Copy number and nucleotide variation of the LILR family of myelomonocytic cell activating and inhibitory receptors

María R. López-Álvarez · Des C. Jones · Wei Jiang · James A. Traherne · John Trowsdale

Abstract Leukocyte immunoglobulin-like receptors (LILR) are cell surface molecules that regulate the activities of myelomonocytic cells through the balance of inhibitory and activation signals. LILR genes are located within the leukocyte receptor complex (LRC) on chromosome 19q13.4 adjacent to KIR genes, which are subject to allelic and copy number variation (CNV). LILRB3 (ILT5) and LILRA6 (ILT8) are highly polymorphic receptors with similar extracellular domains. LILRB3 contains inhibitory ITIM motifs and LILRA6 is coupled to an adaptor with activating ITAM motifs. We analysed the sequences of the extracellular immunoglobulin domain-encoding regions of *LILRB3* and *LILRA6* in 20 individuals, and determined the copy number of these receptors, in addition to those of other members of the LILR family. We found 41 polymorphic sites within the extracellular domains of *LILRB3* and *LILRA6*. Twenty-four of these sites were common to both receptors. *LILRA6*, but not *LILRB3*, exhibited CNV. In 20 out of 48 human cell lines from the International Histocompatibility Working Group, *LILRA6* was deleted or duplicated. The only other LILR gene exhibiting genomic aberration was *LILRA3*, in this case due to a partial deletion.

Keywords LILR · ILT · CNV · Haplotype

Introduction

Leukocyte immunoglobulin (Ig)-like receptors (LILR; also termed LIR or ILT) are mainly expressed on the surface of myelomonocytic cells (Brown et al. 2004) and they have been shown to be important mediators of immunological tolerance (Manavalan et al. 2003; Kim-Schulze et al. 2006). They have been proposed to regulate the function of Toll-like receptors (TLR) and to alter the phenotype and profile of cytokine production by antigen presenting cells (Chang et al. 2002, 2009; Anderson and Allen 2009; Brown et al. 2009), thereby controlling both innate and adaptive immune responses. Their functions have been linked to control of bacterial and viral infection (Brown et al. 2009; Lee et al. 2007; Pilsbury et al. 2010). LILRs are genetically associated with autoimmune diseases such as multiple sclerosis (Koch et al. 2005; Ordóñez et al. 2009), Sjögren's syndrome (Kabalak et al. 2009), and rheumatoid arthritis (Kuroki et al. 2005; Huynh et al. 2007).

LILR receptors can be classified into two main groups, LILRA and LILRB. The LILRA group includes genes that have truncated cytoplasmic tails and associate with the γ-chain of FcεRI through a charged arginine residue in the transmembrane domain, delivering an activating signal through an immunoreceptor tyrosine-based activation motif (ITAM) (Nakajima et al. 1999). One exception is LILRA3, a soluble protein that has no known signalling ability (Borges et al. 1997; Colonna et al. 1999). On the other hand, LILRB

M.R. López-Álvarez and D.C. Jones contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00251-013-0742-5) contains supplementary material, which is available to authorized users.

M. R. López-Álvarez · D. C. Jones · W. Jiang · J. A. Traherne ·
J. Trowsdale
Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK

M. R. López-Álvarez · W. Jiang · J. A. Traherne · J. Trowsdale
Cambridge Institute for Medical Research, University of Cambridge,
Cambridge Biomedical Campus, Wellcome Trust/MRC Building,
Hills Road, Cambridge CB2 0XY, UK

J. Trowsdale (✉)
Immunology Division, Pathology Department,
University of Cambridge, Cambridge CB2 1QP, UK
e-mail: jt233@cam.ac.uk

proteins have long cytoplasmic tails with immunoreceptor tyrosine-based inhibitory motifs (ITIM). Some LILR receptors interact with HLA class I (Group 1: LILRB1, -B2, -A1, -A2 and -A3) and they share significant amino acid sequence homology over class I-binding regions (Borges et al. 1997; Colonna et al. 1998; Fanger et al. 1998; Willcox et al. 2003). The ligands for the other receptors, including LILRB3-B5 and LILRA5-A6 are unknown, except for LILRA4 which binds CD317 (Cao et al. 2009; Tavano et al. 2013). These receptors contain non-conservative substitutions in class I binding regions identified in group 1(Willcox et al. 2003).

LILR are encoded within the leukocyte receptor complex (LRC) on chromosome 19q13.4 (Fig. 1), adjacent to the related killer immunoglobulin (Ig)-like receptor (KIR) genes (Barrow and Trowsdale 2008). KIR genes exhibit considerable sequence polymorphism (Robinson et al. 2010) and extensive copy number variation (CNV) (Jiang et al. 2012). LILRB3 and LILRA6 are paired receptors which potentially deliver opposing signals. The genes encoding these receptors display remarkable diversity within their extracellular domains when compared to either *LILRB1* or *LILRB2* (Colonna et al. 1997). For example, there is evidence for variation in the number of copies of *LILRA6* (Sudmant et al. 2010). In addition, it has been reported that some individuals lack *LILRA3* gene expression as a consequence of a large deletion of 6.7 kbp, resulting in the removal of four Ig domain and two leader peptide exons from the genomic sequence (Torkar et al. 2000).

To gain insight into the genetics and functions of the polymorphic *LILR* that do not bind class I, we characterised variation in *LILRB3* and *LILRA6* genes and examined them for CNV.

Materials and methods

Samples

Sequencing analysis was performed in cDNA samples from 20 healthy individuals from the Cambridge Blood Centre. Informed consent was obtained from all individuals. CNV assay was carried out in DNA samples from 48 human cell lines from the International Histocompatibility Working Group from different ethnic origins (see Table S1).

Sequencing

RNA extraction and subsequent cDNA synthesis was performed as described previously (Jones et al. 2009). Transcripts were initially amplified from macrophage cDNA using primers NK1078 and NK1091 for *LILRB3* (see Table S2 for primer details) and, NK1078 and NK636 for *LILRA6*. These initial polymerase chain reactions (PCRs) were performed using Phusion polymerase (Finnzymes) with the following cycling parameters: 30 s at 98 °C followed by 35 cycles of 98 °C for 10 s, 68 °C for 30 s, and 72 °C for 60 s. All PCRs were performed on the MJ Research (Reno, NV, USA) PTC-200 thermal cycler. Products were assessed by cycle sequencing using BigDye Terminator version 3.1 methodology (Applied Biosystems) and an Applied Biosystems 3730xl DNA analyser using the primers NK1078 and NK1495 (*LILRB3*), or NK1078, NK1404 and NK1494 (*LILRA6*) (Table S2).

To determine the sequences of single alleles within heterozygote individuals, we performed PCRs on individual cDNAs, using forward primers with specificity for single nucleotide polymorphisms (SNPs) (Table S2). These SNP-specific primers (SSP) were used in conjunction with either the *LILRB3*-specific reverse primer NK1091 or the *LILRA6*-specific NK636. The cycling conditions were: 180 s at 96 °C, followed by 5 cycles of 96 °C for 20 s, 70 °C for 45 s, and 72 °C for 25 s. This was followed by 31 cycles of 96 °C for 25 s, 65 °C for 50 s and 72 °C for 30 s; followed by 4 cycles of 96 °C for 30 s, 55 °C for 60 s and 72 °C for 90 s. All amplicons from PCR-SSP carried an M13R tag at the 5′ end, which was exploited for subsequent cycle sequencing using the primers M13R, in addition to NK1495, NK1404 and NK1494.

Sequence analysis: calculation of dN/dS ratios and statistical tests

To assess putative evolutionary selection pressure on polymorphism acting on the extracellular immunoglobulin-like region of *LILRB3* and *LILRA6* alleles, non-synonymous (dN) and synonymous (dS) substitution rates for all pairwise

Fig. 1 *LILR* genes. The LILR family is located in the LRC on chromosome 19q13.4 neighbouring the KIR family of genes. It is composed of two inverted clusters separated by *LAIR* and other genes. The transcription of each of these *LILR* clusters is in a head to tail fashion, as indicated by the *arrows*

comparisons of alleles from both loci were calculated using the KaKs Calculator (Zhang et al. 2006), applying Nei and Gojobori's method and incorporating Jukes–Cantor correction (Nei and Gojobori 1986).

One-tailed Fisher's exact tests for positive selection for sequence pairs were performed on the entire extracellular Ig region of *LILRB3* and *LILRA6* using Graphpad (http://www. graphpad.com/quickcalcs/) as previously described (Zhang et al. 1997). Results were considered statistically significant with $p < 0.05$.

Quantitative PCR

LILR copy number was determined by quantitative PCR (qPCR), on genomic DNA extracted from 48 human cell lines from the International Histocompatibility Working Group (http://www.ihwg.org/hla/) using the QIAamp DNA Blood Midi Kit (QIAGEN) following the manufacturer's instructions. The following genes were typed for CNV: *LILRA1*, *LILRA2*, *LILRA3*, *LILRA4*, *LILRA5*, *LILRA6*, *LILRB1*, *LILRB2*, *LILRB3*, *LILRB4* and *LILRB5*.

Forward and reverse primers, and a dual-labelled probe were designed to specifically amplify each LILR gene (Table S3), avoiding any allelic variation identified to date. *LILR* sequences were analysed for specificity using the primerBLAST tool from the National Centre for Biotechnology Information (http://www.ncbi.nlm.nih.gov/ tools/primer-blast). In addition, all reactions contained specific primers and a probe for the *STAT6* gene, which has two copies per diploid human genome, and was used as an endogenous reference gene. All reactions were performed in quadruplicate for each sample to increase the accuracy of copy number scoring.

A total of 10 ng of genomic DNA was amplified under the following PCR conditions: 5 min at 95 °C; followed by 40 cycles of 95 °C for 15 s and 66 °C for 50s; followed by 10s at 40 °C, using the LightCycler 480 System (Roche Diagnostics Ltd., Burgess Hill, UK). *LILR* copy number was determined by a quantitative PCR comparative Ct method (Schmittgen and Livak 2008).

SSP-PCR genotyping

Genotyping for presence of the SNP that encodes a threonine (T) at residue 94 within *LILRA6* was performed on genomic DNA samples using the primer pair NK1513 (5′-CCCCCT GGAGCTGGTGAC-3′) and NK559 (5′-TCATCAGAACAA AATGGTGATATCT-3′) or NK560 (5′-CATCAGAACAAA ATGGTGATATCC-3′). Detection of a 570-bp deletion located within the intergenic region between *LILRA6* and *LILRB3* was performed using the primer pair NK1520 (5′-CTGGTC CCTGCAGTGGCA-3′) and NK1518 (5′- GCCTTAGACT

TCCTATCCTGAAAC-3′). PCR cycling conditions were as previously described (Jones et al. 2006).

Results

LILRB3 and *LILRA6* immunoglobulin-like domains are highly polymorphic

By exploiting polymorphic sites and unique regions within the 3′ ends of *LILRB3* and *LILRA6*, we were able to characterise the sequences encoding the Ig domains of single alleles. This approach made it possible to identify individual allele sequences regardless of any variation in copy number (see later).

Analysis of the extracellular Ig domain-coding regions of the *LILRB3* and *LILRA6* genes, using cDNA samples from 20 healthy individuals, identified substantial variation in the nucleotide sequences of both receptors (Tables 1 and 2). This analysis generated 25 different sequences, 21 of which were novel and have subsequently been deposited in Genbank with accession numbers from KF294233 to KF294253.

We elucidated a total of 41 variable nucleotides in *LILRB3* and *LILRA6* sequences. Twenty-one of the non-synonymous and three of the synonymous substitutions affected the same nucleotide position in the coding sequence of both genes. None of these polymorphisms displayed significant linkage disequilibrium between *LILRB3* and *LILRA6* alleles (data not shown). The inhibitory receptor *LILRB3* encompassed 40 polymorphic sites with 34 non-synonymous and six synonymous substitutions. The activating receptor *LILRA6* showed less variation; 29 polymorphic sites, of which 25 were non-synonymous and four synonymous substitutions (Tables 1 and 2). Moreover, the analysis of these cDNA sequences revealed that four individuals carried three alleles of *LILRA6*, of which *LILRA6*01* and **03* were common to all the samples and in three out of the four samples, the third allele was *LILRA6*04*. *LILRA6*03* was present in the cDNA samples of all individuals with three alleles and it was the only *LILRA6* allele with the SNP 94 T (Table 1), suggesting that this SNP correlated with the increase in the number of copies of *LILRA6*. We performed SSP-PCR to detect the polymorphism at position 94 in the genomic DNA samples from cell lines but did not find any association between the presence of T at position 94 and the increase in the number of copies of *LILRA6* in our cohort (Table 3).

Protein evolution by selective pressure can be measured by considering the replacement of nucleotides within codons and by analysing site-by-site the dN/dS substitution ratio between sequences (Yang et al. 2000). This analysis can reveal the presence of either strictly conserved (dN/dS <1) or rapidly

Table 1 Predicted amino acids substitutions within the extracellular Ig region of *LILRB3* and *LILRA6* alleles

Allele	Ig 1								Ig 2										Ig 3						Ig 4										Freq
	30	36	38	46	65	67	92	94	97	99	120	126	128	142	152	164	178	182	225	229	265	267	291	294	295	300	318	325	326	327	377	382	386	394	
LILRB3 01	L	R	D	L	M	E	L	M	F	N	R	G	H	R	G	N	M	Q	D	D	P	H	N	M	A	D	V	W	W	Q	Y	H	F	M	0.3
02	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	V	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.2
03	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	R	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.15
04	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	N	–	–	–	–	–	–	–	–	–	–	–	–	0.1
05	Q	–	H	–	–	–	M	–	A	S	–	–	–	–	R	–	T	W	–	N	–	N	–	–	–	–	–	–	–	–	–	–	H	V	0.05
06	Q	–	H	–	–	–	M	–	A	S	–	–	–	–	R	–	T	W	–	N	–	N	–	–	–	–	–	–	–	–	–	–	H	–	0.05
07	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	H	–	0.05
08	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	G	–	–	–	–	–	–	–	–	0.05
09	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	R	–	–	–	–	–	–	–	–	–	–	–	–	–	0.35
10	–	Q	–	–	–	Q	–	–	A	S	–	R	–	–	–	–	–	R	–	–	–	–	–	–	–	–	–	–	–	–	R	Y	H	V	0.15
11	–	Q	–	–	–	Q	–	–	–	–	–	–	D	Q	–	–	–	R	–	H	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0.05
12	–	Q	–	W	–	Q	–	–	–	–	–	–	–	–	–	T	T	W	–	–	R	Y	D	I	T	–	M	R	G	Y	R	–	–	–	0.35
13	–	Q	–	W	–	Q	–	–	–	–	–	–	–	–	–	T	T	W	–	–	R	Y	D	I	T	–	M	R	G	Y	R	–	–	–	0.05
14	–	Q	–	W	–	Q	–	–	–	–	–	–	–	–	–	–	T	W	–	–	R	Y	D	I	T	–	M	R	G	Y	R	–	–	–	0.05
LILRA6 01	–	–	–	W	I	–	–	–	A	S	Q	R	–	–	–	–	–	R	–	–	–	–	–	–	–	–	–	R	G	Y	–	–	–	–	0.75
02	–	Q	–	W	–	Q	–	T	A	S	–	–	–	–	–	–	–	R	–	–	–	–	–	–	–	–	–	R	G	Y	L	–	–	–	0.05
03	–	Q	–	W	–	Q	–	–	–	–	–	–	–	–	–	T	–	R	–	–	–	–	–	–	–	–	–	R	G	Y	–	–	–	–	0.35
04	–	Q	–	W	–	Q	–	–	–	–	–	–	–	–	–	T	–	R	–	–	–	–	–	–	–	–	–	R	G	Y	–	–	–	–	0.3
05	–	Q	–	W	–	Q	–	–	–	–	–	–	–	–	–	T	–	–	–	–	–	–	–	–	–	–	–	R	G	Y	–	–	–	–	0.05
06	–	Q	–	W	–	Q	–	–	–	–	–	–	–	–	–	T	T	W	–	–	R	Y	–	–	–	–	–	R	G	Y	R	–	H	–	0.1
07	–	Q	–	W	–	Q	–	–	–	–	–	–	–	–	–	T	T	W	–	–	R	Y	D	–	–	–	–	R	G	Y	R	–	H	–	0.1
08	–	Q	–	W	–	Q	–	–	–	–	–	–	–	–	–	T	T	W	–	–	R	Y	D	–	–	–	–	R	G	Y	R	–	–	–	0.05
09	–	Q	–	W	–	Q	–	–	–	–	–	–	–	–	–	T	T	R	–	–	R	Y	D	–	–	–	–	R	G	Y	R	–	H	–	0.05
10	–	Q	–	W	–	Q	–	–	–	–	–	–	–	–	–	–	T	R	–	–	R	Y	D	I	T	–	M	R	G	Y	–	–	–	–	0.05
11	–	Q	–	W	–	Q	–	–	–	–	–	–	–	–	–	T	–	R	–	–	R	Y	D	I	T	–	M	R	G	Y	–	–	–	-	0.05

Residues are numbered in relation to the predicted mature protein sequence. The phenotype frequency of each allele is provided in the final column (*n*=20 individuals). *Ig* immunoglobulin domain

Table 2 Synonymous polymorphism within the extracellular Ig-coding region of *LILRB3* and *LILRA6* alleles

	Allele	Ig 1 70(A) GCG	Ig 2 98(Y) TAC	Ig 3 267(H/N/Y)[a] H: CAC	274(Y) TAT	287(S) AGC	Ig 4 326(G)[a] G: GGG	392(E) GAA
LILRB3	01	- - -	- - -	H: - - -	- - -	- - -	W	- - -
	01	- - -	- - -	H: - - -	- - -	- - -	W	- - -
	03	- - -	- - -	H: - - -	- - -	- - -	W	- - -
	04	- - -	- - -	N: A - T	- - -	- - -	W	- - -
	05	- - A	- - -	N: A - T	- - C	- - -	W	- - G
	06	- - A	- - -	N: A - T	- - C	- - -	W	- - -
	07	- - -	- - -	H: - - -	- - -	- - -	W	- - -
	08	- - -	- - -	H: - - -	- - -	- - -	W	- - -
	09	- - -	- - -	H: - - -	- - -	- - -	W	- - -
	10	- - A	- - T	N: A - T	- - C	- - -	W	- - G
	11	- - A	- - -	N: A - T	- - -	- - -	W	- - G
	12	- - A	- - -	Y: T - -	- - -	- - T	G: - - -	- - -
	13	- - A	- - -	Y: T - -	- - -	- - T	G: - - -	- - -
	14	- - A	- - -	Y: T - -	- - -	- - T	G: - - -	- - -
LILRA6	01	- - -	- - T	H: - - -	- - -	- - -	W	- - -
	02	- - A	- - T	H: - - -	- - -	- - -	W	- - -
	03	- - A	- - -	H: - - -	- - -	- - -	G: - - T	- - -
	04	- - A	- - -	H: - - -	- - -	- - -	G: - - T	- - -
	05	- - A	- - -	H: - - -	- - -	- - -	G: - - T	- - -
	06	- - A	- - -	H: - - -	- - -	- - -	G: - - T	- - G
	07	- - A	- - -	Y: T - -	- - -	- - -	G: - - T	- - G
	08	- - A	- - -	Y: T - -	- - -	- - -	G: - - T	- - -
	09	- - A	- - -	Y: T - -	- - -	- - -	G: - - T	- - G
	10	- - A	- - -	Y: T - -	- - -	- - -	G: - - -	- - -
	11	- - A	- - -	Y: T - -	- - -	- - -	G: - - -	- - -

Residue number, amino acid and codon sequences are displayed (residues are numbered in relation to the predicted mature protein sequence). In all cases, synonymous polymorphism occurs at the third position within each codon (*underlined*)

[a] Sites where both synonymous and non-synonymous polymorphisms occur in conjunction

evolving regions in genes (dN/dS >1). *LILRB3* and *LILRA6* sequences displayed in Tables 1 and 2 were analysed in this manner. There were predominantly non-synonymous polymorphisms, so the dN/dS ratios were consistent with positive selection (Fig. 2).

LILRA3 and LILRA6 genes, but not LILRB3, show variation in copy number

We designed assays to detect variation in the number of copies of *LILR* genes by qPCR. CNV was detected in the *LILRA3* and *LILRA6* loci (Fig. 3c and f, respectively). *LILRA3* varied between zero, one or two copies per diploid genome. Around 21 % of the samples lacked *LILRA3*, while ~19 % had only one copy of this gene. CNV in *LILRA6* varied between one, two, three and four copies per diploid genome; ~8 % of the

samples had a single copy of the gene, whilst ~33 % carried a duplication of *LILRA6* (three or four copies). Despite its high rate of allelic variation, *LILRB3* had a constant number of two copies per genotype (Fig. 4c).

We identified a fosmid sequence deposited in Genbank (accession number: AC236241) that contains two sequences of *LILRA6*, the arrangement of which is represented in Fig. 5. The additional *LILRA6* locus (*LILRA6* [2]) occurred ~15 kb upstream of the common *LILRA6* locus (*LILRA6* [1]) within the region normally occupied by *LILRB3*. Sequence AC236241 also possesses a deletion within a tandem repeat sequence that is normally located within the intergenic region between *LILRA6* and *LILRB3*. We used the Tandem repeats finder program (Benson 1999) to detect the 33-bp tandem repeat sequence, which has a consensus sequence of TCTA TTGAGATCCTATGGAGGTCCTGTGGGGGT and a copy

Table 3 *LILRA6* CNV and presence of tandem repeat sequence and SNP 94 T

Cell line	CN	TRS	94 T	Cell line	CN	TRS	94 T
KT12	2.89	Deletion	Negative	RSH	3.32	Deletion	Negative
TAB089	2.06	Complete	Negative	TL	3.04	Deletion	Positive
XL1-ND	2.31	Complete	Negative	PAR	2.06	Complete	Negative
RML	3.02	Deletion	Negative	WT24	3.19	Deletion	Positive
KY	1.68	Complete	Negative	SA	2.10	Complete	Positive
HOM2	1.76	Complete	Negative	LATIF	1.86	Complete	Negative
KT17	2.01	Deletion	Negative	OZB	1.05	Complete	Negative
JESTHOM	1.98	Complete	Negative	AMAI	2.84	Complete	Positive
AKIBA	1.92	Complete	Negative	COX	1.96	Complete	Positive
WWPAK	1.90	Complete	Negative	KAS011	0.38	Complete	Negative
GRC	3.28	Deletion	Negative	CO47	2.11	Deletion	Positive
BM9	2.27	Complete	Negative	FPAF	1.97	Deletion	Negative
SLE005	2.10	Deletion	Positive	SNA BLL	2.01	Complete	Negative
DUCAF	3.09	Deletion	Negative	DUGE	2.98	Deletion	Negative
PMH 161	2.00	Complete	Negative	WJR076	1.93	Complete	Negative
LB	1.95	Complete	Negative	VAVY	2.01	Complete	Negative
BTB	1.96	Deletion	Negative	HOR	1.10	Complete	Negative
IBW9	3.09	Deletion	Positive	MANIKA	3.89	Deletion	Positive
DKB	3.13	Deletion	Positive	MGAR	2.32	Deletion	Negative
DEM	3.08	Deletion	Positive	H0104	1.84	Deletion	Negative
LBF	2.90	Deletion	Positive	M7	2.70	Deletion	Negative
BOLETH	1.16	Complete	Negative	LCL 721	2.13	Complete	Positive
HO301	2.01	Deletion	Positive	CRGE	2.04	Complete	Negative
31708	3.75	Deletion	Negative	MCF	1.99	Deletion	Positive

CN calculated copy number, *TRS* tandem repeat sequence

number of 19.9. This repeat was absent from AC236241. By utilising PCR primers that flank this deletion (Fig. 5b), we found that its presence correlated strongly with the duplication of *LILRA6* within our cohort, occurring in around 93 % of the samples with more than two copies of *LILRA6* (Table 3), indicating that fosmid AC236241 represents a commonly occurring haplotype that carries a duplication of *LILRA6*.

No other *LILR* gene shows CNV

To determine whether the CNV we observed was a more widespread feature of *LILR* genes, we carried out CNV analysis of all other activating and inhibitory *LILR* genes (*LILRA1*, *-A2*, *-A4*, *-A5*, *-B1*, *-B2*, *-B4* and *-B5*). This analysis did not identify any further variation in the number
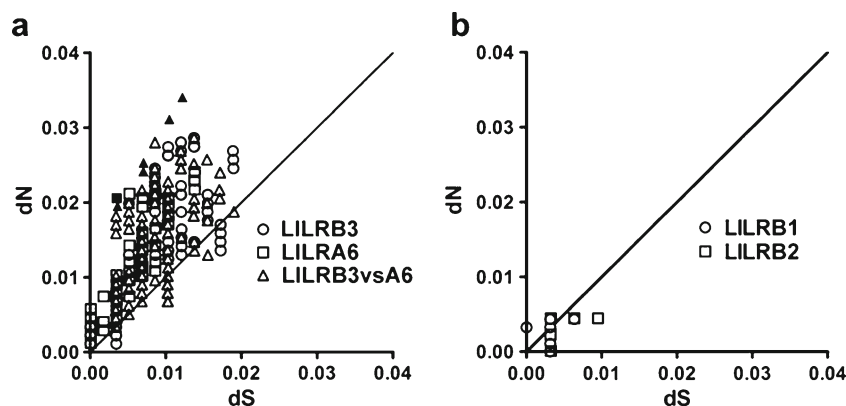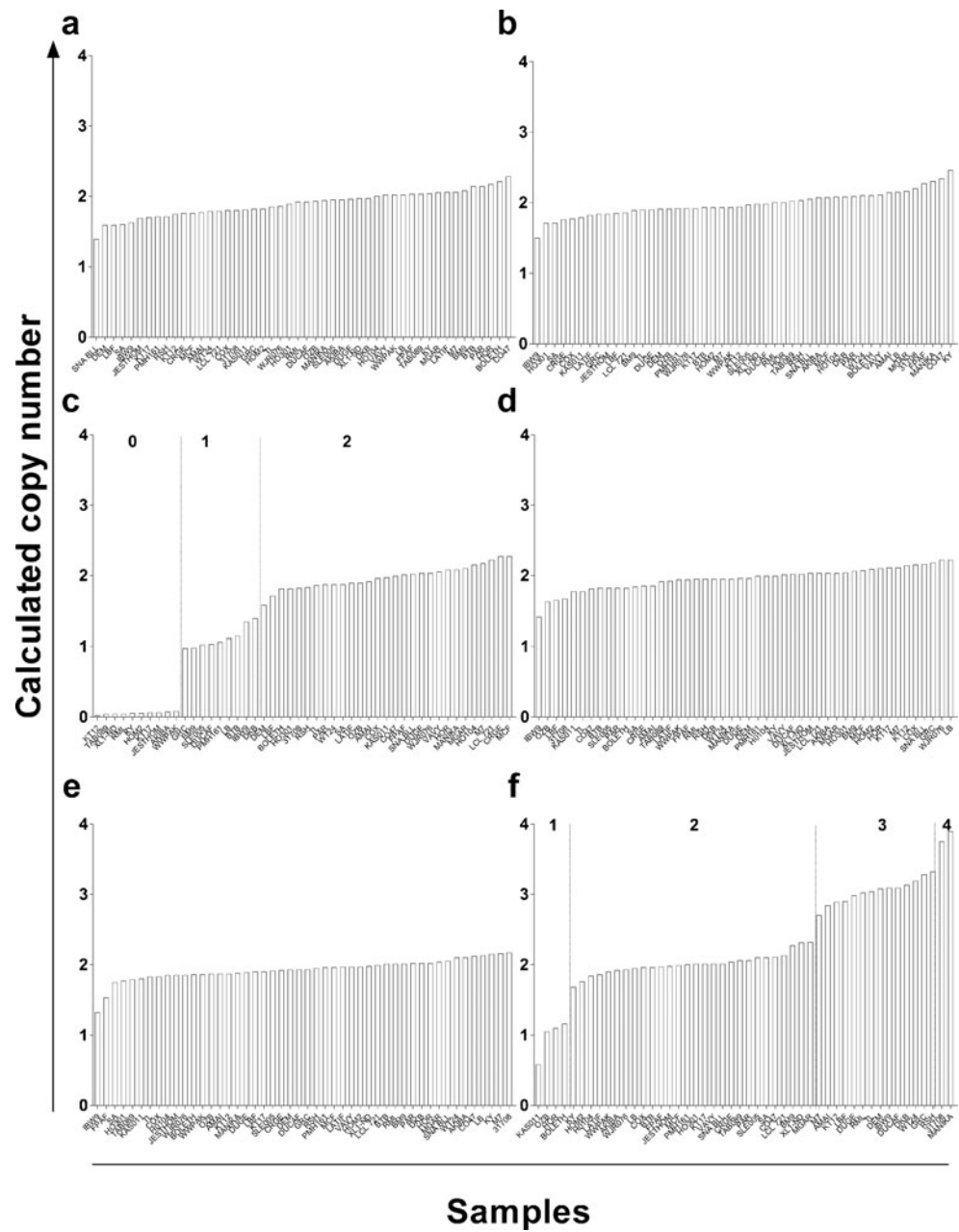

Fig. 2 dN/dS ratios of the extracellular regions of *LILRB3* and *LILRA6* indicate positive selection. **a** dN and dS values from allelic pairwise comparisons of each gene, *LILRB3* (*circles*) and *LILRA6* (*squares*), and inter-gene comparison (*triangles*). Only sequences detected in this study were included in the analysis. The *central plot line* represents balancing selection (where dN=dS). *Shaded symbols* indicate significance using a Fisher's exact test for positive selection pressure (Zhang et al. 1997). **b** dN and dS values for previously reported *LILRB1* (*circles*) and *LILRB2* (*squares*) sequences deposited in Genbank (http://www.ncbi.nlm.nih.gov/genbank/)

Fig. 3 Copy number variation of activating *LILR* receptor genes. Summaries of the results of copy number assays for *LILRA1* (**a**), *LILRA2* (**b**), *LILRA3* (**c**), *LILRA4* (**d**), *LILRA5* (**e**) and *LILRA6* (**f**). The number of copies per genome was constant (two copies) in activating LILR receptors with the exception of *LILRA3* and *LILRA6,* which displayed variation. *LILRA3* showed from zero to two copies per genome. *LILRA6* showed from one to four copies per genome. Each bar represents the calculated copy number obtained for one sample (see Table S1). Number inside plot is the predicted copy number for samples

Discussion

Amplification and sequencing of *LILRB3* and *LILRA6* cDNA from 20 individuals returned 25 different sequences with 41 polymorphic sites, consistent with these receptors being highly polymorphic in their extracellular domains (Colonna et al. 1997). Most of variable positions were represented in both *LILRB3* and *LILRA6*, including several synonymous

of copies of other *LILR* genes (Figs. 3 and 4), indicating that CNV is specific to the *LILRA3* and *LILRA6* loci, or rare for other genes.

substitutions. This indicates the possibility of genetic transfer between these two genes and the additional copies of *LILRA6*. The patchwork nature of the variation is reminiscent of the variation in some MHC class I or class II genes, in which case it may be achieved by allele or gene conversion (Traherne et al. 2006). Several non-synonymous polymorphisms within *LILRB3* and *LILRA6* occur at positions known to form part of the binding sites of Group 1 LILR to MHC class I: 36, 38, 67, 97, 99 and 126 (Willcox et al. 2003; Yang and Bjorkman 2008). Although no specific ligand for both LILRB3 and LILRA6 receptors has yet been identified, it is tempting to speculate that these variations may have functional consequences. Additionally, we found three alleles of *LILRA6* in
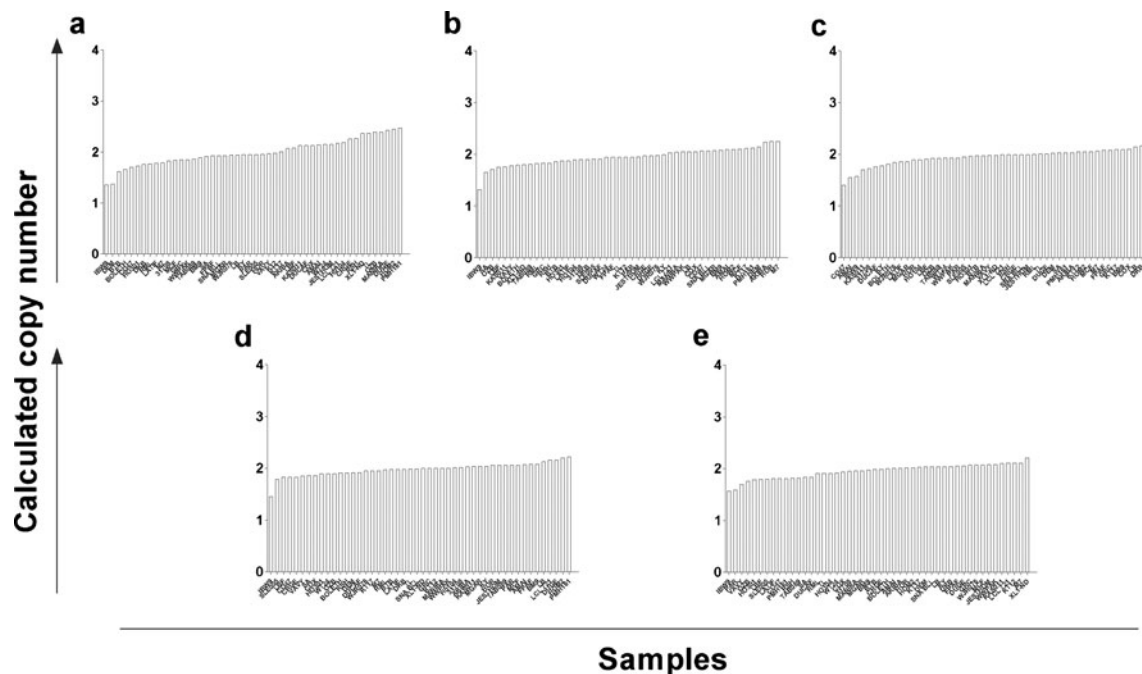
Samples

Fig. 4 Copy number variation of inhibitory *LILR* receptor genes. Copy number variation was analysed for the inhibitory receptor genes *LILRB1* (**a**), *LILRB2* (**b**), *LILRB3* (**c**), *LILRB4* (**d**) and *LILRB5* (**e**). The figure shows the results of this analysis where each of these receptor genes are present in two copies per genotype. Each bar represents the calculated copy number obtained for one sample (see Table S1)

four cDNA samples, indicating that the duplicated copies of this gene are expressed and are presumably functional. The fact that three out of four samples had the combination of *LILRA6*03* and *LILRA6*04* concurs with the duplication of *LILRA6* found in a fosmid (accession number AC236241). However, this allele pairing was not found in all tri-allelic individuals. Moreover, the SNP 94 T, characteristic of *LILRA6*03,* which was carried by all four individuals with

three alleles, was not associated with the increase in the number of copies of *LILRA6* found in cell lines. Therefore, it would appear that the allele content within the haplotype comprising duplicated *LILRA6* genes is variable.

The method of CNV analysis we describe has a number of advantages: it is simple to carry out, low cost, specific and sensitive and is high-throughput, as up to a thousand samples can be analysed per week. Variation in *LILR* gene copy
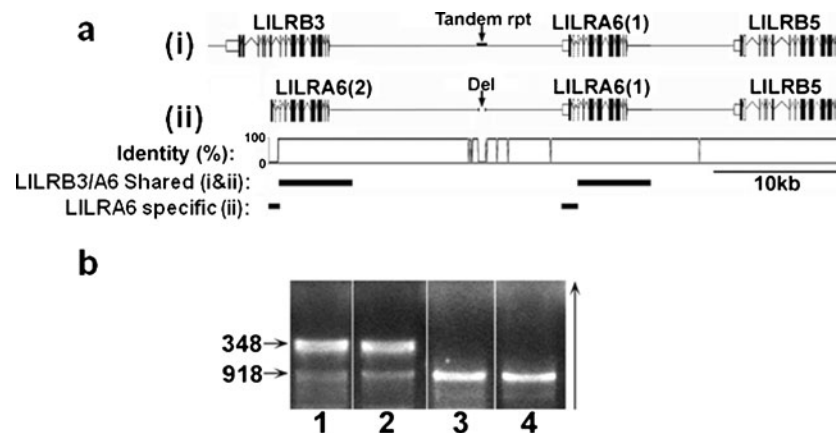

Fig. 5 Genetic organisation of the centromeric region of the *LILR* gene cluster encompassing *LILRB5*, *LILRB3* and *LILRA6*. **a** Genes are arranged according to their positions as determined by analysis of the representative sequence CU151838 (i) and within fosmid AC236241 (ii). Sequence AC236241 carries two copies of *LILRA6* (labelled *1* and *2*) and a 570-bp deletion within the tandem repeat region present in CU151838. The comparative homology between these two sequences is also presented. Regions occurring within both *LILRB3* and *LILRA6* that share >95 % identity are indicated, together with the locations of *LILRA6*-specific sequence within AC236241. The AC236241 sequence terminates within the 3′ region of *LILRA6* (2). **b** PCR-based detection of the deleted region identified within fosmid AC236241. Amplification of the tandem repeat sequence between *LILRB3* and *LILRA6* (918 bp) or two *LILRA6* sequences (348 bp). Samples are DEM (*1*), DUCAF (*2*), HOR (*3*), and SA (*4*). The *arrow* indicates the direction of DNA migration. Refer to Table 3 for information about the results of *LILRA6* copy number, the deletion within the tandem repeat sequence and results for the SNP 94 T by SSP-PCR

number was detected within *LILRA3*, which encodes a soluble molecule (Borges et al. 1997; Colonna et al. 1999), and the activating *LILRA6*, while all inhibitory *LILR* genes, including the highly polymorphic *LILRB3*, had a constant number of two copies per diploid genome. The gene frequency of deletion of *LILRA3* was around 40 %, in concordance with previously published results where the frequency of deletions varied from 6 % to 84 % in different populations (Torkar et al. 2000; Wiśniewski et al. 2013). *LILRA6* was more variable; 8.3 % of the samples showed a heterozygous deletion of *LILRA6*, and in 33.3 % of the cases, we observed duplications of this gene (three or four copies).

The variation in the number of copies of *LILRA6* could be due to non-allelic homologous recombination (NAHR) or crossing-over between sequences that are not in allelic positions, a phenomenon proposed for CNV in *KIR* (Traherne et al. 2010). These events are mediated by genomic structures that are flanked by paralogous repeat sequences, low-copy repeats, segmental duplications or tandem repeat regions (Jeffreys et al. 2004; Liu et al. 2012). We found a tandem repeat sequence between *LILRB3* and *LILRA6* genes (Fig. 5a), that was deleted in sequence contained in fosmid AC236241, which carries a duplication of *LILRA6*. Therefore, it could be hypothesised the deletion within the tandem repeat region may be mechanistically linked to the duplication event of *LILRA6* within this haplotype. From analysis of the AC236241 sequence an obvious insertion point of the duplicated *LILRA6* gene could not be identified indicating that unequal crossover could have occurred within the ~4.9-kb homologous region that *LILRA6* shares with *LILRB3* (Fig. 5a). Recombination centred upon the repeat region may account for the lack of significant linkage disequilibrium observed between alleles of *LILRB3* and *LILRA6*, as determined by the analysis of cDNA from 20 individuals (data not shown).

Our CNV screen indicates that *LILRA6* is absent on some haplotypes. There are two sequences in Genbank (accession numbers: AC235034 and AC153469) that lack *LILRA6*, and instead carry *LILRB3* at that genomic location. The *LILRB3* allele within AC235034 encodes a predicted protein sequence that is identical to *LILRA6*01* throughout its extracellular region, while the predicted protein encoded by AC153469 only differs with *LILRA6*01* due to the presence of 164 T. *LILRA6*01* is the most frequently occurring allele according to our cDNA screen (Table 1) and carries two unique SNPs (65 I and 120 Q), and a unique combination of variable positions 36 R, 46 W and 67 E. We did not detect these polymorphic features within any *LILRB3* allele, suggesting that they are predominantly a characteristic of *LILRA6*. Their presence within the *LILRB3* sequences of AC153469 and AC235034 is evidence of an unequal crossover event that has occurred within the homologous region following the Ig2-encoding exon of *LILRA6*, resulting in the replacement of *LILRA6*-

specific sequence within the 3′ of the gene with that of *LILRB3*. This indicates that unequal crossover is a mechanism by which genetic information may be exchanged between these two genes (Fig. 6a). An additional possibility could be the formation of single-stranded DNA (ssDNA) secondary structures or hairpin loops in the repeat tandem sequence (Fig. 6b). These structures are accessible substrates for nucleases, making them susceptible to strand breaks. This may lead to repair processes that can cause deletion of regions or genetic transfer of homologous sequences between chromosomes. Such recombination processes have previously been proposed in the evolution of the KIR gene cluster (Traherne et al. 2010).

Functionally, the increase in the number of copies of *LILRA6* could affect the balance between activating and inhibitory signals after ligand binding. How could the combination of CNV and variable sequences be interpreted in terms of function? So far, it is difficult to identify haplotype-specific combinations of variable sequences over the sets of genes, as we do not have sequence information over single haplotypes. If these exist they would be consistent with haplotype-specific interactions with a ligand and integration of signals from the different receptors, as may be the case for some neighbouring KIR A and B haplotypes. They would suggest that the
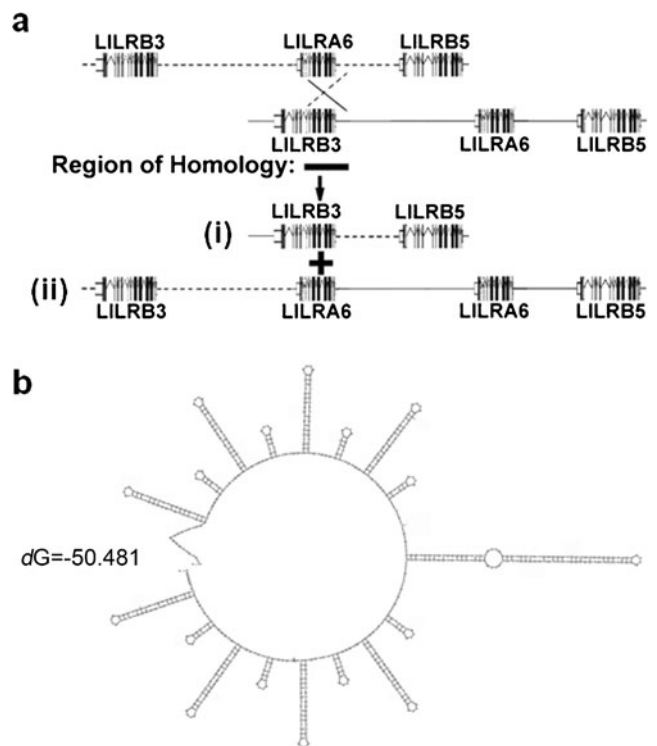

Fig. 6 Non-allelic homologous recombination and potential ssDNA secondary structures produced at the tandem repeat sequence. **a** The process of unequal crossing-over of homologous sequences proposed to generate deletions (i) or duplications (ii) of *LILRA6*. **b** Seventeen hairpin structures were predicted using the Mfold program (Zuker 2003). Twenty copies of the repeat tandem sequence TCTATTGAGATCCTATGGAG GTCCTGTGGGGGT was used as the input

constellations of variable activating and inhibitory receptors on haplotypes are driven by interaction with a variable ligand, such as a variable pathogen.

In conclusion, we have shown high levels of polymorphism and homology between *LILRB3* and *LILRA6* in their extracellular domains, with *LILRA6* also displaying alterations in the number of copies. All the variations observed in these genes may affect ligand binding, the equilibrium between activating and inhibitory signals, and therefore, the balance of immune response. The allelic variation and CNV in the *LILRB3/LILRA6* pair we have described is consistent with strong selection for variation and high dN/dS ratios are characteristic of other immune system genes, such as MHC or KIR, that are involved with resistance to pathogens (Park et al. 2012).

Conflict of interest   The authors declare they have no conflict of interest.

References

Anderson KJ, Allen RL (2009) Regulation of T-cell immunity by leucocyte immunoglobulin-like receptors: innate immune receptors for self on antigen-presenting cells. Immunology 127:8–17

Barrow AD, Trowsdale J (2008) The extended human leukocyte receptor complex: diverse ways of modulating immune responses. Immunol Rev 224:98–123

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acid Res 27(2):573–580

Borges L, Hsu ML, Fanger N, Kubin M, Cosman D (1997) A family of human lymphoid and myeloid Ig-like receptors, some of which bind to MHC class I molecules. J Immunol 159:5192–5196

Brown D, Trowsdale J, Allen R (2004) The LILR family: modulators of innate and adaptive immune pathways in health and disease. Tissue Antigens 64(3):215–225

Brown DP, Jones DC, Anderson KJ et al (2009) The inhibitory receptor LILRB4 (ILT3) modulates antigen presenting cell phenotype and, along with LILRB2 (ILT4), is upregulated in response to *Salmonella* infection. BMC Immunol 10:56

Cao W, Bover L, Cho M et al (2009) Regulation of TLR7/9 responses in plasmacytoid dendritic cells by BST2 and ILT7 receptor interaction. J Exp Med 206(7):1603–1614

Chang CC, Ciubotariu R, Manavalan JS et al (2002) Tolerization of dendritic cells by T(S) cells: the crucial role of inhibitory receptors ILT3 and ILT4. Nat Immunol 3(3):237–243

Chang CC, Liu Z, Vlad G et al (2009) Ig-like transcript 3 regulates expression of proinflammatory cytokines and migration of activated T cells. J Immunol 182:5208–5216

Colonna M, Navarro F, Bellon T et al (1997) A common inhibitory receptor for major histocompatibility complex class I molecules on human lymphoid and myelomonocytic cells. J Exp Med 186:1809–1818

Colonna M, Samaridis J, Cella M et al (1998) Human myelomonocytic cells express an inhibitory receptor for classical and nonclassical MHC class I molecules. J Immunol 160:3096–3100

Colonna M, Nakajima H, Navarro F, Lopez-Botet M (1999) A novel family of Ig-like receptors for HLA class I molecules that modulate function of lymphoid and myeloid cells. J Leukoc Biol 66:375–381

Fanger NA, Cosman D, Peterson L, Braddy SC, Maliszewski CR, Borges L (1998) The MHC class I binding proteins LIR-1 and LIR-2 inhibit Fc receptor-mediated signaling in monocytes. Eur J Immunol 28:3423–3434

Huynh OA, Hampartzoumian T, Arm JP et al (2007) Down-regulation of leucocyte immunoglobulin-like receptor expression in the synovium of rheumatoid arthritis patients after treatment with disease-modifying anti-rheumatic drugs. Rheumatology (Oxford) 46(5):742–751

Jeffreys AJ, Holloway JK, Kauppi L et al (2004) Meiotic recombination hot spots and human DNA diversity. Philos Trab R Soc London B Biol 59(1441):141–152

Jiang W, Johnson C, Jayaraman J et al (2012) Copy number variation leads to considerable diversity for B but not A haplotypes of the human KIR genes encoding NK cell receptors. Genome Res 22(10):1845–1854

Jones DC, Edgar RS, Ahmad T et al (2006) Killer Ig-like receptor (KIR) genotype and HLA ligand combinations in ulcerative colitis susceptibility. Genes Immun 7(7):576–582

Jones DC, Roghanian A, Brown DP et al (2009) Alternative mRNA splicing creates transcripts encoding soluble proteins from most LILR genes. Eur J Immunol 39(11):3195–3206

Kabalak G, Dobberstein SB, Matthias T, Reuter S, The YH, Dörner T, Schmidt RE, Witte T (2009) Association of immunoglobulin-like transcript 6 deficiency with Sjögren's syndrome. Arthritis Rheum 60(10):2923–2925

Kim-Schulze S, Scotto L, Vlad G et al (2006) Recombinant Ig-like transcript 3-Fc modulates T cell responses via induction of Th energy and differentiation of CD8[+] T suppressor cells. J Immunol 176(5):2790–2798

Koch S, Goedde R, Nigmatova V et al (2005) Association of multiple sclerosis with ILT6 deficiency. Genes Immun 6:445–447

Kuroki K, Tsuchiya N, Shiroishi M et al (2005) Extensive polymorphisms of LILRB1 (ILT2, LIR1) and their association with HLA-DRB1 shared epitope negative rheumatoid arthritis. Hum Mol Genet 14(16):2469–2480

Lee DJ, Sieling PA, Ochoa MT et al (2007) LILRA2 activation inhibits dendritic cell differentiation and antigen presentation to T cells. J Immunol 179(12):8128–8136

Liu P, Carvalho CM, Hastings PJ, Lupski JR (2012) Mechanisms for recurrent and complex human genomic rearrangements. Curr Opin Genet Dev 22(3):211–220

Manavalan JS, Rossi PC, Vlad G et al (2003) High expression of ILT3 and ILT4 is a general feature of tolerogenic dendritic cells. Transpl Immunol 11(3–4):245–258

Nakajima H, Samaridis J, Angman L, Colonna M (1999) Human myeloid cells express an activating ILT receptor (ILT1) that associates with Fc receptor gamma-chain. J Immunol 162(1):5–8

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3(5):418–426

Ordóñez D, Sánchez AJ, Martínez-Rodríguez JE et al (2009) Multiple sclerosis associates with LILRA3 deletion in Spanish patients. Genes Immun 10(6):579–585

Park S, Yang JS, Kim J et al (2012) Evolutionary history of human disease genes reveals phenotypic connections and comorbidity among genetic diseases. Sci Rep 2:757

Pilsbury LE, Allen RL, Vordermeier M (2010) Modulation of Toll-like receptor activity by leukocyte Ig-like receptors and their effects during bacterial infection. Mediat Inflamm 2010:536478

Robinson J, Mistry K, McWilliam H, Lopez R, Marsh SG (2010) IPD — the Immuno Polymorphism Database. Nucleic Acids Res 38:D863–D869

Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 3:1101–1108

Sudmant PH, Kitzman JO, Antonacci F et al (2010) Diversity of human copy number variation and multicopy genes. Science 330(6004):641–646

Tavano B, Galao RP, Graham DR et al (2013) Ig-like transcript 7, but not bone marrow stromal cell antigen 2 (also known as HM1.24, tetherin, or CD317), modulates plasmacytoid dendritic cell function in primary human blood leukocytes. J Immunol 190(6):2622–2630

Torkar M, Haude A, Milne S, Beck S, Trowsdale J, Wilson MJ (2000) Arrangement of the ILT gene cluster: a common null allele of the ILT6 gene results from a 6.7-kbp deletion. Eur J Immunol 30:3655–3662

Traherne JA, Horton R, Roberts AN et al (2006) Genetic analysis of completely sequenced disease-associated MHC haplotypes identifies shuffling of segments in recent human history. PLoS Genet 2(1):e9

Traherne JA, Martin M, Ward R et al (2010) Mechanisms of copy number variation and hybrid gene formation in the KIR immune gene complex. Hum Mol Genet 19(5):737–751

Willcox BE, Thomas LM, Bjorkman PJ (2003) Crystal structure of HLA-A2 bound to LIR-1, a host and viral major histocompatibility complex receptor. Nat Immunol 4(9):913–919

Wiśniewski A, Wagner M, Nowak I et al (2013) 6.7-kbp deletion in LILRA3 (ILT6) gene is associated with later onset of the multiple sclerosis in a Polish population. Hum Immunol 74(3):353–357

Yang Z, Bjorkman PJ (2008) Structure of UL18, a peptide-binding viral MHC mimic, bound to a host inhibitory receptor. Proc Natl Acad Sci U S A 105(29):10095–10100

Yang ZH, Nielsen R, Goldman N, Pedersen AMK (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155:431–449

Zhang J, Kumar S, Nei M (1997) Small-sample tests of episodic adaptive evolution: a case study of primate lysozymes. Mol Biol Evol 14(12):1335–1338

Zhang Z, Li J, Zhao XQ, Wang J, Wong GK, Yu J (2006) KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. Genomics ProteomE Bioinforma 4(4):259–263

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31(13):3406–3415