

# Sussex Research

## A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia

Helen E Speedy, Maria Chiara Di Bernardo, Georgina P Sava, Martin J S Dyer, Amy Holroyd, Yufei Wang, Nicola J Sunter, Larry Mansouri, Gunnar Juliusson, Karin E Smedby, Göran Roos, Sandrine Jayne, Aneela Majid, Claire Dearden, Andrew G Hall, Tryfonia Mainou-Fowler, Graham H Jackson, Geoffrey Summerfield, Robert J Harris, Andrew R Pettitt, David J Allsup, James R Bailey, Guy Pratt, Christopher Pepper, Chris Fegan, Richard Rosenquist, Daniel Catovsky, James M Allan, Richard S Houlston

### Publication date

09-06-2023

### Licence

This work is made available under the **Copyright not evaluated** licence and should only be used in accordance with that licence. For more information on the specific terms, consult the repository record for this item.

### Document Version

Accepted version

### Citation for this work (American Psychological Association 7th edition)

Speedy, H. E., Di Bernardo, M. C., Sava, G. P., Dyer, M. J. S., Holroyd, A., Wang, Y., Sunter, N. J., Mansouri, L., Juliusson, G., Smedby, K. E., Roos, G., Jayne, S., Majid, A., Dearden, C., Hall, A. G., Mainou-Fowler, T., Jackson, G. H., Summerfield, G., Harris, R. J., ... Houlston, R. S. (2013). *A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia* (Version 1). University of Sussex. <https://hdl.handle.net/10779/uos.23455808.v1>

### Published in

Nature Genetics

### Link to external publisher version

<https://doi.org/10.1038/ng.2843>

### Copyright and reuse:

This work was downloaded from Sussex Research Open (SRO). This document is made available in line with publisher policy and may differ from the published version. Please cite the published version where possible. Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners unless otherwise stated. For more information on this work, SRO or to report an issue, you can contact the repository administrators at [sro@sussex.ac.uk](mailto:sro@sussex.ac.uk). Discover more of the University's research at <https://sussex.figshare.com/>

## **A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia**

Helen E Speedy<sup>1\*</sup>, Maria Chiara Di Bernardo<sup>1\*</sup>, Georgina P Sava<sup>1</sup>, Martin J S Dyer<sup>2</sup>, Amy Holroyd<sup>1</sup>, Yufei Wang<sup>1</sup>, Nicola J Sunter<sup>3</sup>, Larry Mansouri<sup>4</sup>, Gunnar Juliusson<sup>5</sup>, Karin E Smedby<sup>6</sup>, Göran Roos<sup>7</sup>, Sandrine Jayne<sup>8</sup>, Aneela Majid<sup>8</sup>, Claire Dearden<sup>9</sup>, Andrew G Hall<sup>3</sup>, Tryfonia Mainou-Fowler<sup>10</sup>, Graham H Jackson<sup>11</sup>, Geoffrey Summerfield<sup>12</sup>, Robert J Harris<sup>13</sup>, Andrew R Pettitt<sup>13</sup>, David J Allsup<sup>14</sup>, James R Bailey<sup>15</sup>, Guy Pratt<sup>16</sup>, Chris Pepper<sup>17</sup>, Chris Fegan<sup>18</sup>, Richard Rosenquist<sup>4</sup>, Daniel Catovsky<sup>9</sup>, James M Allan<sup>3</sup>, Richard S Houlston<sup>1+</sup>

1. Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton, Surrey, UK
2. Department of Cancer Studies and Molecular Medicine, University of Leicester, Leicester, UK
3. Northern Institute for Cancer Research, Newcastle University, Newcastle upon Tyne, UK
4. Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden.
5. Lund Strategic Research Center for Stem Cell Biology and Cell Therapy, Hematology and Transplantation, Lund University, Lund, Sweden
6. Unit of Clinical Epidemiology, Department of Medicine, Karolinska Institutet, Stockholm, Sweden
7. Department of Medical Biosciences, Pathology, Umeå University, Umeå, Sweden
8. Medical Research Council Toxicology Unit, Leicester University, Leicester, UK
9. Haemato-Oncology, Division of Pathology, Institute of Cancer Research, Sutton, Surrey, UK
10. Haematological Sciences, Medical School, Newcastle University, Newcastle-upon-Tyne, UK
11. Department of Haematology, Royal Victoria Infirmary, Newcastle-upon-Tyne, UK
12. Department of Haematology, Queen Elizabeth Hospital, Gateshead, Newcastle-upon-Tyne, UK
13. Department of Molecular and Clinical Cancer Medicine, University of Liverpool, Liverpool, UK.
14. Department of Haematology, Hull Royal Infirmary, Hull, UK

15. Hull York Medical School and University of Hull, Hull, UK
16. Department of Haematology, Birmingham Heartlands Hospital, Birmingham, UK
17. Department of Haematology, School of Medicine, Cardiff University, Cardiff, UK
18. Cardiff and Vale National Health Service Trust, Heath Park, Cardiff, UK

\*Equal authorship at this position.

<sup>+</sup>Corresponding author:

E-mail: [Richard.houlston@icr.ac.uk](mailto:Richard.houlston@icr.ac.uk); Tel: +44 (0) 208-722-4175; Fax: +44 (0) 208-722-4365

**Supplementary Table 1: Evidence for association at previously reported CLL susceptibility loci.** Odds ratios and *P*-values shown are derived from the harmonised UK-CLL-2 and UK-CLL-1 datasets. At each locus values are given for the previously reported SNP (or a proxy) and the lead SNP in this study.

Locus	Nearest Gene(s)	Published SNP	Lead SNP in harmonised data <sup>a</sup>	LD ( <i>r</i> <sup>2</sup> ) <sup>b</sup>	Position (hg19)	Alleles <sup>c</sup>	RAF <sup>d</sup>	Individual datasets <sup>e</sup>					Meta-analysis			
								UK-CLL-2		UK-CLL-1			Fixed effects		Random effects	
								OR	<i>P</i> <sub>trend</sub>	OR	<i>P</i> <sub>trend</sub>	IS	OR	<i>P</i> <sub>trend</sub>	OR	<i>P</i> <sub>trend</sub>
2p22.2	<i>QPCT, PRKD3</i>	rs3770745			37,596,089					SNP absent from OEE array						
		rs6734118 <sup>f</sup>		0.45	37,559,355	A/C	0.21	1.17	6.90E-03	1.15	8.20E-02	DT	1.17	1.33E-03	1.17	1.33E-03
			rs2041840	0.11	37,467,264	C/T	0.62	1.24	2.72E-05	1.13	7.40E-02	1.00	1.20	9.23E-06	1.20	4.36E-05
2q13	<i>ACOXL, BCL2L11</i>	rs17483466			111,797,458	G/A	0.21	1.36	2.23E-07	1.43	1.52E-05	DT	1.38	1.72E-11	1.38	1.72E-11
		rs13401811			111,616,104					SNP failed QC						
		rs13395354 <sup>f</sup>		1.00	111,600,519	C/T	0.82	1.29	6.91E-05	1.44	5.22E-05	1.00	1.34	2.24E-08	1.34	2.24E-08
			rs1439287	0.02	111,871,897	A/G	0.48	1.40	1.08E-11	1.31	7.60E-05	1.00	1.37	5.16E-15	1.37	5.16E-15
2q33.1	<i>CASP10/CASP8</i>	rs3769825			202,111,380	A/G	0.43	1.09	7.52E-02	1.16	3.53E-02	DT	1.11	7.69E-03	1.11	7.69E-03
			rs1035142	0.13	202,153,078	T/G	0.36	1.13	1.09E-02	1.26	1.06E-03	1.00	1.18	7.61E-05	1.18	1.41E-03
2q37.1	<i>SP110, SP140</i>	rs13397985	rs13397985		231,091,223	G/T	0.19	1.47	2.64E-10	1.36	3.13E-04	DT	1.43	4.55E-13	1.43	4.55E-13
2q37.3	<i>FARP2</i>	rs757978	rs757978		242,371,101	T/C	0.11	1.15	6.22E-02	1.48	3.68E-04	DT	1.25	3.47E-04	1.29	3.73E-02
4q25	<i>LEF1</i>	rs898518			109,016,824	A/C	0.59	1.07	1.90E-01	1.23	3.01E-03	DT	1.12	5.02E-03	1.14	6.51E-02
			rs12503104	0.44	109,062,510	T/C	0.45	1.09	7.90E-02	1.25	2.02E-03	0.88	1.14	1.46E-03	1.16	3.41E-02
6p25.3	<i>IRF4</i>	rs872071	rs872071		411,064	G/A	0.51	1.37	2.42E-10	1.43	1.87E-07	DT	1.39	2.74E-16	1.39	2.74E-16
6p21.32	<i>HLA</i>	rs674313			32,578,082					SNP absent from OEE array						
		rs9273012 <sup>f</sup>		0.52	32,611,641	G/A	0.26	1.19	1.48E-03	1.15	1.24E-01	0.79	1.18	4.34E-04	1.18	4.34E-04
			rs926070	0.00	32,257,566	A/G	0.66	1.32	1.84E-07	1.17	3.20E-02	DT	1.27	3.97E-08	1.26	5.84E-05
6p21.31	<i>BAK1</i>	rs210134	rs210134		33,540,209	G/A	0.68	1.32	1.98E-07	1.30	3.79E-04	DT	1.31	3.02E-10	1.31	3.02E-10
8q24.21	<i>POU5F1B</i>	rs2456449			128,192,981					SNP absent from OEE array						
		rs2466024 <sup>f</sup>	rs2466024	0.71	128,188,019	A/G	0.41	1.19	4.02E-04	1.24	2.26E-03	DT	1.21	3.24E-06	1.21	3.24E-06
9p21.3	<i>CDKN2B-AS1</i>	rs1679013			22,206,987					SNP failed QC						
		rs1359741 <sup>f</sup>		0.52	22,336,954	G/A	0.54	1.07	1.54E-01	1.15	4.30E-02	DT	1.10	1.93E-02	1.10	1.93E-02
			rs1609591	0.03	22,291,931	C/A	0.09	1.22	1.99E-02	1.33	1.78E-02	0.97	1.25	1.08E-03	1.25	1.08E-03
10q23.31	<i>ACTA, FAS</i>	rs4406737			90,759,724	G/A	0.57	1.25	1.06E-05	1.19	1.43E-02	DT	1.23	5.54E-07	1.23	5.54E-07
			rs1800682	0.43	90,749,963	A/G	0.54	1.36	5.81E-10	1.07	3.32E-01	0.99	1.25	1.96E-08	1.21	1.07E-01
11p15.5	<i>C11orf21, TSPAN32</i>	rs7944004			2,311,152	T/G	0.50	1.16	2.26E-03	1.04	6.25E-01	0.78	1.10	2.09E-02	1.07	4.83E-01
			rs11022157	0.40	2,322,829	A/C	0.36	1.31	1.66E-07	1.04	5.85E-01	0.84	1.22	2.55E-06	1.18	1.41E-01

Locus	Nearest Gene(s)	Published SNP	Lead SNP in harmonised data <sup>a</sup>	LD ( $r^2$ ) <sup>b</sup>	Position (hg19)	Alleles <sup>c</sup>	RAF <sup>d</sup>	Individual datasets <sup>e</sup>					Meta-analysis			
								UK-CLL-2		UK-CLL-1			Fixed effects		Random effects	
								OR	$P_{\text{trend}}$	OR	$P_{\text{trend}}$	IS	OR	$P_{\text{trend}}$	OR	$P_{\text{trend}}$
11q24.1	<i>SCN3B, GRAMD1B</i>	rs735665	rs735665		123,361,397	A/G	0.19	1.71	5.29E-19	1.51	7.02E-07	DT	1.64	4.42E-24	1.63	1.17E-15
15q15.1	<i>BMF, BUB1B</i>	rs8024033			40,403,657			SNP absent from OEE array								
		rs11637681 <sup>f</sup>		0.36	40,387,971	A/G	0.73	1.21	7.78E-04	1.20	5.31E-02	0.66	1.20	1.06E-04	1.20	1.06E-04
			rs8023845	0.28	40,414,116	G/T	0.78	1.28	6.08E-05	1.25	8.59E-03	DT	1.27	1.65E-06	1.27	1.65E-06
15q21.3	<i>RFX7, NEDD4</i>	rs7169431			56,340,896	A/G	0.08	1.36	3.83E-04	1.52	7.04E-04	DT	1.41	1.21E-06	1.41	1.21E-06
			rs16976734	0.30	56,382,295	G/A	0.11	1.33	1.34E-04	1.45	6.63E-04	0.95	1.36	3.97E-07	1.36	3.97E-07
15q23	<i>RPLP1</i>	rs7176508	rs7176508		70,018,990	A/G	0.38	1.42	1.95E-12	1.42	7.76E-07	DT	1.42	8.00E-18	1.42	8.00E-18
15q25.2	<i>CPEB1</i>	rs783540			83,254,708			SNP absent from OEE array								
		rs11631963 <sup>f</sup>		1.00	83,318,202	T/C	0.38	1.10	6.96E-02	1.23	2.86E-03	1.00	1.14	1.29E-03	1.15	1.46E-02
			rs7163848	0.08	83,398,315	C/T	0.79	1.17	9.12E-03	1.32	1.34E-03	0.98	1.22	7.16E-05	1.23	5.53E-04
16q24.1	<i>IRF8</i>	rs305061			85,975,659	T/C	0.67	1.11	5.40E-02	1.38	1.10E-05	DT	1.19	3.64E-05	1.23	6.17E-02
			rs1044873	0.00	85,955,671	C/T	0.61	1.36	1.82E-09	1.16	3.68E-02	0.97	1.29	1.02E-09	1.27	2.67E-03
18q21.32	<i>PMAIP1</i>	rs4368253			57,622,287			SNP failed QC								
		rs7231647 <sup>f</sup>		0.81	57,628,926		0.58	1.23	4.80E-05	1.10	1.78E-01	DT	1.18	4.41E-05	1.18	3.21E-03
			rs7240884	0.14	57,568,693		0.58	1.19	5.14E-04	1.20	9.89E-03	DT	1.19	1.52E-05	1.19	1.52E-05
18q21.33	<i>BCL2</i>	rs4987855			60,793,549	C/T	0.90	1.24	1.00E-02	1.21	1.15E-01	0.95	1.23	2.59E-03	1.23	2.59E-03
		rs4987852	rs4987852	0.01	60,793,921	C/T	0.07	1.31	3.33E-03	1.44	7.09E-03	DT	1.35	8.12E-05	1.35	8.12E-05
19q13.3	<i>PRKD2, STRN4</i>	rs11083846			47,207,654			SNP absent from OEE array								
		rs4802322 <sup>f</sup>		0.91	47,242,992	A/G	0.25	1.15	1.44E-02	1.39	4.36E-05	0.99	1.22	1.48E-05	1.25	2.06E-02
			rs3027957	0.48	47,287,507	C/T	0.20	1.20	2.31E-03	1.36	8.79E-04	0.90	1.25	1.25E-05	1.25	1.19E-04

DT, directly typed; CI, confidence intervals; IS, imputation quality score; LD, linkage disequilibrium; OEE, OmniExpressExome; OR, odds ratio; QC, quality control; RAF, risk allele frequency

<sup>a</sup> Lead SNP= SNP within 500kb of published SNP that returned lowest  $P$ -value in meta-analysis of UK-CLL-2 and UK-CLL-1

<sup>b</sup> LD is based upon 1000 Genomes CEU data

<sup>c</sup> Alleles are given as risk allele/other allele

<sup>d</sup> RAF is from UK-CLL-2 controls

<sup>e</sup> Imputation quality score is provided, where applicable, for UK-CLL-1 data (scale: 0 - 1, SNPs with a score  $\leq 0.4$  were excluded). All UK-CLL-2 SNPs were directly genotyped in harmonised dataset

<sup>f</sup> Proxy for the published SNP

**Supplementary Table 2: Loci showing evidence of an association with CLL risk at  $P < 1 \times 10^{-5}$ .** Odds ratios and  $P$ -values for lead SNPs in the harmonised UK-CLL-2 and UK-CLL-1 datasets are listed. Also shown are minor allele frequencies in cases and controls for the individual datasets and genes mapping within a 50kb span of the lead SNP. Genome-wide significant SNPs recovered by 1000 Genomes imputation are indicated.

Locus	Lead SNP <sup>a</sup>	Position (hg19)	UK-CLL-2				UK-CLL-1				Combined		Nearest gene(s)	Imputed SNPs at GWS
			MAF cases	MAF controls	OR	$P_{\text{trend}}$	MAF cases	MAF controls	OR	$P_{\text{trend}}$	OR	$P_{\text{trend}}$		
1q24.3	rs12404183	171,058,946	0.08	0.06	1.40	7.82E-04	0.10	0.06	1.80	2.79E-05	1.53	2.27E-07	<i>FMO3</i>	None
2p21	rs6716622	45,981,426	0.06	0.04	1.51	4.28E-04	0.06	0.04	1.71	1.33E-03	1.58	2.28E-06	<i>PRKCE</i>	None
3q22.3	rs13072691	136,425,600	0.14	0.11	1.25	3.20E-03	0.16	0.11	1.55	3.16E-05	1.34	1.38E-06	<i>STAG1</i>	None
3q26.2	rs10936599	169,492,101	0.21	0.25	1.25	2.08E-04	0.20	0.25	1.24	6.08E-03	1.25	3.96E-06	<i>MYNN, ACTRT3, TERC, LRRC34</i>	None
4q26	rs1476569	114,698,696	0.32	0.27	1.30	2.30E-06	0.31	0.27	1.18	3.15E-02	1.25	3.54E-07	<i>CAMK2D</i>	rs6858698 <sup>b</sup>
5p15.33	rs31490	1,344,458	0.48	0.44	1.19	6.38E-04	0.48	0.43	1.22	3.95E-03	1.20	8.31E-06	<i>CLPTM1L</i>	None
6q12	rs9346246	69,520,926	0.25	0.29	1.20	4.60E-04	0.25	0.29	1.23	5.66E-03	1.22	8.33E-06	<i>BAI3</i>	None
6q25.2	rs2236256	154,478,440	0.49	0.44	1.22	7.48E-05	0.53	0.45	1.37	6.00E-06	1.27	4.89E-09	<i>IPCEF1</i>	rs4869818 <sup>c</sup> , rs9383692 <sup>d</sup>
7q31.33	rs17246404	124,462,661	0.24	0.29	1.27	2.60E-05	0.26	0.29	1.16	5.22E-02	1.23	5.85E-06	<i>POT1</i>	None
11p14.1	rs1425172	30,056,918	0.24	0.20	1.28	4.59E-05	0.24	0.21	1.18	5.38E-02	1.25	8.51E-06	<i>KCNA4</i>	None
19q13.11	rs8112179	34,122,429	0.27	0.31	1.23	7.70E-05	0.30	0.33	1.19	2.86E-02	1.22	6.82E-06	<i>CHST8</i>	None
20p13	rs6082956	288,996	0.48	0.43	1.23	2.49E-05	0.46	0.43	1.14	5.34E-02	1.20	5.18E-06	<i>ZCCHC3, SOX12, C20orf96</i>	None
20q13.33	rs6062501	62,341,931	0.15	0.20	1.43	1.14E-07	0.17	0.20	1.26	8.78E-03	1.36	6.36E-09	<i>ZGPAT, ARFRP1, TNFRSF6B, RTEL1</i>	None

GWS; genome-wide significance; MAF, minor allele frequency; OR, odds ratio

<sup>a</sup> In combined analysis of harmonised UK-CLL-2 and UK-CLL-1 datasets

<sup>b</sup> Combined analysis, OR=1.40,  $P_{\text{trend}}$ =2.87E-08; imputation quality scores, UK-CLL-2=0.77, UK-CLL-1=0.76

<sup>c</sup> Combined analysis, OR=1.28,  $P_{\text{trend}}$ =7.05E-10; imputation quality scores, UK-CLL-2=0.99, UK-CLL-1=0.99

<sup>d</sup> Combined analysis, OR=1.28,  $P_{\text{trend}}$ =1.63E-09; imputation quality scores, UK-CLL-2=1.00, UK-CLL-1=0.99

**Supplementary Table 3: Genotype counts for SNPs taken forward to replication stage.** Also shown are odds ratios and 95% confidence interval for each SNP across the four datasets

				GWAS datasets												UK-CLL-2 + UK-CLL-1 combined			
				UK-CLL-2						UK-CLL-1						Heterogeneity		Meta-analysis <sup>c</sup>	
SNP	Chr	Position (hg19)	Alleles <sup>a</sup>	G/I <sup>b</sup>	<i>P</i> <sub>trend</sub>	OR (95% CI)	N <sub>Ca</sub>	N <sub>Co</sub>	G/I <sup>b</sup>	<i>P</i> <sub>trend</sub>	OR (95% CI)	N <sub>Ca</sub>	N <sub>Co</sub>	<i>I</i> <sup>2</sup> (%)	<i>P</i> <sub>het</sub>	<i>P</i> <sub>trend</sub>	OR (95% CI)		
rs10936599	3	169,492,101	C/T	G	2.08E-04	1.25	CC	774	1413	G	6.08E-03	1.24	CC	319	1547	0	0.99	3.96E-06	1.25
						(1.11-1.40)	CT	407	936			(1.06-1.45)	CT	162	978				(1.13-1.37)
							TT	55	152				TT	22	173				
rs6858698	4	114,683,844	C/G	I	1.02E-04	1.34	CC	43	69	I	3.88E-05	1.54	CC	20	72	19	0.27	2.87E-08	1.40
						(1.15-1.55)	CG	376	638			(1.25-1.90)	CG	164	704				(1.24-1.58)
							GG	817	1794				GG	319	1922				
rs31490	5	1,344,458	A/G	G	6.38E-04	1.19	AA	270	454	I	3.95E-03	1.22	AA	113	519	0	0.75	8.31E-06	1.20
						(1.08-1.31)	AG	638	1272			(1.07-1.40)	AG	260	1305				(1.11-1.30)
							GG	326	773				GG	130	874				
rs2236256	6	154,478,440	C/A	G	7.48E-05	1.22	CC	282	498	G	6.00E-06	1.37	CC	143	521	49	0.16	4.89E-09	1.27
						(1.10-1.34)	AC	648	1217			(1.20-1.57)	AC	244	1388				(1.17-1.37)
							AA	304	785				AA	116	789				
rs17246404	7	124,462,661	C/T	G	2.60E-05	1.27	CC	699	1256	I	5.22E-02	1.16	CC	280	1373	0	0.34	5.85E-06	1.23
						(1.14-1.42)	CT	447	1040			(1.00-1.35)	CT	188	1104				(1.12-1.34)
							TT	73	204				TT	35	221				
rs6062501	20	62,341,931	T/G	G	1.14E-07	1.43	TT	891	1599	G	8.78E-03	1.26	TT	348	1706	21	0.26	6.36E-09	1.36
						(1.25-1.63)	TG	318	800			(1.06-1.50)	TG	143	899				(1.23-1.51)
							GG	25	99				GG	12	93				

				Replication datasets											All datasets combined				
				UK replication						Swedish replication					Heterogeneity		Meta-analysis <sup>c</sup>		
SNP	Chr	Position (hg19)	Alleles <sup>a</sup>	G/I <sup>b</sup>	<i>P</i> <sub>trend</sub>	OR (95% CI)	N <sub>Ca</sub>	N <sub>Co</sub>	G/I <sup>b</sup>	<i>P</i> <sub>trend</sub>	OR (95% CI)	N <sub>Ca</sub>	N <sub>Co</sub>	<i>I</i> <sup>2</sup> (%)	<i>P</i> <sub>het</sub>	<i>P</i> <sub>trend</sub>	OR (95% CI)		
rs10936599	3	169,492,101	C/T	G	1.25E-03	1.26	CC	503	1575	G	3.02E-02	1.31	CC	199	0	0.99	1.74E-09	1.26	
						(1.09-1.45)	CT	252	1044			(1.03-1.67)	CT	130				140	(1.17-1.35)
							TT	33	141				TT	12				32	
rs6858698	4	114,683,844	C/G	G	5.30E-02	1.16	CC	20	58	G	3.79E-02	1.39	CC	6	39	0.18	3.07E-09	1.31	
						(1.00-1.36)	CG	226	731			(1.02-1.90)	CG	91				80	(1.20-1.44)
							GG	516	1958				GG	225				285	
rs31490	5	1,344,458	A/G	G	2.10E-03	1.19	AA	173	507	G	8.52E-01	1.02	AA	69	0	0.54	1.72E-07	1.18	
						(1.07-1.33)	AG	416	1325			(0.83-1.26)	AG	176				182	(1.11-1.26)
							GG	213	877				GG	96				111	
rs2236256	6	154,478,440	C/A	G	2.43E-03	1.19	CC	192	555	G	5.92E-01	1.06	CC	74	34	0.21	1.50E-10	1.23	
						(1.06-1.34)	AC	368	1350			(0.86-1.31)	AC	176				176	(1.15-1.30)
							AA	194	801				AA	87				103	
rs17246404	7	124,462,661	C/T	G	8.26E-03	1.19	CC	442	1416	G	7.55E-02	1.24	CC	203	0	0.75	3.40E-08	1.22	
						(1.05-1.35)	CT	297	1087			(0.98-1.57)	CT	111				152	(1.14-1.31)
							TT	47	229				TT	24				26	
rs6062501	20	62,341,931	T/G	G	7.36E-01	1.03	TT	529	1797	G	3.13E-01	0.88	TT	207	83 <sup>d</sup>	0.001 <sup>d</sup>	1.47E-05	1.19	
						(0.89-1.19)	TG	246	851			(0.68-1.13)	TG	112				117	(1.10-1.29)
							GG	24	88				GG	19				15	

Ca, cases; Chr, chromosome; CI, confidence intervals; Co, controls; OR, odds ratio

<sup>a</sup> Alleles given as risk allele/other allele

<sup>b</sup> G=SNP genotyped in given dataset; I=SNP imputed in given dataset

<sup>c</sup> Inverse-variance, fixed effects model

<sup>d</sup> Evidence of between-study heterogeneity. Under a random effects model, *P*<sub>trend</sub>=0.19; combined OR=1.14, 95%CI =0.94-1.40.



**Supplementary Table 4: Summary of genomic annotation by HaploReg v2 and RegulomeDB.** Data are shown for GWAS hits and proxy SNPs ( $r^2 > 0.8$  in 1000Genomes EUR phase 1 data) demonstrating evidence of histone marks, DNase hypersensitivity sites or transcription factor occupancy in HaploReg v2 analysis. Also indicated are GERP scores  $> 2$  and RegulomeDB scores for all SNPs.

SNP <sup>a</sup>	Chr	Pos (hg19)	LD <sup>b</sup>	GERP <sup>c</sup>	Histone marks		DNase HS	Bound proteins	Altered motifs	RefSeq genes	dbSNP annot	RegulomeDB score <sup>d</sup>
					Promoter	Enhancer						
rs2293607	3	169482335	0.81		8 cell types <sup>e</sup>	1 cell type	22 cell types <sup>e</sup>	CMYC, ELF1, NFKB, OCT2 + 10	4	62bp 3' of <i>TERC</i>		4
rs12637184	3	169487437	0.81		8 cell types <sup>e</sup>		19 cell types <sup>e</sup>	CTCF, ELF1, POL2, TAF1 + 12	19	<i>ARPM1</i>	5'-UTR	4
rs3821383	3	169489946	0.90		8 cell types <sup>e</sup>	1 cell type	57 cell types <sup>e</sup>	USF1, ETS1, P300, TBP + 18		906bp 5' of <i>MYNN</i>		2b
rs2251795	3	169491729	0.83		8 cell types <sup>e</sup>	1 cell type		HAE2F1	1	<i>MYNN</i>	intronic	3a
<b>rs10936599</b>	<b>3</b>	<b>169492101</b>	<b>1</b>	<b>3.9</b>	<b>3 cell types<sup>e</sup></b>	<b>6 cell types</b>	<b>5 cell types</b>		<b>3</b>	<b><i>MYNN</i></b>	<b>syn</b>	<b>5</b>
rs3950296	3	169493283	1			4 cell types <sup>e</sup>			2	<i>MYNN</i>	intronic	5
rs13069553	3	169508272	1				2 cell types		3	767bp 3' of <i>MYNN</i>		5
rs7625734	3	169508915	0.90				7 cell types	MAFK	6	1.4kb 3' of <i>MYNN</i>		3a
rs7633750	3	169509244	0.90				2 cell types		4	1.7kb 3' of <i>MYNN</i>		5
rs10936601	3	169528449	0.90				1 cell type		3	<i>LRRC34</i>	intronic	5
rs9831661	3	169528523	0.90				1 cell type		2	<i>LRRC34</i>	intronic	5
<b>rs6858698</b>	<b>4</b>	<b>114683844</b>	<b>1</b>	<b>2.4</b>	<b>9 cell types<sup>e</sup></b>		<b>116 cell types<sup>e</sup></b>	<b>ELF1, IRF4, MAX, NFKB + 29</b>	<b>4</b>	<b>760bp 5' of <i>CAMK2D</i></b>		<b>2b</b>
<b>rs10069690</b>	<b>5</b>	<b>1279790</b>	<b>1</b>				<b>4 cells types</b>		<b>2</b>	<b><i>TERT</i></b>	<b>intronic</b>	<b>5</b>
rs4975616	5	1315660	0.80			4 cell types <sup>e</sup>	5 cell types <sup>e</sup>		1	2.2kb 3' of <i>CLPTM1L</i>		5
rs421629	5	1320136	0.90				1 cell type		4	<i>CLPTM1L</i>	intronic	5
rs380286	5	1320247	0.90				3 cell types			<i>CLPTM1L</i>	intronic	5
rs421284	5	1325590	0.90				4 cell types		6	<i>CLPTM1L</i>	intronic	5
rs466502	5	1325767	0.92				5 cell types		17	<i>CLPTM1L</i>	intronic	5

SNP <sup>a</sup>	Chr	Pos (hg19)	LD <sup>b</sup>	GERP <sup>c</sup>	Histone marks		DNAse HS	Bound proteins	Altered motifs	RefSeq genes	dbSNP annot	RegulomeDB score <sup>d</sup>
					Promoter	Enhancer						
rs465498	5	1325803	0.92				2 cell types		2	CLPTM1L	intronic	5
rs11133729	5	1327101	0.92			6 cell types	10 cell types		4	CLPTM1L	intronic	5
rs452932	5	1330253	0.95			3 cell types			3	CLPTM1L	intronic	6
rs452384	5	1330840	0.93			3 cell types	14 cell types <sup>e</sup>		4	CLPTM1L	intronic	5
rs370348	5	1331219	0.95			3 cell types	5 cell types		5	CLPTM1L	intronic	5
rs2447853	5	1333077	0.88				14 cell types <sup>e</sup>		1	CLPTM1L	intronic	5
rs457130	5	1336178	0.85				1 cell type	POL2	14	CLPTM1L	intronic	2b
rs467095	5	1336221	0.94				2 cell types	POL2	3	CLPTM1L	intronic	4
rs455433	5	1336243	0.95				4 cell types <sup>e</sup>	POL2	2	CLPTM1L	intronic	3a
rs460073	5	1336459	0.95				4 cell types <sup>e</sup>		3	CLPTM1L	intronic	2b
rs462608	5	1336626	0.82				1 cell type		2	CLPTM1L	intronic	5
rs456366	5	1337070	0.95				1 cell type		16	CLPTM1L	intronic	4
rs459961	5	1337106	0.95				1 cell type		6	CLPTM1L	intronic	4
rs31484	5	1337906	0.96				1 cell type		1	CLPTM1L	intronic	6
rs111986123	5	1342157	0.83			1 cell type <sup>e</sup>	9 cell types <sup>e</sup>	MEF2A,POL2,TCF12	7	CLPTM1L	intronic	3a
rs31489	5	1342714	0.88			1 cell type <sup>e</sup>			2	CLPTM1L	intronic	No data
<b>rs31490</b>	<b>5</b>	<b>1344458</b>	<b>1</b>		<b>8 cell types<sup>e</sup></b>	<b>1 cell type</b>	<b>6 cell types<sup>e</sup></b>	<b>MAX</b>		<b>CLPTM1L</b>	<b>intronic</b>	<b>4</b>
rs9479767	6	154476528	1			1 cell type <sup>e</sup>			1	IPCEF1	3'-UTR	6
<b>rs2236256</b>	<b>6</b>	<b>154478440</b>	<b>1</b>			<b>3 cell types<sup>e</sup></b>	<b>10 cell types<sup>e</sup></b>	<b>NFKB</b>	<b>2</b>	<b>IPCEF1</b>	<b>3'-UTR</b>	<b>4</b>
rs2236258	6	154478704	0.97			3 cell types <sup>e</sup>	1 cell type		3	IPCEF1	3'-UTR	5
rs2236259	6	154478893	0.98			1 cell type <sup>e</sup>			7	IPCEF1	3'-UTR	6

SNP <sup>a</sup>	Chr	Pos (hg19)	LD <sup>b</sup>	GERP <sup>c</sup>	Histone marks		DNase HS	Bound proteins	Altered motifs	RefSeq genes	dbSNP annot	RegulomeDB score <sup>d</sup>
					Promoter	Enhancer						
rs6902403	6	154479003	0.98			1 cell type <sup>e</sup>			3	<i>IPCEF1</i>	3'-UTR	No data
rs2239532	7	124405676	0.88	4.4	8 cell types <sup>e</sup>		8 cell types	RAD21,CTCF,POL2	8	<i>GPR37</i>	5'-UTR	4
rs10281582	7	124423688	0.88	2.1		1 cell type		GATA2	3	18kb 5' of <i>GPR37</i>	intronic	5
rs10230610	7	124425398	0.94				97 cell types <sup>e</sup>	CTCF,USF1,RAD21, SMC3 + 3		20kb 5' of <i>GPR37</i>	intronic	4
rs113839563	7	124427335	0.94		1 cell type				8	22kb 5' of <i>GPR37</i>	intronic	No data
rs1014163	7	124443604	0.96					CTCF	3	19kb 3' of <i>POT1</i>		5
rs10228201	7	124449895	0.95			1 cell type			3	13kb 3' of <i>POT1</i>		6
rs1541416	7	124450345	0.91			1 cell type				12kb 3' of <i>POT1</i>		No data
rs720613	7	124452670	0.91			1 cell type				9.8kb 3' of <i>POT1</i>		No data
rs10241173	7	124453258	0.91				1 cell type		8	9.2kb 3' of <i>POT1</i>		6
<b>rs17246404</b>	<b>7</b>	<b>124462661</b>	<b>1</b>						<b>4</b>	<b><i>POT1</i></b>	<b>3'-UTR</b>	<b>6</b>
rs6952721	7	124480075	0.91			3 cell types			2	<i>POT1</i>	intronic	5
rs7801661	7	124483222	0.97			1 cell type			1	<i>POT1</i>	intronic	No data
rs7795943	7	124483796	0.97			1 cell type			2	<i>POT1</i>	intronic	No data
rs28728448	7	124532938	0.87			1 cell type <sup>e</sup>			4	<i>POT1</i>	intronic	No data
rs59294613	7	124554267	0.82			1 cell type <sup>e</sup>			1	<i>POT1</i>	intronic	No data
rs67939974	7	124554364	0.86			1 cell type <sup>e</sup>			3	<i>POT1</i>	intronic	6
rs10233596	7	124556659	0.86			1 cell type <sup>e</sup>			4	<i>POT1</i>	intronic	6
rs56356267	7	124557549	0.86			2 cell types <sup>e</sup>			1	<i>POT1</i>	intronic	No data
rs28844358	7	124562695	0.86			2 cell types <sup>e</sup>			2	<i>POT1</i>	intronic	No data
<b>rs2511714</b>	<b>8</b>	<b>103578874</b>	<b>1</b>			<b>3 cell types<sup>e</sup></b>	<b>4 cell types<sup>e</sup></b>			<b>5.6kb 3' of <i>ODF1</i></b>		<b>5</b>

Alt, alternative; annot, annotation; Chr, chromosome; DNase HS, DNase hypersensitivity; Freq, frequency; GERP, Genomic Evolutionary Rate Profiling; LD, linkage disequilibrium; Pos, position

<sup>a</sup> GWAS hits (i.e. replicated SNPs) are emboldened

<sup>b</sup> LD ( $r^2$ ) is based upon 1000 Genomes EUR data and a threshold of  $r^2 > 0.8$  was imposed to identify correlated SNPs at all four susceptibility loci

<sup>c</sup> GERP scores indicative of evolutionary constraint ( $>2$ ) are listed

<sup>d</sup> RegulomeDB scores: 2b, transcription factor (TF) binding + any motif + DNase Footprint + DNase peak; 3a, TF binding + any motif + DNase peak; 4, TF binding + DNase peak; 5, TF binding or DNase peak; 6, other.

<sup>e</sup> Cell types include lymphoblastoid cell lines

**Supplementary Table 5: Clinico-pathological association testing.** Age, gender and mutational status based on all stages; telomere length analysis based on a subset (N=246) of the Swedish replication cases.

SNP	Locus	Genotype	Gender			Age			IGHV mutation status			Telomere length	
			Male	Female	<i>P</i>	<63	≥63	<i>P</i>	Mutated	Unmutated	<i>P</i>	N	<i>P</i>
rs10936599	3q26.2	CC	1134	638	0.186	813	915	0.847	580	407	0.238	142	0.766
		CT	599	337		432	483		330	212		97	
		TT	65	55		53	63		36	19		7	
rs6858698	4q26	GG	1100	635	0.931	819	883	0.010	589	407	0.262		
		CG	438	235		283	370		257	143			
		CC	31	25		18	33		16	13			
rs10069690 <sup>a</sup>	5p15.33	CC	527	262	0.231	375	409	0.218	135	138	0.596		
		CT	403	226		319	307		105	92			
		TT	68	43		57	54		17	18			
rs31490	5p15.33	GG	463	290	0.283	343	393	0.906	237	170	0.496	68	
		AG	945	526		681	746		523	310		124	
		AA	393	220		277	324		190	155		54	
rs2236256	6q25.2	AA	423	269	0.189	298	375	0.067	237	147	0.102		
		AC	908	510		652	734		470	299			
		CC	439	239		330	338		215	172			
rs17246404	7q31.33	CC	993	609	0.045	761	797	0.183	535	367	0.899	143	
		CT	670	360		438	570		352	218		84	
		TT	118	56		89	84		56	45		19	
rs2511714 <sup>a</sup>	8q22.3	GG	170	94	0.655	126	138	0.758	58	51	0.168		
		GT	451	239		357	329		145	120			
		TT	334	196		262	265		89	103			

<sup>a</sup> Data from UK-CLL-1 and UK-CLL-2 only

**Supplementary Table 6: Association between CLL risk SNPs and transcript levels of *cis*-genes in the MuTHER lymphoblastoid cell line dataset.**

Locus	GWAS SNP	eQTL query SNP <sup>a</sup>	LD ( $r^2$ ) <sup>b</sup>	<i>Cis</i> - gene	Probe	Direction of effect <sup>c</sup>	eQTL <i>P</i> -value
3q26.2	rs10936599	rs10936599	1.00	<i>TERC</i>	ILMN_1766573	↑	0.04
4q26	rs6858698	rs17532826	0.41	<i>CAMK2D</i>	ILMN_1658818	↓	0.004
					ILMN_1683415	↓	0.89
					ILMN_1757060	↓	0.82
5p15.33	rs10069690	no proxy available	-	<i>TERT</i>	-	-	-
5p15.33	rs31490	rs401681	1.00	<i>CLPTM1L</i>	ILMN_1752802	↑	0.44
6q25.2	rs2236256	rs2236256	1.00	<i>IPCEF1</i>	ILMN_1796497	↓	0.13
7q31.33	rs17246404	rs727505	1.00	<i>POT1</i>	ILMN_1692041	↓	0.97
					ILMN_2287707	↓	0.61
					ILMN_2327203	↓	0.08
8q22.3	rs2511714	no proxy available	-	<i>ODF1</i>	-	-	-

LD, linkage disequilibrium

<sup>a</sup> eQTL data is for GWAS SNP (where available) or proxy SNP ( $r^2 > 0.4$  with genotyped SNP)

<sup>b</sup>  $r^2$  based on 1000Genomes EUR data

<sup>c</sup> Direction of change in transcript levels in carriers of risk allele

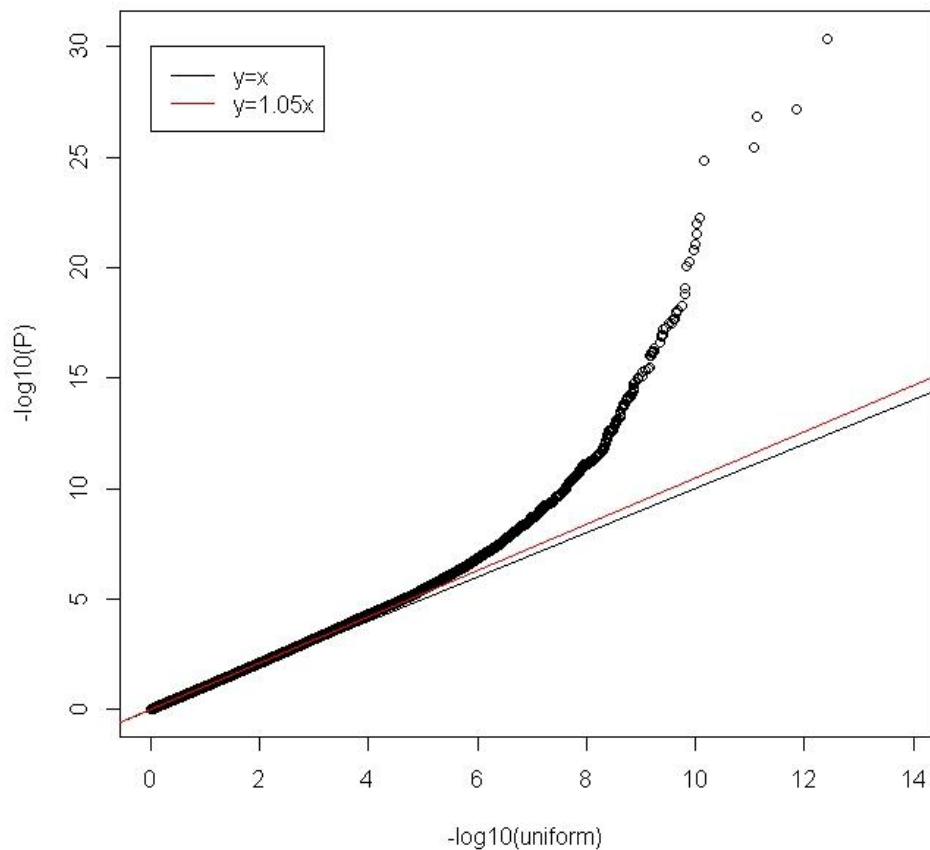
**Supplementary Table 7: Pairwise analysis of rs10936599, rs6858698, rs31490, rs2236256, rs17246404 and previously established loci.** For each row-column combination, numbers show the *P* value and number of samples (N) the result is based on, for inclusion of an interaction term between the two SNPs. The interactions are based on data from UK-CLL-2 and UK-CLL-1.

SNP	rs10936599	rs6858698	rs31490	rs2236256	rs17246404
rs6858698	0.67 (N=5755)	-	-	-	-
rs31490	0.60 (N=6924)	0.93 (N=5745)	-	-	-
rs2236256	0.37 (N=6938)	0.21 (N=5755)	0.76 (N=6924)	-	-
rs17246404	0.36 (N=6938)	0.10 (N=5755)	0.31 (N=6924)	0.09 (N=6938)	-
rs872071	0.88 (N=6938)	0.75 (N=5755)	0.36 (N=6924)	0.99 (N=6938)	0.10 (N=6938)
rs11083846	0.87 (N=6830)	0.99 (N=5667)	0.37 (N=6816)	0.84 (N=6830)	0.34 (N=6830)
rs7176508	0.66 (N=6938)	<b>0.003</b> (N=5755)	0.77 (N=6924)	0.70 (N=6938)	0.73 (N=6938)
rs13397985	0.23 (N=6938)	0.19 (N=5755)	0.24 (N=6924)	0.34 (N=6938)	0.22 (N=6938)
rs735665	0.19 (N=6938)	0.54 (N=5755)	0.94 (N=6924)	0.97 (N=6938)	0.20 (N=6938)
rs17483466	0.33 (N=6938)	0.59 (N=5755)	0.54 (N=6924)	0.63 (N=6938)	0.37 (N=6938)
rs2456449	0.46 (N=6875)	0.80 (N=5703)	0.99 (N=6861)	0.34 (N=6875)	0.94 (N=6875)
rs305061	0.14 (N=6938)	0.13 (N=5755)	0.26 (N=6924)	0.44 (N=6938)	0.66 (N=6938)
rs7169431	0.56 (N=6938)	0.25 (N=5755)	0.41 (N=6924)	<b>0.012</b> (N=6938)	<b>0.014</b> (N=6938)
rs757978	0.47 (N=6938)	0.26 (N=5755)	0.71 (N=6924)	0.62 (N=6938)	0.96 (N=6938)
rs783540	0.43 (N=6938)	0.26 (N=5755)	0.54 (N=6924)	0.27 (N=6938)	0.52 (N=6938)
rs210142	0.47 (N=6905)	0.16 (N=5731)	0.75 (N=6891)	0.72 (N=6905)	0.73 (N=6905)
rs4406737	0.19 (N=6935)	0.56 (N=5753)	0.73 (N=6921)	0.97 (N=6935)	0.15 (N=6935)
rs4987855	0.35 (N=6884)	0.79 (N=5708)	0.42 (N=6870)	0.54 (N=6884)	0.37 (N=6884)
rs4987852	0.08 (N=6938)	0.90 (N=5755)	0.54 (N=6924)	0.24 (N=6938)	0.14 (N=6938)
rs7944004	0.33 (N=6076)	0.17 (N=5060)	0.34 (N=6067)	0.63 (N=6076)	0.38 (N=6076)
rs898518	0.55 (N=6938)	0.92 (N=5755)	0.54 (N=6924)	0.64 (N=6938)	0.39 (N=6938)
rs3769825	0.28 (N=6938)	0.61 (N=5755)	0.60 (N=6924)	0.62 (N=6938)	0.88 (N=6938)
rs1679013	0.58 (N=4227)	0.46 (N=3496)	0.82 (N=4218)	0.78 (N=4227)	0.38 (N=4227)
rs4368253	0.29 (N=6639)	0.20 (N=5505)	0.86 (N=6626)	0.48 (N=6639)	0.25 (N=6639)
rs8024033	0.84 (N=5709)	<b>0.018</b> (N=4746)	0.48 (N=5699)	0.58 (N=5709)	<b>0.018</b> (N=5709)
rs3770745	0.48 (N=6424)	0.88 (N=5326)	<b>0.032</b> (N=6410)	0.85 (N=6424)	0.17 (N=6424)
rs13401811	0.06 (N=6913)	0.35 (N=5735)	0.21 (N=6899)	0.95 (N=6913)	0.34 (N=6913)
rs10069690	0.84 (N=5901)	0.62 (N=4911)	0.72 (N=5899)	0.26 (N=5901)	0.21 (N=5901)

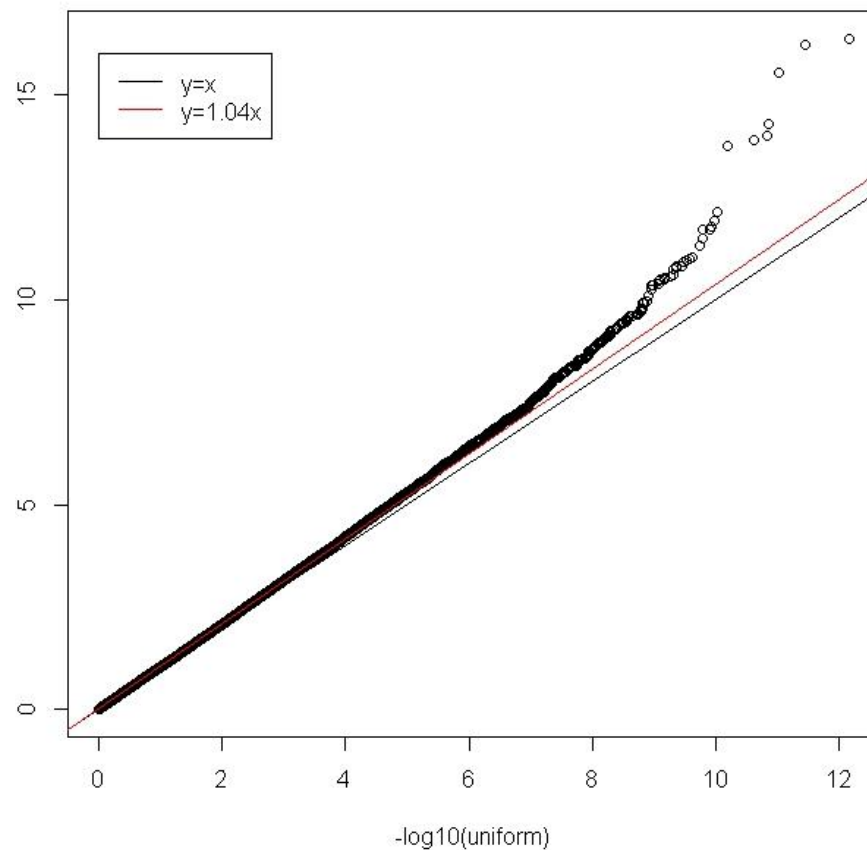
<b>rs2511714</b>	0.74 (N=6060)	0.36 (N=5022)	0.37 (N=6048)	0.91 (N=6060)	0.74 (N=6060)
<b>rs674313</b>	0.32 (N=6746)	0.21 (N=5607)	<b>0.011</b> (N=6734)	<b>0.003</b> (N=6746)	0.69 (N=6746)



(a) QQ plot for UK-CLL-2, Armitage trend test

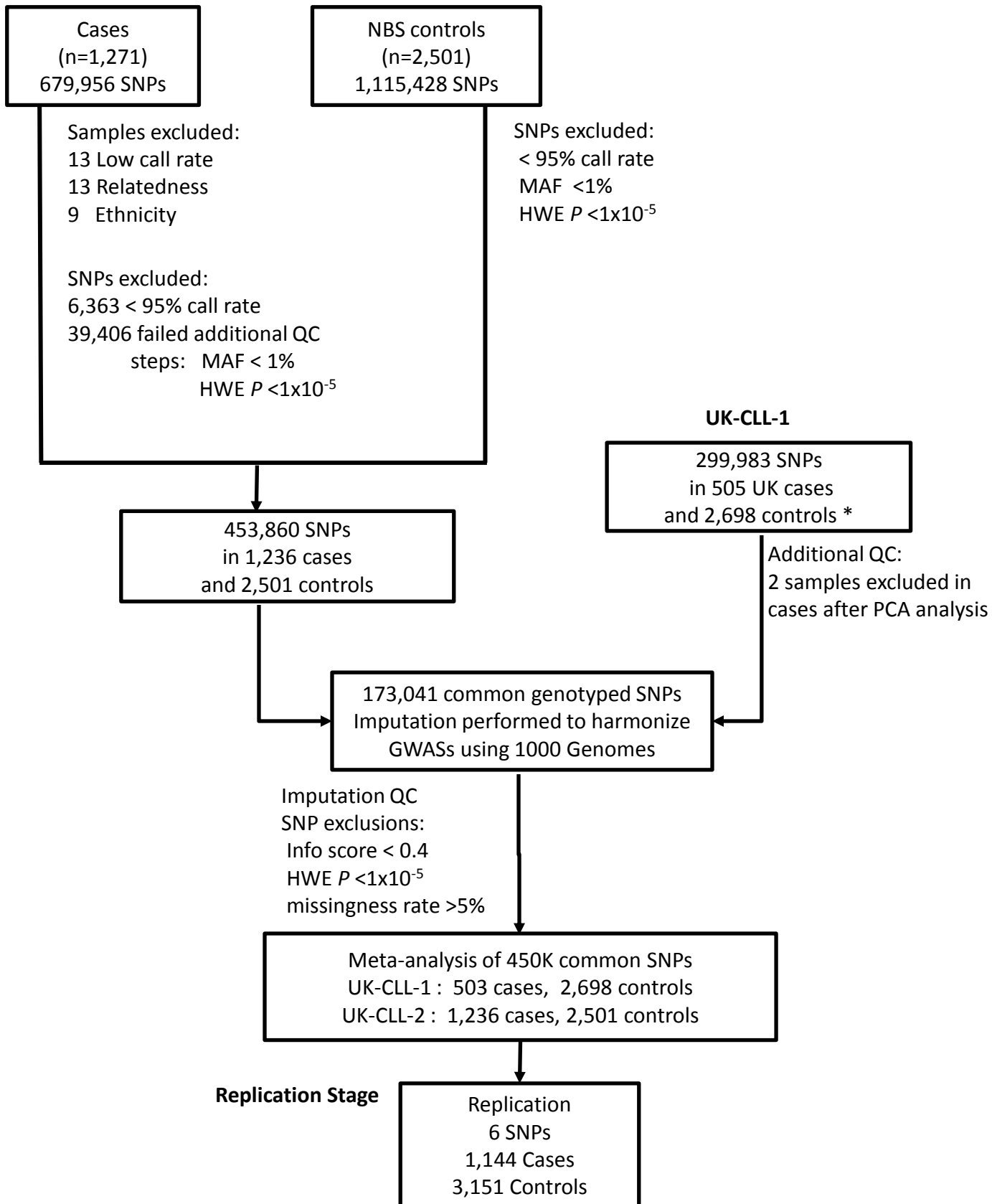


(b) QQ plot for UK-CLL-1, Armitage trend test



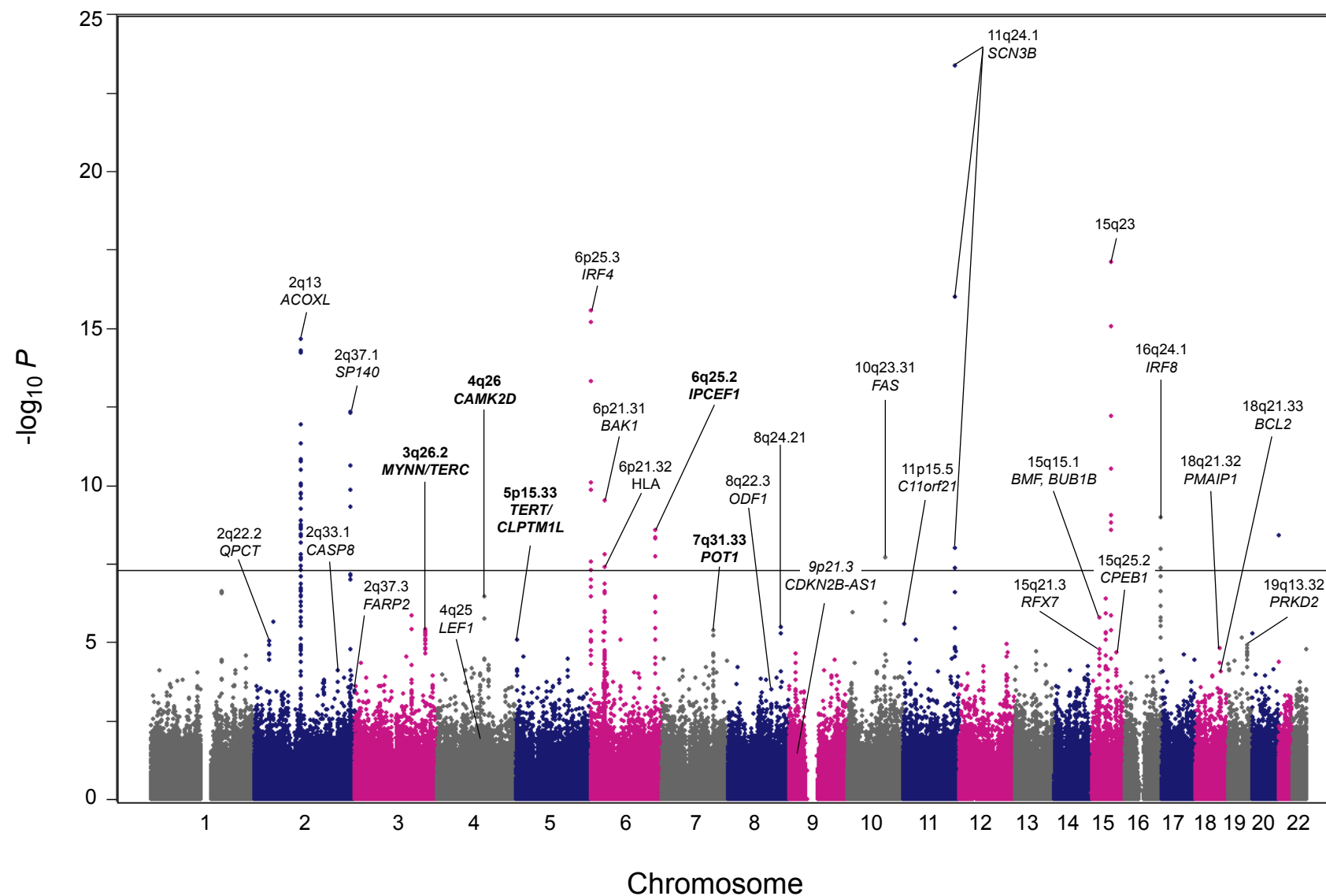
**Supplementary Figure 1:** Q-Q plots of observed  $P$ -values versus expected  $P$ -values from association results in (a) UK-CLL-2 ( $\lambda=1.05$ ) and (b) UK-CLL-1 ( $\lambda=1.04$ ).

## UK-CLL-2



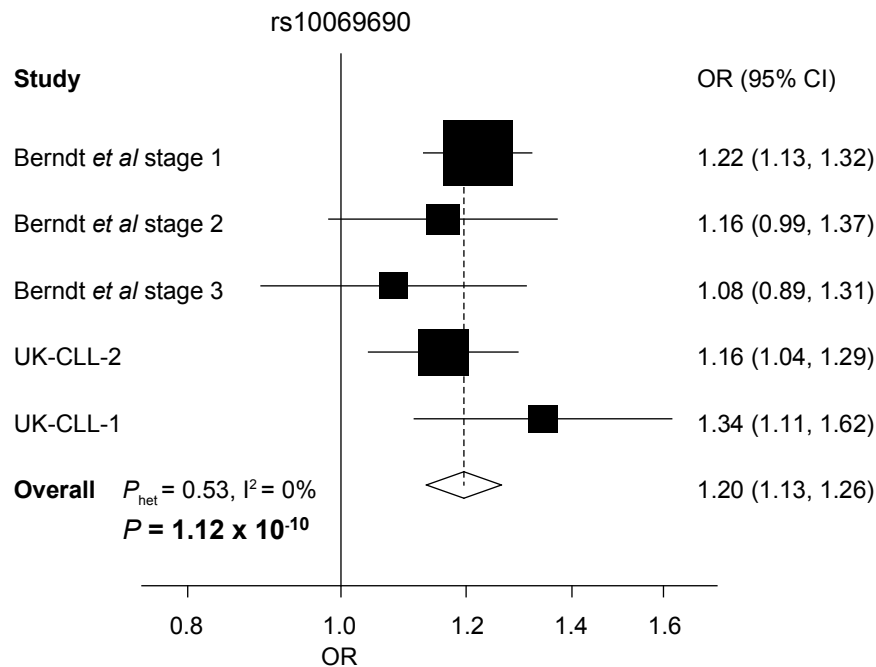
**Supplementary Figure 2: GWAS data quality control.** Details are provided of samples, SNPs and quality control (QC) used in each GWAS.

\* In original study after QC 505 UK cases and 1,438 controls were used

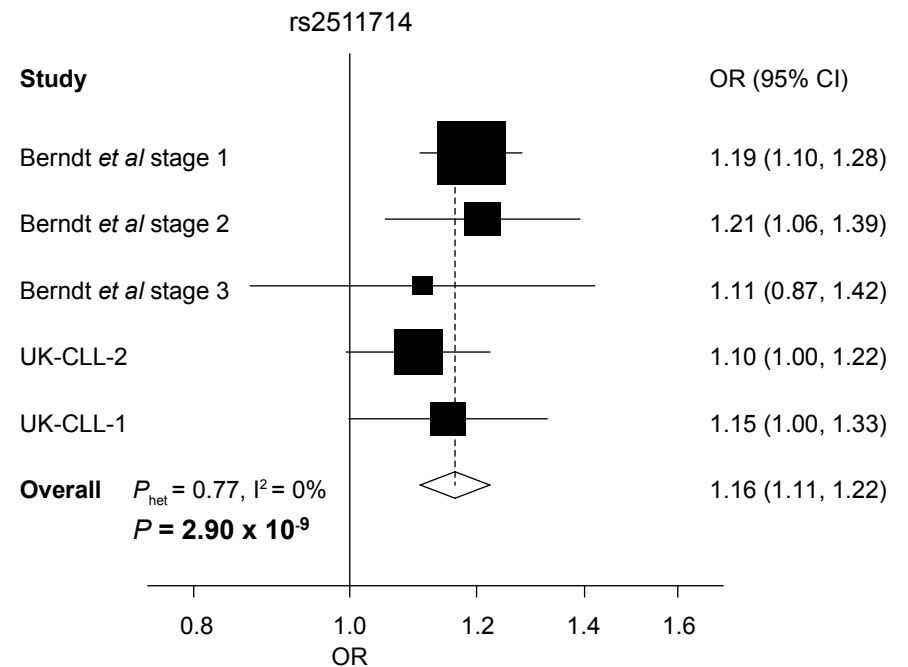


**Supplementary Figure 3: Genome-wide  $P$ -values ( $-\log_{10}P$ , y axis) plotted against their respective chromosomal positions (x axis).** Shown are the genome-wide  $P$ -values (two-sided) obtained in the combined analysis of UK-CLL-1 and UK-CLL-2 (fixed effects meta-analysis). Only autosomal SNPs, directly typed in UK-CLL-2 are illustrated. The horizontal line represents the genome-wide significance threshold level ( $P=5.0 \times 10^{-8}$ ). CLL risk loci reported in this and previous studies are labelled.

a

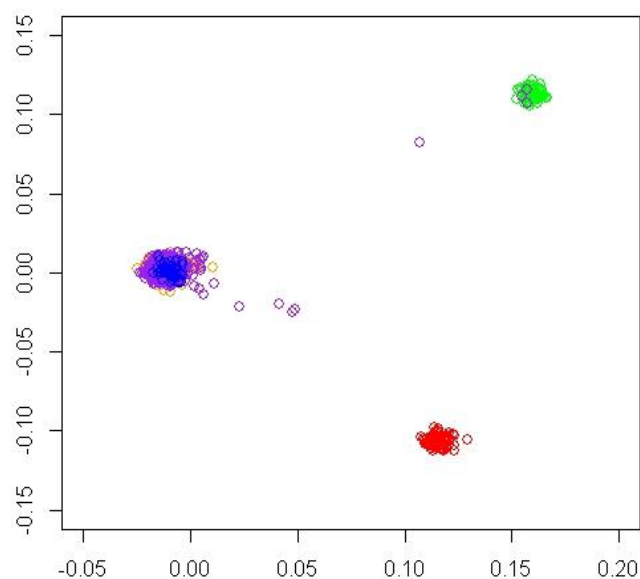


b

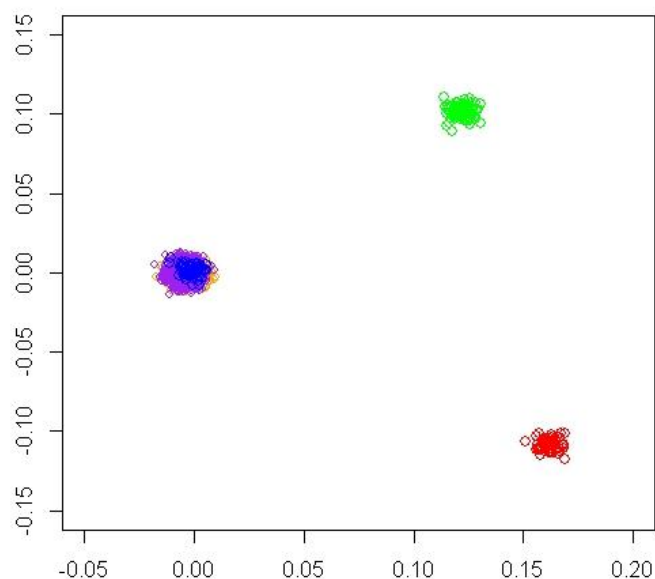


**Supplementary Figure 4: Plot of the odds ratios of chronic lymphocytic leukemia associated with (a) rs10069690 and (b) rs2511714 in this and previously published studies.** Studies were weighted according to the inverse of the variance of the log of the odds ratio (OR) calculated by unconditional logistic regression. *Horizontal lines*: 95% confidence intervals (95% CI). *Box*: OR point estimate; its area is proportional to the weight of the study. *Diamond (and broken line)*: summary OR computed under a fixed effects model, with 95% CI given by its width. *Unbroken vertical line*: null value (OR=1.0).

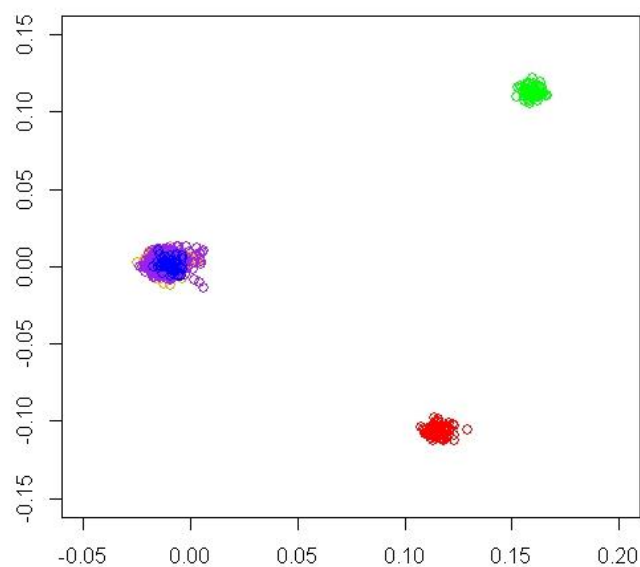
(a)



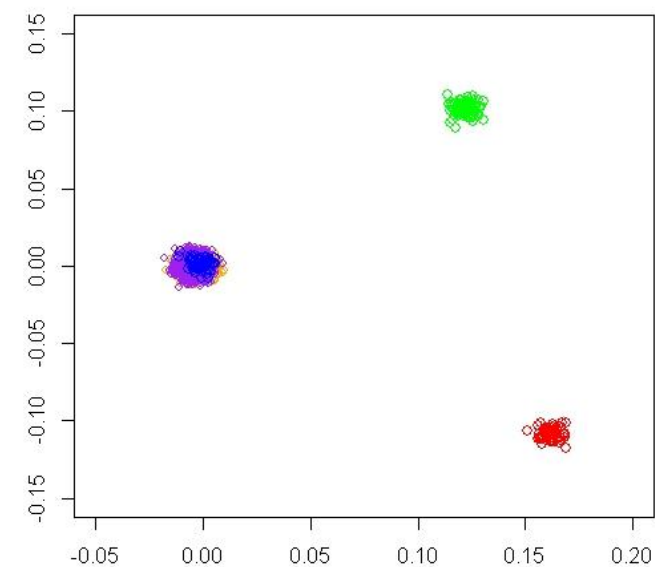
(b)



(c)



(d)



**Supplementary Figure 5: Identification of individuals in the GWA scan of non-European ancestry.** The first two principal components of the analysis were plotted. HapMap CEU individuals are plotted in blue; CHB+JPT are plotted in green; YRI individuals are plotted in red; UK-CLL-2 cases are plotted in purple and UK-CLL-1 cases are plotted in orange before (a) and after (c) removal; controls for UK-CLL-2 are plotted in purple and controls for UK-CLL-1 are plotted in orange before (b) and after (d) removal.