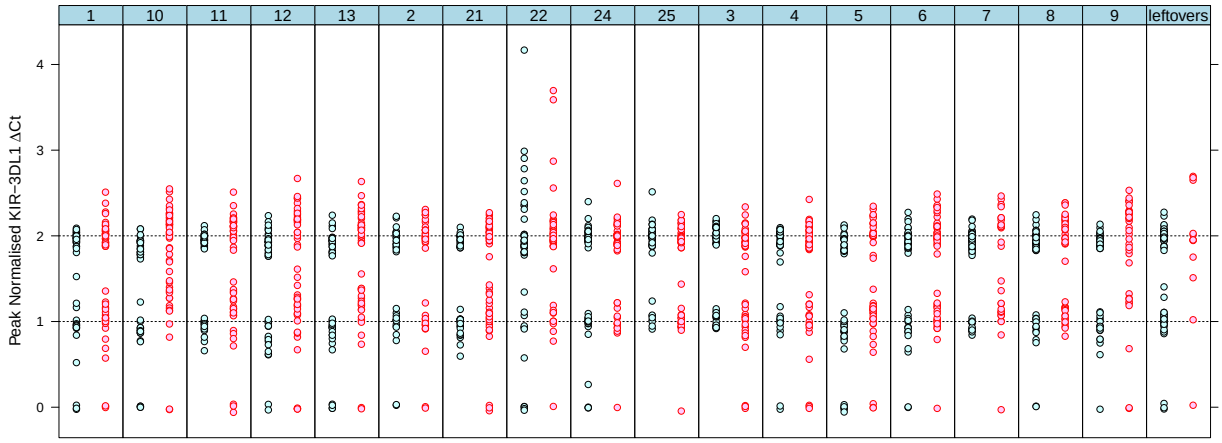
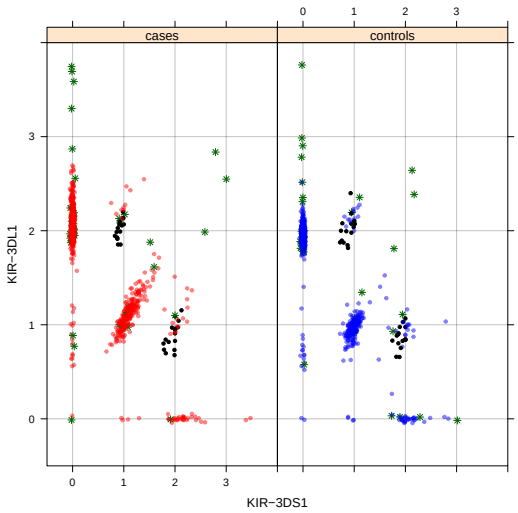
Supplementary

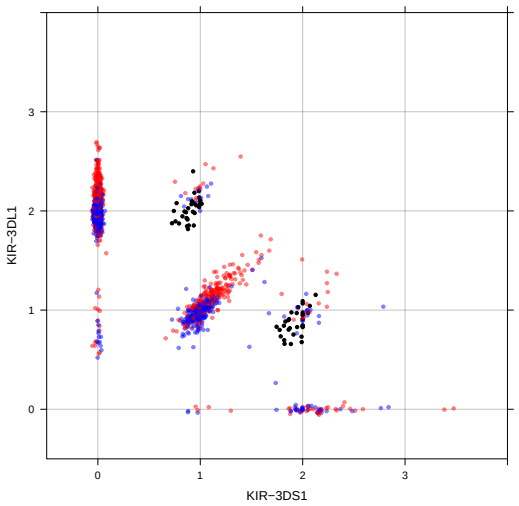
Figures



Supplementary Figure 1. Post-QC normalised *KIR3DL1* ΔCt values for cases (red) and controls (blue) per plate. Plate 22 stands out as the noisiest for *KIR3DL1* which suggests it should be excluded from the analysis.

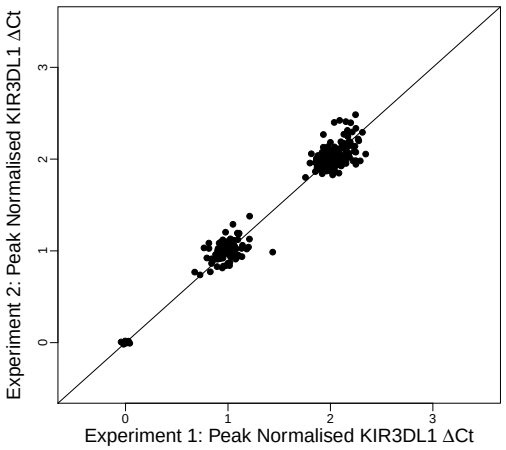


(a) Pre QC: 816 cases (red) and 813 controls (blue).

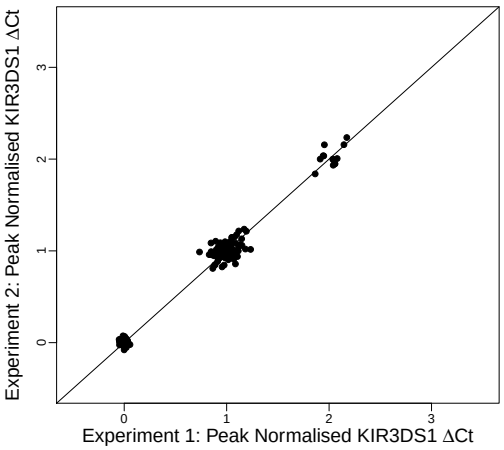


(b) Post QC: 747 cases (red) and 727 controls (blue).

Supplementary Figure 2. (a) Cases are in red, controls are in blue and the samples with known *KIR3DL1/S1* copy number are in black. Samples from plate 22 are represented by the green asterisks. There is a larger spread in cases than in controls for 3DL1. QC involved dropping plate 22 which is very noisy and samples for which we do not have four replicates. (b) After QC we are left with a total of 1,474 unique samples (747 cases and 727 controls).

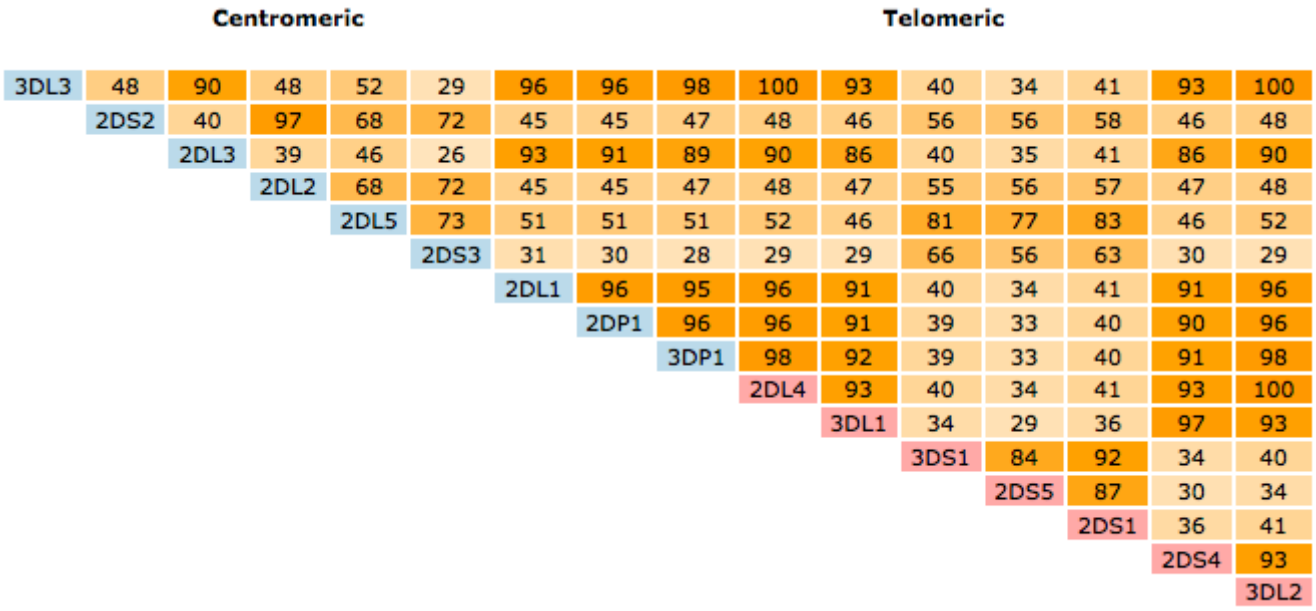


(a) Repeatability of KIR3DL1 ΔCt post normalisation and QC ($r^2 = 0.961$).



(b) Repeatability of KIR3DS1 ΔCt post normalisation and QC ($r^2 = 0.99$).

Supplementary Figure 3. In order to assess the reliability of the qPCR assay 310 samples were re-analysed. We found very high reproducibility of the ΔCt values ($r^2 > 0.96$) confirming the reliability of our qPCR assay.



Supplementary Figure 4. Linkage values 0-100% between KIR genes organised by telomeric (blue) and centromeric (red) content. Gene association values are estimated using the genotype list which is available on the Allele Frequencies Net website (Gonzalez-Galarza et al, 2011).

Tables

Gene	Oligos	Sequence (5'-3')
3DS1	Forward Primer	CATCGGTTCCATGATGCG
	Reverse Primer	GGGAGCTGACAACTGATAGG
	Probe	AACAGAACCGTAGCATCTGTAGGTCCCT
3DL1	Forward Primer	CACAGTTGGATCACTGCGT
	Reverse Primer	CCGTGTACAAGATGGTATCTGTA
	Probe	CCCTTCTCAGAGGCCCAAGACAC
STAT6	Forward Primer	CCAGATGCCTACCATGGTG
	Reverse Primer	CCATCTGCACAGACCACTCC
	Probe	CTGATTCCTCCATGAGCATGCAGCTT

Supplementary Table 1. The qPCR primers used. These were originally developed by Jiang et al (2012).

Epitope		Residues (77-83)	HLA-B	HLA-A
HLA-Bw4	80I	NLR**I**ALR	B*1516 B*1517 B*1524 B*2702 B*3801 B*4901 B*5101 B*5108 B*5201 B*5301 B*5302 B*5701 B*5702 B*5801	A*2301 A*2402 A*2403 A*2407 A*2501 A*3201
HLA-Bw4	80T	DLR**T**LLR	B*1302 B*2701 B*2704 B*2705 B*3701 B*3802 B*4402 B*4403 B*4404 B*4405 B*4414 B*4417 B*4429 B*4435 B*4701	
HLA-Bw6		SLR**N**LRG	B*702 B*703 B*705 B*706 B*708 B*710 B*716 B*726 B*801 B*1401 B*1402 B*1501 B*1503 B*1504 B*1505 B*1507 B*1508 B*1509 B*1510 B*1514 B*1515 B*1518 B*1539 B*1801 B*3501 B*3502 B*3503 B*3508 B*3901 B*3906 B*3928 B*4001 B*4002 B*4006 B*4011 B*4023 B*4101 B*4102 B*4202 B*4501 B*4601 B*4801 B*5001 B*5002 B*5501 B*5601	

Supplementary Table 2. The Martin et al (2002) HLA-Bw4/6 epitope defines the classification of HLA-B and HLA-A alleles based on the amino acid at position 80. HLA-Bw4 groups both HLA-Bw4-80I and HLA-Bw4-80T epitopes.

HLA Epitope	Cases	Controls	Total
N/A	3822 (11)	2681 (70)	6503 (81)
HLA-Bw6	1175 (308)	753 (199)	1928 (507)
HLA-Bw4-80T	651 (162)	754 (174)	1405 (336)
HLA-Bw4-80I	1096 (266)	1174 (284)	2270 (550)
HLA total	2922 (736)	2681 (657)	5603 (1393)

Supplementary Table 3. HLA epitope classification of subjects in study. In parentheses, number of subjects analysed with qPCR post QC. No HLA typing was done for the N/A category. The HLA epitopes are defined in Table 2. An individual is assigned to an HLA epitope group if he is a carrier of at least one allele of that group. So that each individual only belongs to a single HLA epitope group, the assignment priority is first HLA-Bw4-80I, then HLA-Bw4-80T and finally HLA-Bw6 allele if no HLA-Bw4 alleles were found.

	Name	Position	Score	SNP	Ill Strand	Cust Strand	NormID	QC
1	rs597598	60007252	0.89	[A/G]	TOP	TOP	1	ok
2	rs598452	60007428	0.86	[A/G]	TOP	BOT	1	ok
3	t1d-19-60007809-C-G	60007809	0.79	[G/C]	BOT	BOT	100	ok
4	rs55761930	60008141	0.66	[T/C]	BOT	TOP	1	ok
5	rs10500318	60012591	0.72	[A/G]	TOP	TOP	1	ok
6	rs592645	60012739	0.68	[A/T]	TOP	BOT	100	ok
7	rs604077	60013208	0.89	[A/G]	TOP	BOT	1	ok
8	rs604999	60013409	0.89	[A/G]	TOP	TOP	1	ok
9	t1d-19-60014013-A-C	60014013	0.82	[T/G]	BOT	TOP	1	lowcallrate
10	rs3865507	60014188	0.80	[T/G]	BOT	TOP	0	ok
11	rs3865510	60016051	0.87	[A/C]	TOP	TOP	1	ok
12	rs648689	60016286	0.86	[A/G]	TOP	BOT	1	ok
13	rs649216	60016447	0.90	[T/C]	BOT	BOT	1	ok
14	rs581623	60018551	0.86	[A/G]	TOP	BOT	0	ok
15	rs4806568	60022568	0.41	[A/G]	TOP	TOP	1	lowcallrate
16	rs674268	60024002	0.64	[T/C]	BOT	TOP	1	lowcallrate
17	rs12461010	60024413	0.79	[A/G]	TOP	BOT	0	ok
18	rs2295805	60028513	0.83	[T/C]	BOT	TOP	1	lowcallrate
19	rs12976350	60030391	0.61	[T/C]	BOT	BOT	1	lowcallrate
20	t1d-19-60034052-C-T	60034052	0.63	[A/G]	TOP	BOT	1	hwe
21	rs4806585	60038236	0.58	[T/G]	BOT	TOP	0	hwe
22	rs62122181	60039178	0.22	[T/C]	BOT	TOP	1	lowcallrate
23	rs10422740	60052298	0.81	[T/C]	BOT	BOT	0	monomorph
24	rs640345	60054671	0.67	[A/G]	TOP	BOT	0	ok
25	t1d-19-60054973-T-C	60054973	0.69	[A/G]	TOP	BOT	1	ok
26	t1d-19-60056605-A-T	60056605	0.70	[A/T]	TOP	BOT	200	ok
27	t1d-19-60056721-C-T	60056721	0.62	[A/G]	TOP	BOT	1	ok
28	rs10407958	60063974	0.89	[T/A]	BOT	BOT	201	ok
29	rs1654644	60065174	0.82	[T/G]	BOT	BOT	1	ok
30	rs3826878	60069023	0.91	[A/G]	TOP	TOP	0	ok

Supplementary Table 4. The ImmunoChip SNPs which fall in the *KIR3DL1* region according to build36/hg18. *KIR3DS1* is missing from build36/h18.

k	LOO Error %
1	2.68
2	2.89
3	**2.10**
4	2.18
5	2.29
6	2.36
7	2.39
8	2.44
9	2.56
10	2.71
11	2.89
12	3.11
13	3.47
14	3.62
15	3.64
16	3.78
17	3.85
18	3.81
19	3.96
20	3.89

Supplementary Table 5. Average leave-one-out (LOO) error rate obtained from running k nearest neighbour for k from 1 to 20 over 10 multiply imputed qPCR datasets. The smallest average error rate is 2.10% obtained for k=3.

SNP \ qPCR	0-1	0-2	1-0	1-1	1-2	2-0	2-1	3-0
0-1	23.00	0.40	0.00	0.00	0.00	0.00	0.00	0.00
0-2	1.00	886.30	0.00	2.00	1.00	0.00	1.00	0.00
1-0	0.00	0.00	5.00	0.00	0.00	0.00	0.00	0.00
1-1	0.00	2.30	2.00	434.00	2.30	0.00	2.60	0.00
1-2	0.00	1.00	0.00	0.00	21.30	0.00	0.10	0.00
2-0	0.00	0.00	0.00	0.00	0.00	51.50	0.00	2.30
2-1	0.00	0.00	0.00	0.00	2.40	1.00	27.30	0.00
3-0	0.00	0.00	0.00	0.00	0.00	0.80	0.00	3.40

Supplementary Table 6. Using the posterior probabilities obtained from the qPCR genotype calls, 10 datasets were multiply imputed and each was used for training a k nearest neighbour classifier (k=3) to predict genotype from SNP data. This table represents the average over the 10 misclassification tables. Each cell along the diagonal shows the average number of matching predictions and the cells off the diagonal show the contains the average number of mismatches. The average misclassification rate is 1.5%.

	qPCR			SNP		
KIR3DS1-KIR3DL1	count (percentage)			count (percentage)		
Copy Number	cases	controls	total	cases	controls	total
0-2	444 (59.44)	446 (61.35)	890 (60.38)	4091 (60.66)	3220 (60.05)	7311 (60.39)
1-1	229 (30.66)	207 (28.47)	436 (29.58)	2056 (30.49)	1628 (30.36)	3684 (30.43)
2-0	26 (3.48)	28 (3.85)	54 (3.66)	231 (3.43)	223 (4.16)	454 (3.75)
2-1	15 (2.01)	16 (2.2)	31 (2.1)	116 (1.72)	104 (1.94)	220 (1.82)
1-2	13 (1.74)	14 (1.93)	27 (1.83)	116 (1.72)	79 (1.47)	195 (1.61)
0-1	13 (1.74)	11 (1.51)	24 (1.63)	101 (1.5)	73 (1.36)	174 (1.44)
1-0	4 (0.54)	3 (0.41)	7 (0.47)	24 (0.36)	20 (0.37)	44 (0.36)
3-0	3 (0.4)	2 (0.28)	5 (0.34)	9 (0.13)	15 (0.28)	24 (0.2)
.-2	457 (61.18)	460 (63.27)	917 (62.21)	4192 (62.16)	3293 (61.41)	7485 (61.83)
.-1	257 (34.4)	234 (32.19)	491 (33.31)	2288 (33.93)	1811 (33.77)	4099 (33.86)
.-0	33 (4.42)	33 (4.54)	66 (4.48)	264 (3.91)	258 (4.81)	522 (4.31)
0-.	457 (61.18)	457 (62.86)	914 (62.01)	4207 (62.38)	3299 (61.53)	7506 (62)
1-.	246 (32.93)	224 (30.81)	470 (31.89)	2181 (32.34)	1721 (32.1)	3902 (32.23)
2-.	41 (5.49)	44 (6.05)	85 (5.77)	347 (5.15)	327 (6.1)	674 (5.57)
3-.	3 (0.4)	2 (0.28)	5 (0.34)	9 (0.13)	15 (0.28)	24 (0.2)
Total	747 (100)	727 (100)	1,474 (100)	6,744 (100)	5,362 (100)	12,106 (100)

Supplementary Table 7. *KIR3DS1-KIR3DL1* genotype counts and percentages in parentheses obtained from the qPCR and SNP data set.

KIR3DS1-KIR3DL1		dataset	
Copy Number	Jiang et al (2012)	SNP	qPCR
0-2	57.95	60.39	60.38
1-1	30.65	30.43	29.58
2-0	4.07	3.75	3.66
1-2	2.26	1.44	1.83
0-1	2.18	1.61	1.63
2-1	1.81	1.82	2.10
1-0	0.58	0.36	0.47
3-0	0.32	0.20	0.34
0-3	0.10		
3-1	0.02		
2-2	0.02		
0-0	0.02		
4-0	0.01		
1-3	0.00		
0-4	0.00		

Supplementary Table 8. Comparison of *KIR3DS1-KIR3LD1* genotype frequencies calculated from Jiang et al (2012) KIR haplotype frequencies (assuming Hardy-Weinberg) compared to those obtained from our qPCR and SNP dataset.

References

Gonzalez-Galarza FF, Christmas S, Middleton D, and Jones AR. 2011. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic acids research* **39**: D913–9.

Jiang W, Johnson C, Jayaraman J, Simecek N, Noble J, Moffatt MF, Cookson WO, Trowsdale J, and Traherne JA. 2012. Copy number variation leads to considerable diversity for B but not A haplotypes of the human KIR genes encoding NK cell receptors. *Genome Research* **22**: 1845–1854.

Martin MP, Gao X, Lee JH, Nelson GW, Detels R, Goedert JJ, Buchbinder S, Hoots K, Vlahov D, Trowsdale J, et al. 2002. Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. *Nature Genetics* **31**: 429–434.