IPD—the Immuno Polymorphism Database

James Robinson[1], Kavita Mistry[1], Hamish McWilliam[2], Rodrigo Lopez[2] and Steven G. E. Marsh[1,3,*]

[1]Anthony Nolan Research Institute, Royal Free Hospital, Pond Street, Hampstead, London, NW3 2QG, [2]EMBL Outstation, The European Bioinformatics Institute (EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD and [3]UCL Cancer Institute, University College London, Royal Free Campus, Pond Street, Hampstead, London, NW3 2QG, UK

ABSTRACT

The Immuno Polymorphism Database (IPD) (http://www.ebi.ac.uk/ipd/) is a set of specialist databases related to the study of polymorphic genes in the immune system. The IPD project works with specialist groups or nomenclature committees who provide and curate individual sections before they are submitted to IPD for online publication. The IPD project stores all the data in a set of related databases. IPD currently consists of four databases: IPD-KIR, contains the allelic sequences of Killer-cell Immunoglobulin-like Receptors, IPD-MHC, is a database of sequences of the Major Histocompatibility Complex of different species; IPD-human platelet antigens, alloantigens expressed only on platelets and IPD-ESTDAB, which provides access to the European Searchable Tumour cell-line database, a cell bank of immunologically characterised melanoma cell lines. The data is currently available online from the website and ftp directory.

INTRODUCTION

The study of the immune system constitutes many different complex areas of research. The Immuno Polymorphism Database (IPD) is a set of specialist databases related to the study of polymorphic genes of the immune system. The IPD project works with specialist groups or nomenclature committees, which each curate a different section of the project. IPD currently consists of four databases: IPD-KIR, contains the allelic sequences of Killer-cell Immunoglobulin-like Receptors (KIR); IPD-MHC, is a database of sequences of the Major Histocompatibility Complex (MHC) of different species; IPD-human platelet antigens (IPD-HPA), alloantigens expressed only on platelets and IPD-ESTDAB, which

provides access to the European Searchable Tumour cell-line database (ESTDAB), a cell bank of immunologically characterised melanoma cell lines. By providing a centralised resource for the work of different groups it is hoped we can bring together similar data to aid in the study and analysis of this area.

This is done by contacting various expert groups, such as the nomenclature committees for certain genes, and inviting them to contribute their data to a central resource. The individual nomenclature committees are all established within their own field, their continuing role being the identification and naming of new alleles based on the submission of new sequences to generalist databases like the European Molecular Biology Laboratory's nucleotide sequence database (also known as EMBL-Bank) (1), the National Center for Biotechnology Information's GenBank (2) and the DNA DataBank of Japan (DDBJ) (3). The main difference between the data held within a specialist system like IPD and the generalist databases is the further curation of the files by experts in the relevant field. This additional step allows for improvements in data quality and the addition of more specialised information. This may result in there being differences in the entries kept in the different databases; in this case the entry in IPD should be considered the most accurate. The inclusion of some nomenclature committees has meant the online publication of sequence alignments for the first time. This is particularly important when it comes to the sequences of polymorphic genes. Also because the IPD project is able to provide the work of different committees in a common format, it is easier to compare the sequences of different species.

The IPD project stores all the data in a set of related databases. Those sections with similar data, IPD-KIR and IPD-MHC share the same database structure. The sharing of a common database structure makes it easier to implement common tools for data submission and retrieval. Other unrelated sections like IPD-ESTDAB currently have their own unique structure.

---

*To whom correspondence should be addressed. Tel: +44 (0) 20 7284 8321; Fax: +44 (0) 20 7284 8331; Email: steven.marsh@ucl.ac.uk

IPD-KIR

The KIR are members of the immunoglobulin super family (IgSF) formerly called Killer-cell Inhibitory Receptors. KIRs have been shown to be highly polymorphic both at the allelic and haplotypic levels (4). They are composed of two or three Ig-domains, a transmembrane region and cytoplasmic tail, which can in turn be short (activatory) or long (inhibitory). The Leukocyte Receptor Complex (LRC), which encodes KIR genes, has been shown to be polymorphic, polygenic and complex in a manner similar to the MHC. Because of the complexity in the KIR region and KIR sequences a KIR Nomenclature Committee was established in 2002, to undertake the naming of KIR allele sequences. The first KIR Nomenclature report was published in 2003 (5), which coincided with the first release of the IPD-KIR database. The number of officially named KIR alleles has increased since the initial release which contained 89 alleles. As of September 2009, there are over 450 alleles, which code for over 230 unique protein sequences. This has meant that since its initial release there have been nine further releases of the IPD-KIR database, the latest being in February 2009.

The online tools available for IPD-KIR include those for performing allele queries, sequence alignments and cell queries. As the database is based on the work of a nomenclature committee, the website includes links to a portable document format (PDF) file of recent nomenclature reports. From the data contained within these reports the database is also able to provide individual allele reports (Figure 1). These pages contain the official allele name, any previous designations, the EMBL, GenBank, or DDBJ accession number(s) and a reference linked wherever possible to the PubMed abstract. Where possible additional details on the source of sequence are also provided. This source material is normally in the form of a cell-line or DNA from which each allele in the database was isolated and characterised and it's availability. The information contained within this dataset can be searched independently from the allele data.

Within each IPD section alleles of a particular gene may differ from each other by as little a single nucleotide. With sometimes hundreds of alleles for a particular gene we need to be able to graphically represent where these differences lie. These polymorphic positions are also often conserved within certain groups of alleles. For this reason the sequences are displayed as multiple sequence alignments which highlight the polymorphic positions. These alignments allow a visual interpretation of sequence similarity, so that polymorphic positions can easily be identified and motifs found in multiple alleles are easily identified. The sequence alignments are available via a link from the section homepage. The sequence alignment tool uses the same basic interface for both IPD-KIR database and IPD-MHC. The interface provided (Figure 2) lets the user define a number of key variables for the alignments, before producing an online output, which can be printed or downloaded. The first step in any alignment is to select the locus of interest. The tool provides a drop-down list of all loci. The selection of a

locus automatically updates the list of features, which can be aligned, as well as the default reference sequence used for the alignment. The types of feature available for alignment are the nucleotide coding sequence and individual exons, the signal peptide, mature protein and full-length protein sequence. We also include the genomic sequence, which covers some of the 5′ and 3′ untranslated regions, exons and introns into the database for over half of the alleles in the KIR system. The alignment tool options also allow the user to display a subset of alleles of a particular locus, omit alleles unsequenced for a particular region and also to align against a particular reference or consensus sequence. The alignment tool uses standard formatting for the display of sequence alignments. The alignment tool does not perform a sequence alignment each time it is used, but it extracts pre-aligned sequences, allowing for faster access. The alignments adhere to a number of conventions for displaying evolutionary events and numbering. The numbering of the alignments is based upon the sequence of the reference allele. For a nucleotide sequence, the A of the initiation Methionine codon is denoted nucleotide $+1$ and the nucleotide 5′ to $+1$ is numbered -1. There is no nucleotide zero (0). All numbering is based on the ATG of the reference sequence. If a nucleotide sequences is displayed in codons, then the protein numbering is applied. For amino acid-based alignments, the first codon of the mature protein, after cleavage of the signal sequence is labelled codon 1 and the codon 5′ to this is numbered -1. In all sequences the following conventions are used. Where identity to the reference sequence is present the base will be displayed as a hyphen (-). Non-identity to the reference sequence is shown by displaying the appropriate base at that position. Where an insertion or deletion has occurred, this is be represented by a period (.). If the sequence is unknown at any point in the alignment, this is be represented by an asterisk (*). In protein alignments for null alleles, the 'Stop' codons are represented by an X and the sequence following the termination codon, is not marked up and appears blank. The flexibility of the new alignment tool means that unlike in previous alignments you can now display a small subset of sequences against an allele of your choice, using a number of display options (Figure 3).

Further recent additions to the tools available from the IPD-KIR site include a KIR ligand calculator (6). The ligand calculator allows the user to define which KIR ligands are present in a transplant setting based on the HLA typing of a patient and prospective donor. This is because recent transplant strategies based on KIR-ligand mismatch to predict NK cell alloreactivity have resulted in less relapse, less GvHD and better overall survival in patients with acute myeloid leukaemia (AML) (7). The KIR-ligands are HLA molecules that can be grouped into three major categories based on the amino acid sequence determining the KIR-binding epitopes in HLA-C and HLA-B molecules. All expressed HLA-C alleles are of the C1 or C2 group (8) and most HLA-B alleles can be classified as either Bw4 or Bw6 (9). The receptors KIR2DL1, KIR2DL2 and KIR3DL1 bind KIR-ligand C2, C1 and Bw4, respectively, resulting in

2DL1*002			
Previous Designations & Local Names:	p58 cl-42		
IPD Acc No:	KIR00002 (Links to SRS Entry)		
Assigned:	28-JUN-2003	Sequence Last Modified:	28-JUN-2003
Source Entries:	AY789055 , CU459007 , DQ666832 , U24076		
Cells Sequence Derived From:	COX , J56 , NK-CM , RDP35		
Known Ethnic Origin Of Source:	Caucasoid, Oriental		

Publications

Wagtmann N *et al. Immunity* 2:439-49 1995. abstract @
Hou L *et al. Tissue Antigens* 69:250-4 2007. abstract @

Confirmatory Data

Sequence Status:	Complete CDS	Confirmed:	Confirmed
Cells:	4	Laboratories:	4

Protein Sequence (348 aa)

```
MSLLFVSMACVGFFLLQGAWPHEGVHRKPSLLAHPGPLVKSEETVILQCWSDVMFEHFLLHREGMFNDTLRLIGEHHDGV
SKANFSISRMTQDLAGTYRCYGSVTHSPYQVSAPSDPLDIVIIGLYEKPSLSAQPGPTVLAGENVTLSCSSRSSYDMYHL
SREGEAHERRLPAGPKVNGTFQADFPLGPATHGGTYRCFGSFHDSPYEWSKSSDPLLVSVTGNPSNSWPSPTEPSSKTGN
PRHLHILIGTSVVIILFILLFFLLHRWCSNKKNAAVMDQESAGNRTANSEDSDEQDPQEVTYTQLNHCVFTQRKITRPSQ
RPKTPPTDIIVYTELPNAESRSKVVSCP
```

Nucleotide Sequence Data (1047 bps)

```
ATGTCGCTCTTGTTCGTCAGCATGGCGTGTGTTGGGTTCTTCTTGCTGCAGGGGGCCTGGCCACATGAGGGAGTCCACAG
AAAACCTTCCCTCCTGGCCCACCCAGGTCCCCTGGTGAAATCAGAAGAGACAGTCATCCTGCAATGTTGGTCAGATGTCA
TGTTTGAACACTTCCTTCTGCACAGAGAGGGGATGTTTAACGACACTTTGCGCCTCATTGGAGAACACCATGATGGGGTC
TCCAAGGCCAACTTCTCCATCAGTCGCATGACGCAAGACCTGGCAGGGACCTACAGATGCTACGGTTCTGTTACTCACTC
CCCCTATCAGGTGTCAGCTCCCAGTGACCCTCTGGACATCGTGATCATAGGTCTATATGAGAAACCTTCTCTCTCAGCCC
AGCCGGGCCCCACGGTTCTGGCAGGAGAGAATGTGACCTTGTCCTGCAGCTCCCGGAGCTCCTATGACATGTACCATCTA
TCCAGGGAAGGGGAGGCCCATGAACGTAGGCTCCCTGCAGGGCCCAAGGTCAACGGAACATTCCAGGCTGACTTTCCTCT
GGGCCCTGCCACCCACGGAGGGACCTACAGATGCTTCGGCTCTTTCCATGACTCTCCATACGAGTGGTCAAAGTCAAGTG
ACCCACTGCTTGTTTCTGTCACAGGAAACCCTTCAAATAGTTGGCCTTCACCCACTGAACCAAGCTCCAAAACCGGTAAC
CCCCGACACCTGCACATTCTGATTGGGACCTCAGTGGTCATCATCCTCTTCATCCTCCTCTTCTTTCTCCTTCATCGCTG
GTGCTCCAACAAAAAAATGCTGCGGTAATGGACCAAGAGTCTGCAGGAAACAGAACAGCGAATAGCGAGGACTCTGATG
AACAAGACCCTCAGGAGGTGACATACACACAGTTGAATCACTGCGTTTTCACACAGAGAAAAATCACTCGCCCTTCTCAG
AGGCCCAAGACACCCCAACAGATATCATCGTGTACACGGAACTTCCAAATGCTGAGTCCAGATCCAAAGTTGTCTCCTG
CCCATGA
```

Genomic Sequence Data (14215 bps)

```
TAGGTGTCAAATTCTAGCGAGAACAATTTCCAGGAAGCCGTGTTCCGCTCTTGAGCGAGCACCCACTGGGCCTCATGCAA
GGTAGAAAGAGCCTGCGTACGTCACCCTCCCATGATGTGGTCAACATGTAAACTGCATGGGCAGGCGCCAAATAACATC
CTGTGCGCTGCTGAGCTGAGCTCGGTCGCGGCTGCCTGTCTGCTCCGGCAGCACCATGTCGCTCTTGTTCGTCAGCATGG
```

Figure 1. Allele report. The figure shows part of the report provided for each KIR allele. The report provides cross-references to an SRS entry (KIR00002), the source entries in EMBL-BANK (AY789055-U24076) and to the seminal citations in PubMed. Other information provided includes links to the source material, known ethnic origins and genomic DNA, cDNA and translated protein sequence.

inhibition of NK cell mediated lysis. The output lists the ligands associated with the HLA typing entered. In the case of two digit typings the most common ligand for the allele type is displayed. A link to the full list of alleles matching the type given and their associated ligands is also provided. For two digit typings any exceptions to the list are also listed along with their motif. For example B*15 alleles are predominately Bw6 however there are a number of alleles, B*1513 for example, which contain the Bw4 motif. The output classifies alleles in the exceptions list as 'Rare' if they have a gene frequency of <0.001 and have been seen in less than three unrelated individuals (10). This classification allows users to make a judgement call over whether the exceptions are likely to be seen in their samples.

The typing of KIRs is dependant on up to date lists of alleles and primers, and many typing laboratories have spreadsheets detailing probe hit patterns for different

KIR Alignment Tool

Where discrepancies have arisen between reported sequences and those stored in the database, the original authors have been contacted where possible, and necessary amendments to published sequences have been incorporated into this alignment. Future sequencing may identify errors in this list and the Nomenclature Committee would welcome any evidence that helps to maintain the accuracy of these sequence alignments.

KIR Database Alignment Tool		
Select Gene :	Mamu-KIR2DL4	Help
Select the feature to align :	Nucleotide - CDS	Help
Enter the reference sequence :	Mamu-KIR2DL4*001	Help
Enter any specific sequences required :		Help
Select how you wish to view any mismatches :	Show mismatches between sequences	Help
Select how the alignment will be numbered :	In Codons - nucleotide sequence displayed in codons	Help
Do you want to omit alleles unsequenced for this region :	Show all alleles	Help
Proceed with the alignment :	Align Sequences Now Reset Form	

Figure 2. Alignment interface. The alignment interface provides a user-friendly method of viewing sequence alignments with output options easily selected.

Figure 3. Alignment formats available from IPD. In these alignments a dash (−) indicates identity to the reference sequence and an asterisk (*) denotes an unsequenced base. The first alignment shows the default output for the nucleotide sequence of KIR3DL2. The second alignment shows alleles from six different genes. The pipes indicate exon boundaries and the periods represent the non-expressed pseudoexon 3 in the 2DL genes. The final alignment shows the translated protein sequences of the Rhesus Macaque (*Macaca mulatta*) 3DL1 gene.

alleles. Each time a new database is released it is necessary to update these ever-expanding lists. The new Probe and Primer Search Tool allows uses to enter a list of primer sequences and the tool will search the know coding sequences for these and report any matches in file format suitable for cutting and pasting into existing spreadsheets. The tool is currently limited to coding sequence but as the number of genomic sequences in the database expands then the tool will be modified to search these regions as well. A number of groups involved in developing KIR typing have proposed A Community Standard Reporting Format for KIR Genotyping Data, these guidelines are available from the IPD-KIR website.

The IPD-KIR database is also been expanded to include the KIR sequences from other species, most recently work has begun on including the sequences of KIR alleles found in Rhesus Macaques (*Macaca mulatta*) (11–13). These sequences have been compiled through direct submission and through data-mining existing sequences from the generalist sequence databanks. The first release of the official Mamu-KIR nomenclature will include 107 alleles covering 13 loci. Further loci will be added once haplotype studies have allowed the identification of the genes present. Sequences from other primate species like Crab eating Macaque (*Macaca fascicularis*) have also been submitted to the database, and these sequences will be included at a later date. The non-human KIR sequences will be included into the IPD-KIR section and be accessible using the same tools as the human KIR sequences.

IPD-MHC

The MHC sequences of many different species have been reported (14–17), along with different nomenclature systems used in the naming and identification of new genes and alleles in each species. The sequences of the MHC from a number of different species are highly conserved between species (18). The nomenclature for MHC genes and alleles in species other than humans and mice has historically been overseen either informally by groups generating sequences, or by formal nomenclature committees set up by the International Society for Animal Genetics (ISAG). This work is now overseen by the Comparative MHC Nomenclature Committee (19) and is supported by ISAG and the Veterinary Immunology Committee (VIC) of the International Union of Immunological Societies (IUIS). By bringing the work of different nomenclature committees and the sequences of different species together it is hoped to provide a central resource that will facilitate further research on the MHC of each species and on their comparison.

The first version of the IPD-MHC database involved the work of groups specialising in non-human primates (NHP), canines (DLA) (15,20) and felines (FLA) and incorporated all data previously available in the IMGT/MHC database (21). This version included data from five species of ape, 16 species of new world monkey and 17 species of old world monkey. Since the first version we have been able to add sequences from cattle (BoLA), teleost fish, rats (RT1), sheep (OLA) and swine (SLA) (22–24). Future versions are planned to include chicken, horse, prosimians and pinipeds (seal and sea lion) sequences. As well as adding new species to the database, the individual sections are also updated to include new sequences, depending on the number of sequences to be added updates are performed approximately twice a year.

For each species, there are some differences in the spectrum of data covered but all sections provide the core nomenclature pages and sequence alignments. The nomenclature and alignments follow a similar structure to that of the IPD-KIR section, and the same basic tools are used in both sections. Currently the IPD-MHC sequence alignments are limited to species-specific alignments; however, we are working to allow cross species alignments and the inclusion of human sequences from the IMGT/HLA database (25) for comparative purposes.

The initial release of the database was limited to access plain text alignments and simple HTML tables detailing the nomenclature. We are in the process of migrating each section to use tools similar to IPD-KIR, this means user-defined alignments, online search tools and more detailed allele reports. Currently only the NHP section has been moved to this new version, this was undertaken when the nomenclature used for some species was updated. The NHP section now contains information for over 40 species of primate, with nearly 3000 alleles.

IPD-HPA

Human platelet antigens (HPA) are alloantigens expressed only on platelets, specifically on platelet membrane glycoproteins. These platelet-specific antigens are immunogenic and can result in pathological reactions to transfusion therapy. The HPA nomenclature system was adopted in 1990 (26,27) to overcome problems with the previous nomenclature. Since then more antigens have been described and the molecular basis of many has been resolved. As a result the nomenclature was revised in 2003 (28) and included in the IPD project. The IPD-HPA section contains nomenclature information and additional background material. The different genes in the HPA system have not been sequenced to the same level as some of the other projects and so currently only single nucleotide polymorphisms (SNP) are used to determine alleles. This information is presented in a grid of SNP for each gene The IPD and HPA nomenclature committee hope to expand this to provide full sequence alignments when possible.

IPD-ESTDAB

IPD-ESTDAB is a database of immunologically characterised melanoma cell lines. The database works in conjunction with ESTDAB cell bank (29,30), which is housed in Tübingen, Germany and provides immunologically characterised tumour cells. The IPD-ESTDAB section of the website provides an online search facility for cells stored in this cell bank. This enables investigators to identify cells possessing specific parameters important for studies of immunity, immunogenetics, gene expression, metastasis, response to chemotherapy, and other tumour biological experimentation. The search tool allows for searches based on a single parameter, or clusters of parameters on over 250 different markers for each cell. The detailed reports produced can be then be used to identify cells of interest, which can then be obtained from the cell bank.

DISCUSSION

The IPD project provides a resource for those interested in the study of polymorphic sequences in the immune system. By accommodating related systems in a single database, data can be made available in common formats aiding use and interpretation. As the projects grow and more sections are added, the benefit of having expertly curated sequences from related areas stored in a single location is becoming more apparent. This is particularly true of the IPD-MHC project, where cross-species studies are able to utilise the high quality sequences provided by the different nomenclature committees in a common standardised format, ready for use. The initial release of the IPD Database contained only four sections and a small number of tools; however, as the database has grown and more sections and species have been added, more tools have been added to the website. We plan to use the existing database structures to house data for new sections of the IPD project as they become available. The files will also be made available in different formats to download from the website, ftp server and included into SRS, BLAST and FASTA search engines at the European Bioinformatics Institute (31).

REFERENCES

1. Cochrane,G., Akhtar,R., Bonfield,J., Bower,L., Demiralp,F., Faruque,N., Gibson,R., Hoad,G., Hubbard,T., Hunter,C. *et al.* (2009) Petabyte-scale innovations at the European Nucleotide Archive. *Nucleic Acids Res.*, **37**, D19–D25.
2. Benson,D.A., Karsch-Mizrachi,I., Lipman,D.J., Ostell,J. and Sayers,E.W. (2009) GenBank. *Nucleic Acids Res.*, **37**, D26–D31.
3. Sugawara,H., Ikeo,K., Fukuchi,S., Gojobori,T. and Tateno,Y. (2009) DDBJ dealing with mass data produced by the second generation sequencer. *Nucleic Acids Res.*, **37**, D16–D18.
4. Garcia,C.A., Robinson,J., Guethlein,L.A., Parham,P., Madrigal,J.A. and Marsh,S.G.E. (2003) Human KIR sequences 2003. *Immunogenetics*, **55**, 227–239.
5. Marsh,S.G.E., Parham,P., Dupont,B., Geraghty,D.E., Trowsdale,J., Middleton,D., Vilches,C., Carrington,M., Witt,C., Guethlein,L.A. *et al.* (2003) Killer-cell immunoglobulin-like receptor (KIR) nomenclature report, 2002. *Immunogenetics*, **55**, 220–226.
6. Yun,G., Tolar,J., Yerich,A.K., Marsh,S.G.E., Robinson,J., Noreen,H., Blazar,B.R. and Miller,J.S. (2007) A novel method for KIR-ligand typing by pyrosequencing to predict NK cell alloreactivity. *Clin. Immunol.*, **123**, 272–280.
7. Ruggeri,L., Capanni,M., Casucci,M., Volpi,I., Tosti,A., Perruccio,K., Urbani,E., Negrin,R.S., Martelli,M.F. and Velardi,A. (1999) Role of natural killer cell alloreactivity in HLA-mismatched hematopoietic stem cell transplantation. *Blood*, **94**, 333–339.
8. Khakoo,S.I., Thio,C.L., Martin,M.P., Brooks,C.R., Gao,X., Astemborski,J., Cheng,J., Goedert,J.J., Vlahov,D., Hilgartner,M. *et al.* (2004) HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. *Science*, **305**, 872–874.
9. Gumperz,J.E., Barber,L.D., Valiante,N.M., Percival,L., Phillips,J.H., Lanier,L.L. and Parham,P. (1997) Conserved and variable residues within the Bw4 motif of HLA-B make separable contributions to recognition by the NKB1 killer cell-inhibitory receptor. *J. Immunol.*, **158**, 5237–5241.
10. Cano,P., Klitz,W., Mack,S.J., Maiers,M., Marsh,S.G., Noreen,H., Reed,E.F., Senitzer,D., Setterholm,M., Smith,A. *et al.* (2007) Common and well-documented HLA alleles: report of the Ad-Hoc committee of the american society for histocompatiblity and immunogenetics. *Hum. Immunol.*, **68**, 392–417.
11. Bimber,B.N., Moreland,A.J., Wiseman,R.W., Hughes,A.L. and O'Connor,D.H. (2008) Complete characterization of killer Ig-like receptor (KIR) haplotypes in Mauritian cynomolgus macaques: novel insights into nonhuman primate KIR gene content and organization. *J. Immunol.*, **181**, 6301–6308.
12. Blokhuis,J.H., Doxiadis,G.G. and Bontrop,R.E. (2009) A splice site mutation converts an inhibitory killer cell Ig-like receptor into an activating one. *Mol. Immunol.*, **46**, 640–648.
13. Blokhuis,J.H., van der Wiel,M.K., Doxiadis,G.G. and Bontrop,R.E. (2009) Evidence for balancing selection acting on KIR2DL4 genotypes in rhesus macaques of Indian origin. *Immunogenetics*, **61**, 503–512.
14. Kennedy,L.J., Angles,J.M., Barnes,A., Carter,S.D., Francino,O., Gerlach,J.A., Happ,G.M., Ollier,W.E., Thomson,W. and Wagner,J.L. (2001) Nomenclature for factors of the dog major histocompatibility system (DLA), 2000: Second report of the ISAG DLA Nomenclature Committee. *Tissue Antigens*, **58**, 55–70.
15. Kennedy,L.J., Altet,L., Angles,J.M., Barnes,A., Carter,S.D., Francino,O., Gerlach,J.A., Happ,G.M., Ollier,W.E., Polvi,A. *et al.* (1999) Nomenclature for factors of the dog major histocompatibility system (DLA), 1998. First report of the ISAG DLA Nomenclature Committee. International Society for Animals Genetics. *Tissue Antigens*, **54**, 312–321.
16. Davies,C.J., Joosten,I., Bernoco,D., Arriens,M.A., Bester,J., Ceriotti,G., Ellis,S., Hensen,E.J., Hines,H.C., Horin,P. *et al.* (1994) Polymorphism of bovine MHC class I genes. Joint Report of the Fifth International Bovine Lymphocyte Antigen (BoLA) Workshop, Interlaken, Switzerland, 1 August 1992. *Eur. J. Immunogenet.*, **21**, 239–258.
17. Klein,J., Bontrop,R.E., Dawkins,R.L., Erlich,H.A., Gyllensten,U.B., Heise,E.R., Jones,P.P., Parham,P., Wakeland,E.K. and Watkins,D.I. (1990) Nomenclature for the major histocompatibility complexes of different species: a proposal. *Immunogenetics*, **31**, 217–219.
18. Parham,P. (1999) Virtual reality in the MHC. *Immunol. Rev.*, **167**, 5–15.
19. Ellis,S.A., Bontrop,R.E., Antczak,D.F., Ballingall,K., Davies,C.J., Kaufman,J., Kennedy,L.J., Robinson,J., Smith,D.M., Stear,M.J. *et al.* (2006) ISAG/IUIS-VIC Comparative MHC Nomenclature Committee report, 2005. *Immunogenetics*, **57**, 953–958.
20. Kennedy,L.J., Angles,J.M., Barnes,A., Carter,S.D., Francino,O., Gerlach,J.A., Happ,G.M., Ollier,W.E., Thomson,W. and Wagner,J.L. (2001) Nomenclature for factors of the dog major histocompatibility system (DLA), 2000: second report of the ISAG DLA Nomenclature Committee. *Anim. Genet.*, **32**, 193–199.
21. Robinson,J., Waller,M.J., Parham,P., de Groot,N., Bontrop,R., Kennedy,L.J., Stoehr,P. and Marsh,S.G.E. (2003) IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. *Nucleic Acids Res.*, **31**, 311–314.
22. Ho,C.S., Lunney,J.K., Ando,A., Rogel-Gaillard,C., Lee,J.H., Schook,L.B. and Smith,D.M. (2009) Nomenclature for factors of the SLA system, update 2008. *Tissue Antigens*, **73**, 307–315.
23. Smith,D.M., Lunney,J.K., Ho,C.S., Martens,G.W., Ando,A., Lee,J.H., Schook,L., Renard,C. and Chardon,P. (2005) Nomenclature for factors of the swine leukocyte antigen class II system, 2005. *Tissue Antigens*, **66**, 623–639.
24. Smith,D.M., Lunney,J.K., Martens,G.W., Ando,A., Lee,J.H., Ho,C.S., Schook,L., Renard,C. and Chardon,P. (2005) Nomenclature for factors of the SLA class-I system, 2004. *Tissue Antigens*, **65**, 136–149.
25. Robinson,J., Malik,A., Parham,P., Bodmer,J.G. and Marsh,S.G.E. (2000) IMGT/HLA database—a sequence database for the human major histocompatibility complex. *Tissue Antigens*, **55**, 280–287.
26. von dem Borne,A.E. and Decary,F. (1990) ICSH/ISBT Working Party on platelet serology. Nomenclature of platelet-specific antigens. *Vox Sang.*, **58**, 176.

27. von dem Borne,A.E. and Decary,F. (1990) Nomenclature of platelet-specific antigens. *Hum. Immunol.*, **29**, 1–2.
28. Metcalfe,P., Watkins,N.A., Ouwehand,W.H., Kaplan,C., Newman,P., Kekomaki,R., De Haas,M., Aster,R., Shibata,Y., Smith,J. *et al.* (2003) Nomenclature of human platelet antigens. *Vox Sang*, **85**, 240–245.
29. Pawelec,G. and Marsh,S.G.E. (2006) ESTDAB: a collection of immunologically characterised melanoma cell lines and searchable databank. *Cancer Immunol. Immunother.*, **55**, 623–627.
30. Robinson,J., Roberts,C.h., Dodi,A.I., Pawelec,G. and Marsh,S.G.E. (2009) The European Searchable Tumour Line Database (ESTDAB). *Cancer Immunol. Immunother.*, **58**, 1501–1506.
31. Harte,N., Silventoinen,V., Quevillon,E., Robinson,S., Kallio,K., Fustero,X., Patel,P., Jokinen,P. and Lopez,R. (2004) Public web-based services from the European Bioinformatics Institute. *Nucleic Acids Res.*, **32**, W3–W9.

Appendix 1:
Access and contact

IPD homepage:	http://www.ebi.ac.uk/ipd/
IPD-KIR homepage:	http://www.ebi.ac.uk/ipd/kir/
IPD-MHC homepage:	http://www.ebi.ac.uk/ipd/mhc/
IPD-HPA homepage:	http://www.ebi.ac.uk/ipd/hpa/
IPD-ESTDAB homepage:	http://www.ebi.ac.uk/ipd/estdab/
Contact:	ipd@ebi.ac.uk

If you are interested in contributing to the project, there are specific guidelines for the inclusion of new sections, and interested parties should contact Prof. SGE Marsh, steven.marsh@ucl.ac.uk for further information.