ALK** (DC BA)(2p23) / CEP7 (SA)[a]	8	3-6DC/2-3A [8]	Gain of **ALK** relative to CEP7 in 8/8 cells
		loss 6q	**MYB** (DC,BA)(6q23) / CEP7 (SA)[d]	5	4DC/4A [3]; 2DC/4A [1];3DC/2A [1]	Loss of **MYB** relative to CEP7 in only 1/5 cells

[a] LSI: locus specific identifier; DC: dual color; BA: break apart; CEP: centromeric probe; SA: spectrum aqua (A); SG: spectrum green (G); SO: spectrum orange (R); TR: texas red (R); DC,BA: dual color break-apart. The probe in bold hybridizes with the region gained/lost in cfDNA analysis. As targets for control probes, we chose regions not affected by CNA by ichorCNA, preferentially on the chromosome of interest.

[b] The FISH patterns of the cells analyzed is shown, with the number of cells with a given pattern between square brackets. G: spectrum green signal; R: spectrum orange signal; A: spectrum aqua signal.

[c] Due to failed hybridization of CEP6 probes, the number of control signals was derived from the paired FISH experiment (CEP12 for HL19; *EGR1* for HL26).

[d] Due to technical problems with CEP6 (failed hybridization) and with CEP9 (cross-hybridization) CEP7/CEP8 were used instead.

Supplemental Table 3. Recurrent gains and losses in our patient cohort and comparison with literature.

Locus	GISTIC wider region	Size (Mb)	Encompassed genes*	n[$] (%)	Reported incidence[&] (range)	References
GAINS						
1q33.1 - q32.2	186490000 - 208010000	21,5	*IL10*, ***PIK3C2B***, ***MDM4***,	36 (22%)	19% (8-30)	5,14,20
2p16.1 - p15	58990000 - 63510000	4,5	***REL***, ***BCL11A***, *XPO1*	114 (69%)	50% (28-70)	4-5,10-12,14,20,23
4q23 - q24	99490000 - 106010000	6,5	*EIF4E*, ***NFKB1***	16 (10%)	8%	10
5p15.33 - p13.3	1 - 30010000	30	*TERT*	83 (50%)	32% (14-50)	10,20,23
8q24.13 - q24.3	123490000 - 146364022	22,9	***MYC***, *SHARPIN*	31 (19%)	34% (9-67)	5,14-15,20
9p24.3 - p24.1	1 - 6010000	6	***JAK2***, ***PD-L1***	82 (50%)	48% (23-97)	5,10,12,14,18-19, 20-22,24,
11q13.1 - q13.3	65490000 - 68500000	3	***CLCF1***, *PACS1*	31 (19%)	25% (10-40)	14,20
12q13.13 - q14.1	52490000 - 62010000	9,5	***GLI1***, ***LRP1***, ***STAT2***, ***STAT6***	82 (50%)	39% (10-67)	10,14,16,20
15q21.3 - q26.3	55490000 - 102531392	47	*FES*, *PIAS1*, *PRC1*	19 (11%)	X	X

LOSSES

1p36.33 - p36.13	1 - 19170000	19,2	*RPL22, TNFRSF4, TNFRSF8, TNFRSF14, TP73*	59 (36%)	17% (13-20)	5,20,23
1p34.1 - p21.1	45460000 - 104620000	59,2	*BCL10*	39 (23%)	13% (7-20)	5,10,20
2q35 - q37.3	217360000 - 243199373	25,8	***BOK**, **INPP5D**, **TRAF3IP1***	12 (7%)	20%	20
3p26.3 - p13	1 - 71600000	71,6	*RASSF1, TLR9, TRAIP, VHL*	29 (17%)	X	X[#]
4q34.1 - q35.2	173960000 - 191154276	17,2	*CASP3*	83 (50%)	40% (30-50)	14,20
6q22.1 - q27	116980000 - 171115067	54,1	***BCLAF1**, *SGK1*, **TNFAIP3***	79 (48%)	41% (21-60)	5,12-14,20,23,25
6q25.3 - q27	160320000 - 171115067	10,8	*DLL1*	92 (55%)	50%	20,23
7q31.31 - q34	117880000 - 141540000	23,7	*HIPK2*	47 (28%)	26% (17-30)	16,21,25
8p22.2 - p21.2	16980000 - 26610000	9,6	*CNOT7*	71 (43%)	25% (20-30)	14,20

| 11q21 - q24.3 | 93910000 - 130030000 | 36 | *EI24, TIRAP, TBRG1* | 73 (44%) | 29% (10-58) | 5,12,20 |
| 12q21.31 - q24.33 | 85410000 - 133851895 | 48,5 | *NCOR2, SH2B3, SOCS2* | 11 (7%) | X | X[#] |

* a selection of genes within the region of interest is shown, based on known or potential involvement in HL or NHL. Genes located within the MCR are indicated in bold.

$ number of cases and frequency of events including all genes in the GISTIC wider region among the aberrant diagnostic cases from our cohort (n=166).

& for events reported in more than 1 study, the average frequency and range is given.

these aberrations were reported in one case in Juskevicius *et al* (2018) and therefore not recurrent.

Supplemental Table 4. FISH validation of novel GISTIC regions.

Case	Estimated clonal fraction (%)	Observation Ichor CNA	Probes (fluorochrome)(location)	# of HRS cells analyzed	FISH pattern [# of HRS cells]	Interpretation
HL33	6.7	loss 3p	**FOXP1**(DC,BA)(3p14) / CEP7(SA)[a]	10	1DC/2A [5];2DC/3A [1];2DC/4A [4]	Loss of **FOXP1** relative to CEP7 in 10/10 cells
HL43	16,7	loss 3p	**FOXP1**(DC,BA)(3p14) / CEP8(SA)[a]	10	1 DC /3A [7];1 DC /2A [1]; 2 DC /4A [1];2 DC /8A [1]	Loss of **FOXP1** relative to CEP8 in 10/10 cells
HL13	5,5	gain 15p	LSI **PML** (SO)(15q24) / *RARA* (SG)(DC DF)(17q21)[a]	12	3R/2G [7];6R/4G [5]	Gain of **PML** relative to *RARA* in 12/12 cells
HL49	7	gain 15p	LSI **PML** (SO)(15q24) / *RARA* (SG)(DC DF)(17q21)[a]	17	3R/2G [3];4R/3G [11]; 5R/2G [1];5R/3G [2]	Gain of **PML** relative to *RARA* in 17/17 cells

[a] As targets for control probes, we chose regions not affected by CNA according to ichorCNA on other chromosomes, as the centromeric region of the chromosome with CNA was also affected by this CNA.

Supplemental Figures
Supplemental Figure 1
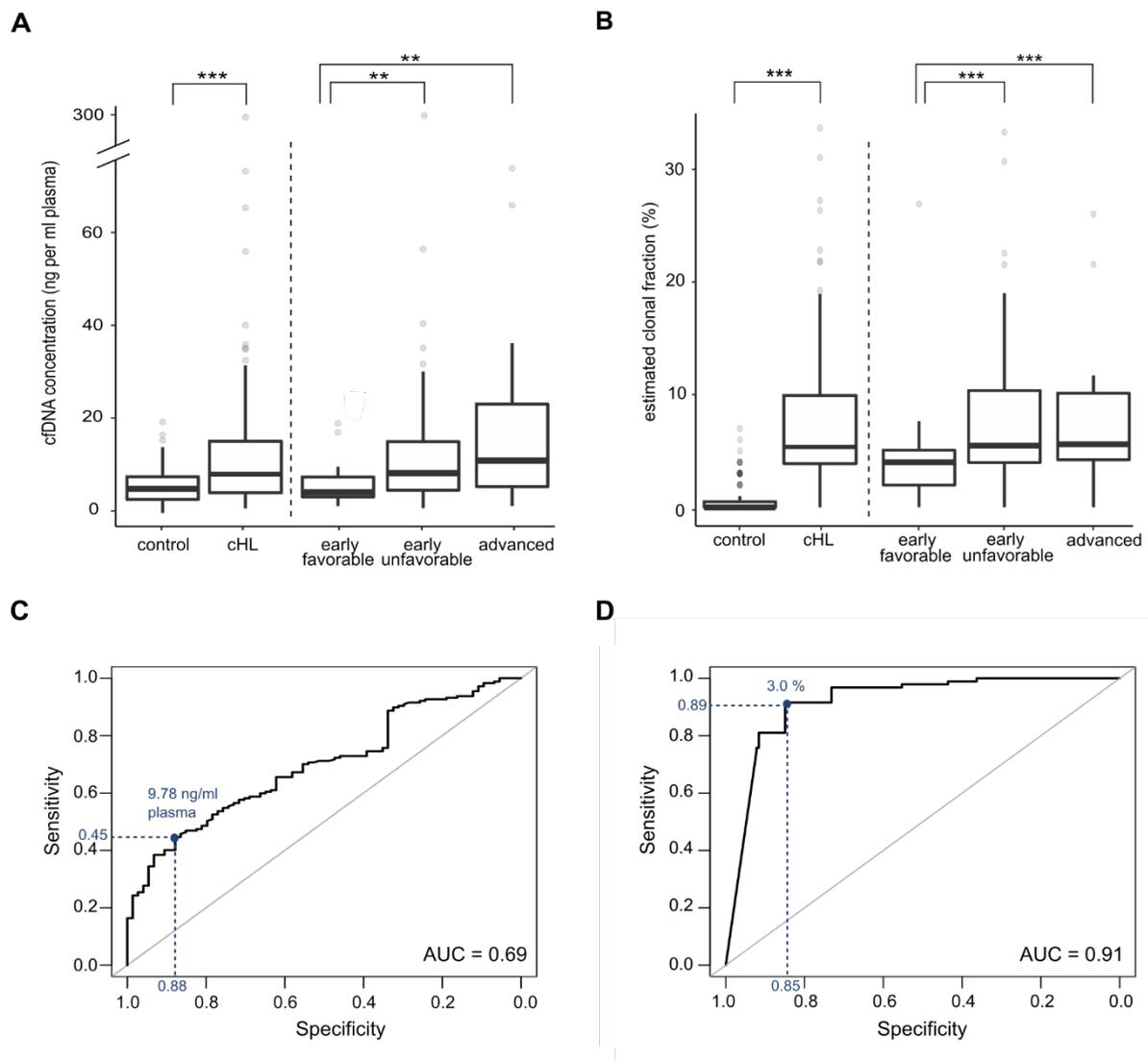
Supplemental Figure 1. **cfDNA and estimated clonal fractions in cHL patients versus controls.** (A-B) Box and whisker plots for cfDNA concentrations (A) and estimated clonal fractions (B) in controls versus cHL patients, grouped by disease category. Outliers are defined as values 1.5 x IQR below the first or above the third quartile. *P < .05, **P < .01, ***P < .001. (C-D) ROC analysis for total cfDNA levels (C) and estimated clonal fractions (D) determines the optimal cut-off between cHL cases and controls based on the Youden index (blue). Corresponding sensitivity and specificity are indicated by dashed lines in blue. AUC: area under the curve.

Supplemental Figure 2
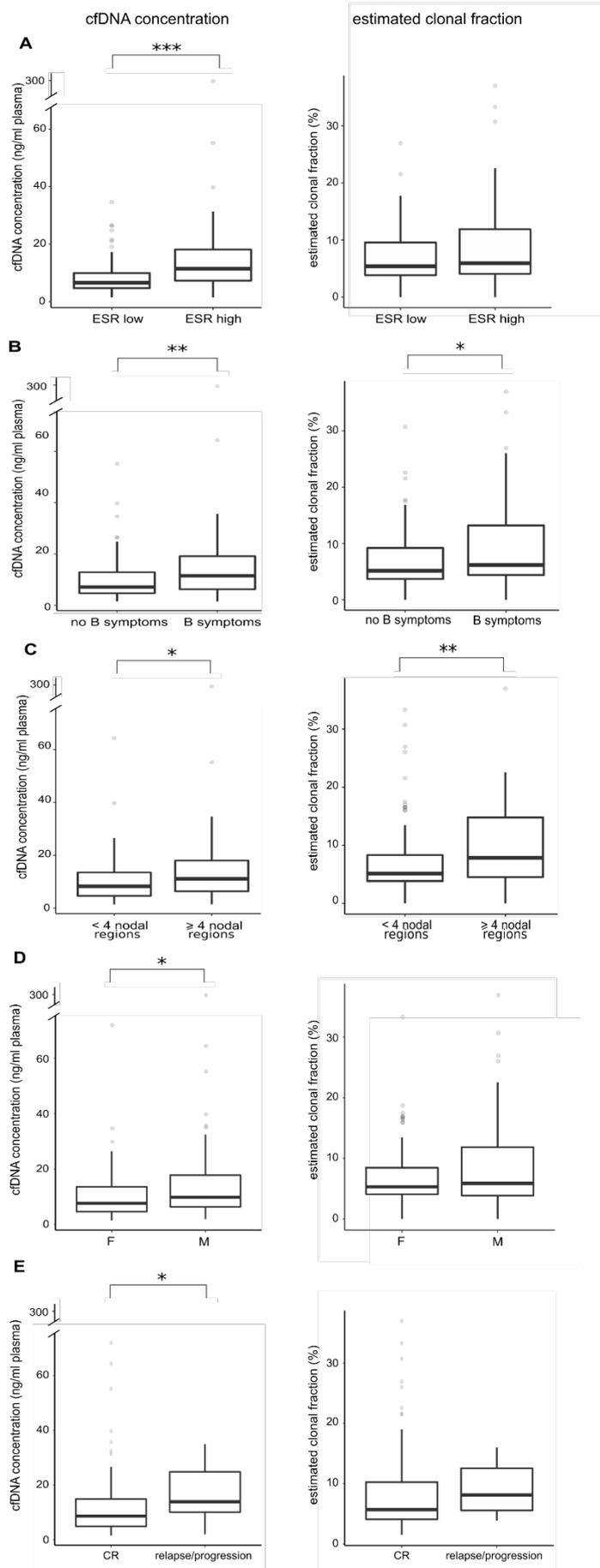
Supplemental Figure 2. **Association between clinical parameters and total cfDNA levels and estimated clonal fractions in cHL**. Box and whisker plots of total cfDNA levels (left) and estimated clonal fractions (right) at diagnosis for cases with (A) normal versus elevated ESR levels; (B) presence versus absence of B symptoms; (C) limited versus extensive nodal involvement; (D) female versus male gender; (E) treatment response versus treatment failure. Outliers are defined as values 1.5 x IQR below the first or above the third quartile. *P < .05, **P < .01, ***P < .001.

A

B
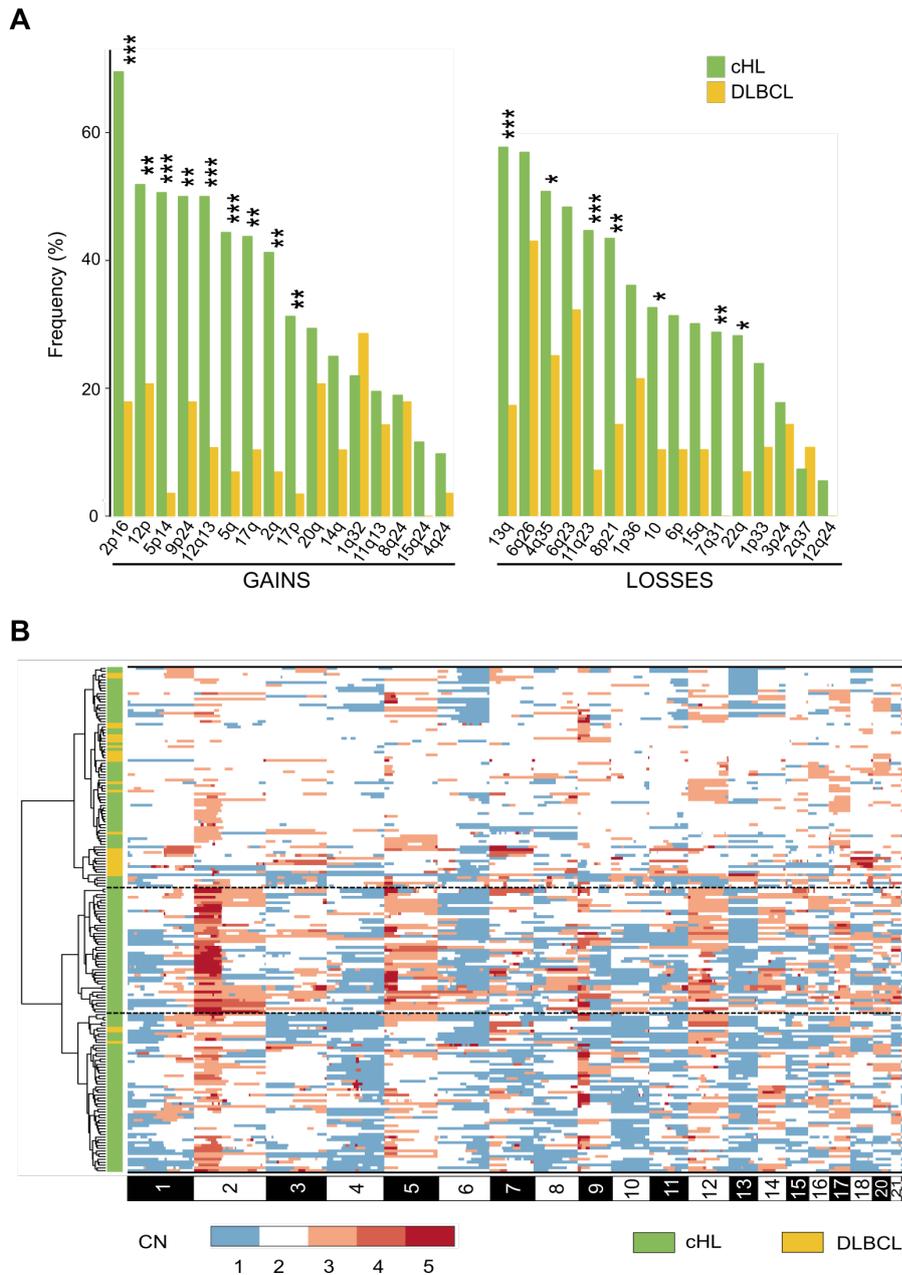
Figure 3. CNA patterns in HL and DLBCL. (A) Bar plot of the frequencies at which recurrent gains and losses in HL were observed in HL and DLBCL cases. CNA are plotted in descending order and statistically interpreted using Fischer's exact test. *P < 0.05, **P < 0.01, ***P < 0.001. (B) Heat map of hierarchically clustered abnormal cHL (n= 164) and DLBCL (n=29) CNA profiles. Rows represent genome-wide CN profiles (blue = loss, white = copy neutral, red = gain/amplification). Samples are annotated according to disease (green = cHL, yellow = DLBCL). Unsupervised hierarchical clustering of CNA profiles segregated cHL from DLBCL in cases with higher clonal fractions (lower part of the graph). In cases with low clonal fraction, segregation was less successful (the upper third of the graph).

Supplemental Figure 4
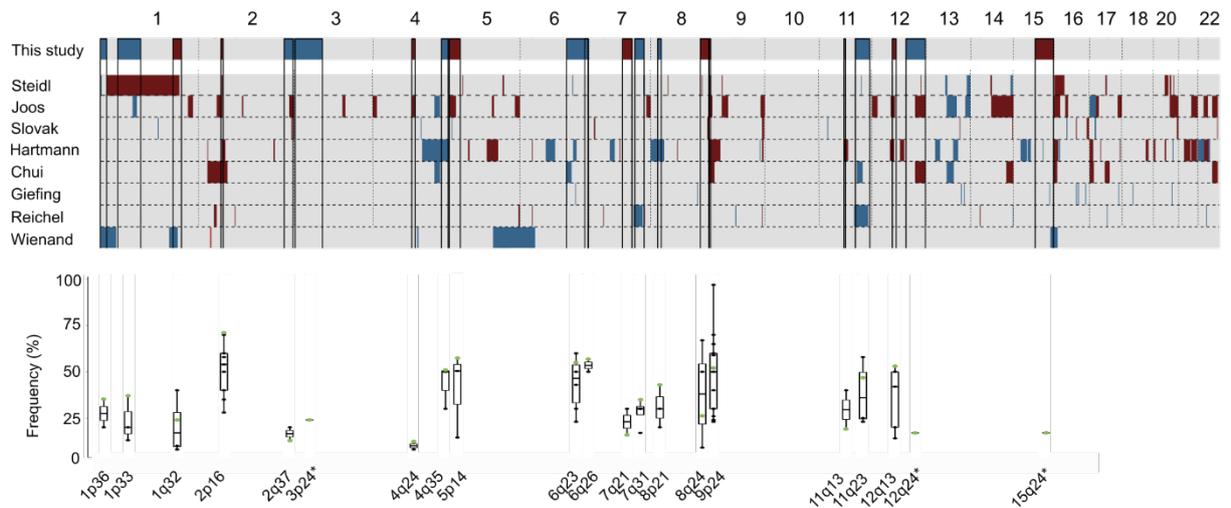
Supplemental Figure 4. **Recurrent CNA in previous studies compared to our cohort**. Heatmap showing recurrent gains (red) and losses (blue) as described in this study and in previous genome-wide CNA studies in cHL with accessible data. Recurrent CNA in our cohort are delineated by black rectangles for comparison purposes. The range of frequencies at which these regions have been reported in previous studies are plotted (see also Supplemental Table 1 and 2). The recurrence in our cohort is indicated in green. *Three regions that are reported for the first time in our cohort.

Supplemental figure 5
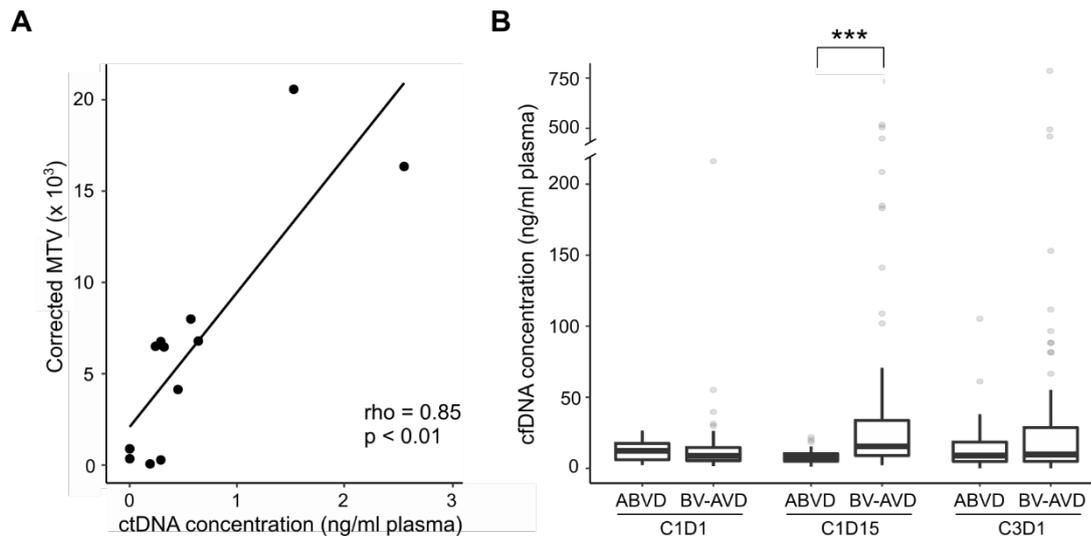
Supplemental Figure 5. **ctDNA as a tool for disease follow-up**. (A) Correlation between HRS density-corrected MTV and ctDNA concentration. MTV is corrected for HRS density by multiplying the MTV with the number of HRS cells per unit of volume. (B) Box and whisker plots of cfDNA concentrations in ABVD versus BV-AVD arms at diagnosis, C1D15 and C3D1. ***P < .001