Immunogenetics (2018) 70:625–632
https://doi.org/10.1007/s00251-018-1073-3

ORIGINAL ARTICLE

Comparative MHC nomenclature: report from the ISAG/IUIS-VIC committee 2018

Keith T. Ballingall[1] · Ronald E. Bontrop[2] · Shirley A. Ellis[3] · Unni Grimholt[4] · John A. Hammond[3] · Chak-Sum Ho[5] · Jim Kaufman[6] · Lorna J. Kennedy[7] · Giuseppe Maccari[3,8] · Donald Miller[9] · James Robinson[8,10] · Steven G. E. Marsh[8,10]

Received: 25 May 2018 / Accepted: 13 July 2018 / Published online: 24 July 2018

Abstract

Significant progress has been made over the last decade in defining major histocompatibility complex (MHC) diversity at the nucleotide, allele, haplotype, diplotype, and population levels in many non-human species. Much of this progress has been driven by the increased availability and reduced costs associated with nucleotide sequencing technologies. This report provides an update on the activities of the comparative MHC nomenclature committee which is a standing committee of both the International Society for Animal Genetics (ISAG) and the International Union of Immunological Societies (IUIS) where it operates under the umbrella of the Veterinary Immunology Committee (VIC). A previous report from this committee in 2006 defined the role of the committee in providing guidance in the development of a standardized nomenclature for genes and alleles at MHC loci in non-human species. It described the establishment of the Immuno Polymorphism Database, IPD-MHC, which continues to provide public access to high quality MHC sequence data across a range of species. In this report, guidelines for the continued development of a universal MHC nomenclature framework are described, summarizing the continued development of each species section within the IPD-MHC project.

Keywords Comparative MHC · Nomenclature · IPD-MHC

Background

The Comparative Major Histocompatibility Complex (MHC) Nomenclature Committee met for the first time at the Institute for Animal Health, Compton, UK, in January 2003, and its first workshop was held in 2004 at the International Veterinary Immunology Symposium (IVIS) in Quebec. Comparative MHC workshops have been held, at most, at the International Society for Animal Genetics (ISAG) and IVIS symposia since. The committee seeks to provide a guiding framework for allelic nomenclature across all non-human species. The composition of the current committee is shown in Table 1. As detailed in the 2006 report (Ellis et al. 2006), the requirement at the time for a standardized nomenclature framework to support the increase in allelic sequence data from many different species was clear. Much of the approach taken by the Comparative MHC Committee mirrors aspects of the work of the WHO Nomenclature Committee for Factors of the HLA System. The human MHC known as HLA, with its central role in transplant tolerance, has by clinical necessity required an accurate system of allelic nomenclature where unambiguous assignment of an official name to a sequence

Shirley A. Ellis is no longer at The Pirbright Institute

✉ Keith T. Ballingall
keith.ballingall@moredun.ac.uk

[1] Moredun Research Institute, Midlothian, UK and Chair of the Comparative MHC Nomenclature Committee, Edinburgh, Scotland, UK

[2] Biomedical Primate Research Centre, Rijswijk, The Netherlands

[3] The Pirbright Institute, Pirbright, Surrey, UK

[4] Norwegian Veterinary Institute, Oslo, Norway

[5] Gift of Life Michigan, Ann Arbor, MI, USA

[6] University of Cambridge, Cambridge, UK

[7] Centre for Integrated Genomic Medical Research, Manchester, UK

[8] Anthony Nolan Research Institute, London, UK

[9] Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA

[10] UCL Cancer Institute, Royal Free Campus, London, UK

Table 1 The composition of the ISAG/IUIS comparative MHC nomenclature committee

Composition of the Comparative MHC Nomenclature Committee, 2018		
Name	Taxa of interest	Affiliation
Keith Ballingall (Chair)	Sheep/ruminants	Moredun Research Institute, UK
Mike Stear	Sheep/ruminants	La Trobe University, Australia
John Hammond	Cattle/ruminants	The Pirbright Institute, UK
Shirley Ellis (Past Chair)	Cattle/ruminants	The Pirbright Institute, UK[#]
Chak-Sum Ho	Swine	Gift of Life, Michigan, USA
Ronald Bontrop	Non-human primates	Biomedical Primate Research Centre, The Netherlands
Unni Grimholt	Salmonid fish	Norwegian Veterinary Institute
Jim Kaufman	Chicken	University of Cambridge, UK
Donald Miller	Horses	Baker Institute for Animal Health, Cornell University, NY, USA
Lorna Kennedy	Dogs and cats	Centre for Integrated Genomic Medical Research, UK
Lutz Walter	Rat	German Primate Centre
Tomas Bergström	Horses	Swedish University of Agricultural Sciences
Chris Davies (Past Chair)	Ruminants	College of Agriculture and Applied Sciences, Utah State University, USA
Shin-nosuke Takeshima	Ruminants	Viral Infectious Disease Unit, RIKEN, Japan
Bertrand Bed'Hom	Chickens	INRA, Jouy en Josas, France
*Steven Marsh	Humans IPD-IMGT/HLA and IPD-MHC Database	Anthony Nolan Research Institute and UCL Cancer Institute, UK
*James Robinson	IPD-MHC Bioinformatics	Anthony Nolan Research Institute and UCL Cancer Institute, UK
*Giuseppe Maccari	IPD-MHC Bioinformatics	The Pirbright Institute and Anthony Nolan Research Institute, UK

[#] No longer at this Institute. *Co-opted members

occurs following a number of quality assessments (WHO Nomenclature Committee 1968; Marsh et al. 2010). HLA allelic sequence data are maintained in a public database (IPD-IMGT/HLA, Robinson et al. 2015). Similarly, the Comparative MHC Committee works closely with the developers and curators of the IPD-MHC Database to provide high-quality allelic sequence data with official names to the research community (Robinson et al. 2005). Currently, the IPD-MHC database includes over 8000 sequences from non-human primates, canids, bovids, equids, salmonids, rattus, ovids, suids, and gallus. The importance of the IPD-MHC Database to the research community has been recognized by receipt of significant funding from the UK's Biotechnology and Biological Sciences Research Council (BBSRC) to support its upgrade and future development (Maccari et al. 2017).

The purpose of this report

In 1990, Jan Klein proposed a system for MHC nomenclature in which MHC genes from any species could be prefixed by a four-letter abbreviation composed of the first two letters of the genus and the first two letters of the species scientific name (Klein et al. 1990). This system of nomenclature is now widely accepted with only a few historical exceptions. The purpose of this report and its partner paper (Maccari et al. 2018) is to provide an update on the committee's work in partnership with the developers of the IPD-MHC Database to elaborate a universal nomenclature framework which is flexible enough to deal with the complexities of MHC diversity in gnathostomes, or jawed vertebrates. To achieve this, a set of rules have been developed that allows a standardized approach to MHC nomenclature and ensures that the quality standards of data held on the IPD-MHC Databases are maintained for the benefit of the research community.

Overview of the nomenclature system from 2006

Prior to 2017, most allelic designation for species included in the IPD-MHC database followed the pre-2010 HLA system of nomenclature. This included details of the species from which the sequence was derived by incorporating the first two letters of the genus and species names. For example, sequences derived from domestic sheep, genus *Ovis*, species *O. aries*, were prefixed *Ovar* while sequences derived from the Atlantic salmon, genus *Salmo*, species *S. salar*, were prefixed *Sasa*.

The next set of letters, linked by a hyphen, described the locus to which the sequence belongs. Where homologs or orthologues of the HLA loci are identifiable such as the MHC class II loci of most mammalian species, the same locus

nomenclature was used. This included the MHC class II *DRA*, *DRB*, *DQA*, *DQB* loci, and the *DP* locus in some non-human primates (NHP). As many of these loci are duplicated, a number was assigned by individual species nomenclature committees to define each of the duplicated genes. For example, the highly transcribed and polymorphic MHC class II *DRB* locus in cattle is termed *DRB3* while in sheep it is *DRB1*. This reflects the chronological order of their initial identification. Apart from some NHP, orthologues of the HLA class I loci were not readily identifiable in other species and individual species nomenclature committees have thus developed their own MHC class I nomenclature systems.

To separate the genus/species and locus designation from the allele name, an asterisk (*) is incorporated. This is followed by a number describing the allelic group to which the sequence belongs. While each species-specific nomenclature committee has defined its own rules, in general members of an allelic group are defined as having four or fewer amino acid substitutions in the region encoded by the first extracellular domain (alpha 1 or beta 1 domains) of mature MHC class II proteins and the alpha 1 and alpha 2 domains of mature MHC class I proteins. In practice, the predicted amino acid sequence derived from the second exon of MHC class II alleles and the second and third exons of MHC class I alleles has been used to define the allelic group.

Once the allelic group has been assigned, the next number defines the order within the allelic group. This number is based on the order of submission to the IPD-MHC Database. Two additional numbers may be used to define synonymous substitutions and to define diversity within introns or 5-prime and 3-prime untranslated regions respectively.

The need to update the nomenclature system

The nomenclature system described above worked well through the commitment of species nomenclature committees and those at the European Bioinformatics Institute - EMBL Outstation (EBI) and the Anthony Nolan Research Institute who maintained and updated the IPD-MHC Databases. However, this approach has become unsustainable due to the widespread availability and cost effectiveness of Sanger sequencing and the development of next-generation sequence technologies over the last 12 years. This necessitated the development of a more standardized approach to providing MHC allelic nomenclature in non-human species which remains compatible with the new IPD-MHC Database bioinformatics infrastructure recently described by Maccari et al. (2017). Here, we present the committee's recommendations for the further development of MHC nomenclature in non-human species, which have been developed through discussions at successive workshops and IPD-MHC Database steering group meetings over the last 5 years. For the development of nomenclature in species not currently incorporated into the IPD-MHC Database, the committee recommends that the following guidelines are followed.

Incorporating new species sections into the IPD-MHC Database

The committee, with support from the IPD-MHC Database provides the following recommendations for inclusion of new species sections into the database:

- An individual curator with experience of working on the new species should be identified to take the lead. The curator will be encouraged to publish their intention to develop the new species section on the IPD-MHC Database as a brief note to inform the research community, to allow formation of a species nomenclature committee, and to encourage sequence submission
- This species committee will then develop the nomenclature rules with support from the nomenclature committee and the IPD-MHC Database
- The sequence size and quality should conform to the standards specified below
- Ideally, the reference sequence for each locus should be a full-length transcript
- The committee suggests that a single polymorphic locus is used to start the process with additional loci added as required

Recommendations for changes in nomenclature

Shared genus/species prefixes

As detailed in the accompanying paper (Maccari et al. 2018), additional flexibility is required to deal with sequence data from potentially thousands of species. We propose to maintain the four-letter genus/species designation; however, in cases where a prefix is shared, the third or subsequent letters of the species name will be used until a unique prefix is generated. The IPD-MHC Database acts as the official repository of already assigned names. When a request for the addition of a particular organism is received, a unique prefix is generated and stored in the database.

Standardized system of MHC class I nomenclature

With the exception of some non-human primates, orthologues of the *HLA-A*, *HLA-B* and *HLA-C* loci cannot easily be defined in other species due to high levels of gene duplication,

gene loss, inter- and intra-locus recombination, and functional diversification of MHC class I loci over time. Therefore, bespoke naming systems are required for each species. The committee recommends that a numerical tag be used for each MHC class I locus to avoid confusion with the HLA loci. Where MHC class I sequences cannot be assigned to individual loci, the sequences should be assigned to a single allelic series under the prefix N, for example *class I-N*01:01:01*. When haplotype data become available that allow sequences to be assigned to individual loci, the N may be replaced by a number, as recently recommended for cattle class I sequences (Hammond et al. 2012). Where non-classical class I genes are identified, these should be designated NC, for example *class I-NC*01:01*.

A standardized system of MHC class II nomenclature across mammalian MHC

Where orthologues or homologs of the HLA class II *DP*, *DQ*, and *DR* loci are identifiable in other species, we recommend that this nomenclature continues to be used, rather than a bespoke system such as those which, for historical reasons were developed for *Mus musculus*, (reviewed in Klein 1986) and *Rattus norvegicus* (Hurt et al. 2004).

Recommendations for non-mammalian species

Of the estimated 22,000 species of marine fish, nomenclature systems have been developed only for Atlantic salmon (*Salmo salar*) and rainbow trout (*Oncorhynchus mykiss*, Grimholt et al. 2000, Grimholt et al. 2015). As sequence data become available from many more species, shared patterns of diversity may be identified in related taxonomic groups allowing some standardization of nomenclature. For avian species, a bespoke nomenclature system has been developed only for the chicken (*Gallus gallus*, Miller et al. 2004). The committee recommends that the nomenclature guidelines initially proposed by Klein et al. in 1990 are followed and endorses the refinements described here and in Maccari et al. (2018).

The use of the colon

As detailed in the accompanying paper (Maccari et al. 2018), the committee endorses the use of the colon (:) as a field separator in allele names for use in all non-human MHC allelic nomenclature.

A standard for MHC class I and class II allelic nomenclature across species

The committee recommends that the approach described in "Overview of the nomenclature system from 2006" to assign alleles to allelic groups is standardized across taxa. In keeping with HLA nomenclature, the current approach is based on amino acid diversity in the combined α1 and α2 domains of the mature class I protein and the α1 or β1 domains of mature class II protein. Generally, an allelic group includes sequences with four or fewer amino acid substitutions across these regions.

Quality standards for submission of sequences to the IPD-MHC database

Sequences should conform to a minimum length standard

The sequence length requirements for inclusion in the IPD-MHC Database remain the prerogative of individual species nomenclature committees. Data quality remains the highest priority over sequence length. The aspiration of the committee is the complete genomic sequence of a number of haplotypes for each species. This would provide a reference sequence for each locus, identify locus-specific features associated with gene duplication, define the basis of haplotype diversity and increase the accuracy of the assignment of alleles to specific loci. With rapid developments in sequencing technologies, larger fragments of class I and class II genes may now be sequenced from genomic or cDNA templates. For class I and II loci, submission of small fragments of individual exons are of less value and the committee continues to recommend the submission of full-length transcripts. However, for genes such as class II *DRB* in sheep, cattle, and dogs where hundreds of alleles have already been described, submission of a complete second exon sequence is sufficient to receive an official name.

Sequences should conform to a minimum quality standard

As described in the previous report (Ellis et al. 2006), most sequences are derived through PCR amplification, direct sequence of the PCR product followed by validation by cloning and bidirectional sequencing. Data quality is maintained through analysis of three or more independent clones and through identification of the allele in more than one animal. On submission of the sequence to the IPD-MHC Database, sequences are checked by an expert curator for unique SNPs, indels, and reading frame and for compatibility with other sequences. If problems are identified, additional validation data may be requested prior to assignment of an official name. The committee continues to recommend these quality control measures which ensure that only high-quality, validated alleles are assigned official names. Such quality standards are also recommended for sequence derived through next-generation sequencing technologies.

To accept data into the IPD-MHC Database, the committee recommends that the following quality measures are followed irrespective of the sequencing technology employed:

- The sequence must include a complete second exon sequence for MHC class II loci or complete second and third exons for MHC class I loci
- The sequence must be derived in both directions, and if generated from multiple fragments, the tiling technology employed should be capable of accurately phasing sequence from heterozygous individuals
- The number of animals in which an identical sequence was identified must be included in the submission. Unique sequences generated from a single animal must be validated using a different method
- The number and quality of the sequences generated for the allele in question must be stated. Alleles that differ by single nucleotides must be validated
- The measures employed to eliminate chimeric alleles must be detailed
- The sequencing platform and chemistry employed must be stated
- DNA or RNA quality measures employed must be stated

Nomenclature summaries for individual species

Salmonid Fish, curated by Unni Grimholt, Norwegian Veterinary Institute, Oslo, Norway

In bony fish, the MHC class Ia and II genes are not located in a complex, but are found on different chromosomes (Bingulac-Popovic et al. 1997). This unique feature sets the bony fish apart from other vertebrates and may have occurred in the split between Spotted gar and teleosts as both MHC class I and class II genes seem linked in Spotted gar (Grimholt et al. 2015). As in other vertebrates, both classical and non-classical MHC class I and class II genes have been identified in teleosts (reviewed in Grimholt 2016). The classical MHC class I genes all belong to the U lineage while the classical MHC class II genes all belong to what is now defined as the A lineage (Grimholt 2016). Although the function of the classical genes appears as expected, the function of the non-classical genes remains unresolved. A surprising feature of teleost MHC class I is the sharing of alpha 1 domain sequences between distantly related species such as the cyprinid zebrafish and the neoteleosts medaka (Grimholt et al. 2015, Nonaka et al. 2011). This unique sequence feature has not yet been implemented in the nomenclature, but should be kept in mind when analyzing new teleost MHC class I sequences.

There are currently only two teleost species with registered MHC class I and II alleles in the IPD-MHC Database, namely those from rainbow trout (*Oncorhynchus mykiss*) and Atlantic salmon *(Salmo salar)*. In both, the classical MHC class I and class II *A* and *B* genes exist as single copies. Nomenclature used for rainbow trout is Onmy-*DAA* and Onmy-*DAB* for the classical MHC class II *A* and *B* loci while the classical MHC class I locus is designated Onmy-*UBA*. For Atlantic salmon, the nomenclature is similar apart from each gene being preceded by Sasa-*DAA*, Sasa-*DAB*, and Sasa-*UBA*. Information on the location and organization of the class I and class II genes has been described in Atlantic salmon (Grimholt 2016) while the genomic organization MHC class II genes in rainbow trout is currently unknown; the location of Onmy-*UBA* has been described (Shiina et al. 2005). The number of alleles per species in the IPD-MHC Database ranges from 3 to 48.

Gallus gallus, curated by Jim Kaufman and Hassnae Afrache, University of Cambridge, Cambridge, UK

Originally, serology was used to define 24 B haplotypes in egg-laying chickens; alleles within these haplotypes were given names based on the haplotype. Eventually, a nomenclature for the two class I genes, BF1 and BF2, and two class II B genes, BLB1 and BLB, was adopted. However, the existence of recombinant haplotypes and new alleles from other populations including meat-type (broiler) chickens (Hosomichi et al. 2008; Lima-Rosa et al. 2004; Livant et al. 2004; Simonsen et al. 1980; Simonsen et al. 1982; Worley et al. 2008) led to the initial concept of using stable haplotypes to name the alleles becoming cumbersome and misleading.

Therefore, a system to name each allele and then build haplotypes from these alleles has been adopted for the IPD-MHC Database, in which the first field of digits identifies a series of closely related alleles, the second field of digits identifies the exact allele sequence at the protein level, and the third field of digits identifies the exact allele sequence based on the nucleotide sequence. In order to maintain as much of the information from the old system as possible, the original standard haplotypes were assessed in order. As an example, the minor class I gene from the B4 haplotype is now named BF1*004:01, the identical gene found in the B13 haplotype is given the same name, while the closely related gene found in the B21 haplotype is called BF1*004:02. A similar procedure was used for the BF2, BLB1, and BLB2 loci.

Thus far, 20 BLB1, 20 BLB2, 20 BF1, and 33 BF2 alleles are considered to be validated and have been assigned names. There are many other sequences in the public databases that are only present once and increasing amounts of data are becoming available from whole genome sequencing efforts. Understanding how to incorporate these data as well as other polymorphic loci in the chicken and other avian MHC will be future steps.

Non-human primates, curated by Ronald Bontrop, Biomedical Primate Research Centre, Rijswijk, The Netherlands

At present, the non-human primate (NHP) section of the IPD-MHC Database comprises the great apes and Old and New World monkeys and covers 55 species (Maccari et al. 2017). The NHP section contains annotated sequence information of more than 7000 alleles and the rules concerning nomenclature, submission, etc., have been updated recently (de Groot et al. 2012).

Equids, curated by Donald miller, Cornell University, Ithaca, New York, USA.

The equine section of the IPD-MHC Database contains information on the ELA complex of the domestic horse, *Equus caballus*, and other member species of the genus *Equus*. There are currently 57 sequences in the database. All of the sequences in the database were obtained from either Thoroughbred or Standardbred horses. The physical genomic structure of the ELA region was determined using information from the horse reference DNA sequence (Wade et al. 2009). Twilight, the DNA donor horse for the reference sequence, is homozygous for the ELA-A3 haplotype (Tallmadge et al. 2005). Most of the polymorphic MHC class I (Tallmadge et al. 2010) and class II sequences (Miller et al. 2016) in the IPD-MHC Database were identified through sequencing of cDNA clones from horses homozygous for other MHC haplotypes.

A total of 18 sequences from MHC class I loci have been deposited, and 16 of these have been assigned to loci. The remaining two sequences have not been assigned to loci and are prefixed with "N" for "not assigned." For the MHC class II region, 49 sequences have been deposited, 13 class II A alleles and 26 class II B alleles, and all have been assigned to loci.

Nomenclature for the horse sequences follows the HLA system, and sequences in the database are full or nearly full length. In the future, sequence from other equid species including donkey and zebra will be included. The Equine Nomenclature Committee decided to endorse the recommendations of the Comparative MHC Nomenclature Committee and re-designate each allele by adopting the nomenclature which includes colons as separators, for example, allele *Eqca-DRB1*00101* will become *Eqca-DRB1*001:01.*

Bovids, curated by John A. Hammond and Shirley A. Ellis[#], The Pirbright Institute, Surrey, UK.[#]No longer at this Institute.

Since the previous report was published in 2006, the BoLA section has more than doubled in size and now contains almost 500 allele sequences from four distinct groups, cattle (*Bos taurus* and *Bos indicus*), gayal (*Bos frontalis*), yak (*Bos grunniens*), and water buffalo (*Bubalus bubalis*). The previous species prefix before each allele from *Bos taurus* (Bota) and *Bos indicus* (Boin) have been removed and amalgamated into one prefix: BoLA. Several alleles were shared between these two groups and it is sometimes difficult to determine the precise genetic background of study material as these sub-species are often interbred. Therefore, *Bos indicus* and *Bos taurus* are now treated as breeds which more accurately reflect the multiple domestication processes and subsequent cross-breeding that have generated the diversity of domesticated cattle seen worldwide.

The other major nomenclature change has been to remove the prefix "N" (to indicate "not assigned") from the MHC class I genes. As in other species where there is gene presence/absence polymorphism, it has always been problematic to assign allele sequences to an individual locus, or even define individual genes. However, the greater amount of sequence data in combination with multiple other lines of evidence all indicated six loci. Therefore, a robust classification, based on phylogenetic analysis of the cDNA sequence from the alpha 3 domain to the stop codon, which fits with these six genes was adopted in 2012 (Hammond et al. 2012). All MHC class I alleles are still named in a single series to avoid confusion, and all submissions post this revision have fallen into one of the six allele groups.

More minor changes include the standardization of nomenclature between species. BoLA has adopted the HLA naming convention to simplify the overall system and allow multi-species comparison within the new IPD-MHC Database framework. This is essential to allow standardized submission and analysis tools for all species groups. As more sequence data is submitted from next-generation sequencing projects (both RNAseq and whole genome), it is hoped that these changes will allow the same standards to be adopted for BoLA curation with significantly improved scalability.

Suids, curated by Chak-Sum Ho, Gift of Life Michigan, Ann Arbor, Michigan, USA

The swine leucocyte antigen (SLA) system is among the most well characterized MHC systems in non-human animal species. Since the publication of the initial nomenclature reports in 2005 (Smith et al. 2005a, b) and subsequently an update in 2009 (Ho et al. 2009), the swine section of the IPD-MHC Database has grown rapidly in size, which currently contains 259 sequences of the class I loci (*SLA-1*, *SLA-2*, *SLA-3*, *SLA-4*, *SLA-5*, *SLA-6*, *SLA-7*, *SLA-8*, *SLA-9*, *SLA-11*, *SLA-12*) and 227 sequences of the class II loci (*DMA*, *DMB*, *DOA*, *DOB1*, *DOB2*, *DQA1*, *DQB1*, *DQB2*, *DRB1*, *DRB2*, *DRB3*, *DRB4*, *DRB5*, *DYB*). There are also 61 class I (*SLA-1-3-2*) and 49 class II (*DRB1-DQB1*) haplotypes designated at the allele level resolution.

The SLA Nomenclature Committee last met at the 35th ISAG Conference in Salt Lake City, UT, USA, and made some major revisions to the allele naming system. The Committee decided to endorse the recommendations of the Comparative MHC Nomenclature Committee and retired the provisional alphanumerical naming system and re-designated each allele an official number with the use of colons as field separators, for example, allele *SLA-1*01rh28* will now become *SLA-1*01:03*. In addition, The Committee decided that Bayesian phylogeny will remain as the primary approach for assigning allelic sequences of *SLA-1*, *SLA-2*, *SLA-3*, *DRA*, *DRB1*, *DQA*, and *DQB1* into allele groups with similar sequence motifs (i.e., the first field, e.g., *SLA-1*01*), while alleles of the other loci are designated sequentially in the order of description (e.g., *SLA-6*01:01*, *02:01*, *03:01*).

Ovids, curated by Keith Ballingall, Moredun Research Institute, Penicuik Midlothian, UK

The ovine section of the IPD-MHC Database now contains 135 sequences representing both MHC class I and class II alleles from a wide range of domestic sheep breeds (*Ovis aries*). Twenty-six full-length MHC class I sequences are named in a single allelic series as assigning alleles to individual loci remains difficult. However, *Ovar-class I-N*01:01* to *N*21:01* are thought to represent alleles at a minimum of three classical class I loci while *Ovar-N*50* to *Ovar-N*53 are likely to represent alleles at a minimum of two non-classical class I loci (Miltiadou et al. 2005).

The sequences of 106 class II *DRB1* alleles, nine of which are full length, within 26 allelic groups and three *DRA* alleles, are currently held in the IPD-MHC Database. The *DQA* and *DQB* loci are duplicated in sheep and both sets of loci are polymorphic and appear to be functional (Ballingall et al. 2015, 2018). Full-length sequences at each of the *DQ* loci will be deposited in the IPD-MHC Database in the near future to begin the development of an ovine DQ nomenclature system.

The Ovine Nomenclature Committee has also endorsed the recommendations of the Comparative MHC Nomenclature Committee and re-designated each allele by including colons as separators, for example, allele *Ovar-DRB1*0101* will now become *Ovar-DRB1*01:01*.

Canids, curated by Lorna Kennedy, Centre for Integrated Genomic Medical Research, Manchester, UK

The Canidae section of the IPD-MHC Database is in the process of being updated with additional allelic sequences for each of the MHC class I and MHC class II loci. In accordance with the recommendations of the Comparative MHC Committee, the allelic nomenclature has been updated to include the use of the colon.

Three MHC class I loci have been defined in the canids, *DLA-88*, *DLA-12* and *DLA-64*. Forty-eight *DLA-88* alleles are currently held on the IPD-MHC Database. MHC class II alleles and haplotypes and been defined in over 300 breeds of dogs, revealing considerable interbred diversity. Some alleles and haplotypes are common to many breeds, while others have a more restricted distribution. To date, 211 *DLA-DRB1*, 38 *DLA-DQA1*, and 107 *DLA-DQB1* alleles have been assigned official names. These alleles are clustered in at least 270 three-locus haplotypes: *DLA-DRB1*, *DQA1*, and *DQB1*. Individual breeds generally have between two and ten haplotypes, with one or two occurring at high frequency. The most frequent haplotypes are *DLA-DRB1*006:01*, *DLA-DQA1*005:01:1*, *DLA-DQB1*007:01*, which has been found in over 70% of dog breeds.

There is clear allele sharing between canids as the most common haplotype in the domestic dog is also found in gray wolves and coyotes. Such sharing of alleles between species complicates canid allelic nomenclature. Currently, all canine alleles are named in a single series for each locus; however, an alternative system of nomenclature which includes the species prefix is proposed in the accompanying paper (Maccari et al. 2018).

Felids, curated by Lorna Kennedy, Centre for Integrated Genomic Medical Research, Manchester, UK

The data to support the feline MHC section on the IPD website are currently under revision. Once completed and a nomenclature system agreed upon, it is hoped that the site will be available as soon as possible.

Acknowledgements The support of all those who provide their time to assign official names to MHC alleles across many different species is acknowledged. The Comparative MHC Nomenclature Committee would also like to acknowledge the Biotechnology and Biological Sciences Research Council (BBSRC) for supporting the maintenance and development of the IPD-MHC through grant BB/M011488/1 and IUIS-VIC for supporting the comparative MHC workshops at ISAG and IVIS meetings.

Funding information KB receives funding from the Scottish Government's strategic research program and the European Union's Horizon 2020 research and innovation program under grant agreement No. 731014. JAH and GM were also supported by the BBSRC through projects BBS/E/I/00001710, BBS/E/I/00007030, BBS/E/I/00007038, and BBS/E/I/00007039.

References

Ballingall KT, Steele P, Lantier I, Cotelli M, Todd H, Lopez G, Martin E, Lantier F (2015) An ancient inter-locus recombination increases class II MHC DQA diversity in sheep and other bovidae. Anim Genet 46:333–336

Ballingall KT, Lantier I, Todd H, Lantier F, Rocchi M (2018) Structural and functional diversity arising from intra and inter haplotype combinations of duplicated *DQA* and *B* loci within the Ovine MHC. Immunogenetics 70:257–269

Bingulac-Popovic J, Figueroa F, Sato A, Talbot WS, Johnson SL, Gates M, Postlethwait JH, Klein J (1997) Mapping of mhc class I and class II regions to different linkage groups in the zebrafish, Danio rerio. Immunogenetics 46:129–134

de Groot NG, Otting N, Robinson J, Blancher A, Lafont BAP, Marsh SGE, O'Connor DH, Shiina T, Walter L, Watkins DI, Bontrop RE (2012) Nomenclature report on the major histocompatibility complex genes and alleles of Great Ape, Old and New World monkey species. Immunogenetics 64:615–631

Ellis SA, Bontrop RE, Antczak DF, Ballingall K, Davies CJ, Kaufman J, Kennedy LJ, Robinson J, Smith DM, Stear MJ, Stet RJM, Waller MJ, Walter L, Marsh SGE (2006) ISAG/IUIS-VIC comparative MHC nomenclature committee report, 2005. Immunogenetics 57: 953–958

Grimholt U (2016) MHC and evolution in teleosts. Biology 5:1. https://doi.org/10.3390/biology5010006

Grimholt U, Getahun A, Hermsen T, Stet RJ (2000) The major histocompatibility class II alpha chain in salmonid fishes. Dev Comp Immunol 24:751–763

Grimholt U, Tsukamoto K, Azuma T, Leong J, Koop BF, Dijkstra JM (2015) A comprehensive analysis of teleost MHC class I sequences. BMC Evol Biol 15:32

Hammond JA, Marsh SG, Robinson J, Davies CJ, Stear MJ, Ellis SA (2012) Cattle MHC nomenclature: is it possible to assign sequences to discrete class I genes? Immunogenetics 64:475–480

Ho CS, Lunney JK, Ando A, Rogel-Gaillard C, Lee JH, Schook LB, Smith DM (2009) Nomenclature for factors of the SLA system, update 2008. Tissue Antigens 73:307–315

Hosomichi K, Miller MM, Goto RM, Wang Y, Suzuki S, Kulski JK, Nishibori M, Inoko H, Hanzawa K, Shiina T (2008) Contribution of mutation, recombination, and gene conversion to chicken MHC-B haplotype diversity. J Immunol 181:3393–3399

Hurt P, Walter L, Sudbrak R, Klages S, Müller I, Shiina T, Inoko H, Lehrach H, Günther E, Reinhardt R, Himmelbauer H (2004) The genomic sequence and comparative analysis of the rat major histocompatibility complex. Genome Res 14:631–639

Klein J (1986) Natural history of the major histocompatibility complex. Published by J Wiley and Sons, New York

Klein J, Bontrop RE, Dawkins RL, Erlich HA, Gyllensten UB, Heise ER, Jones PP, Parham P, Wakeland EK, Watkins DI (1990) Nomenclature for the major histocompatibility complexes of different species: a proposal. Immunogenetics 31:217–219

Lima-Rosa CA, Canal CW, Streck AF, Freitas LB, Delgado-Cañedo A, Bonatto SL, Salzano FM (2004) B-F DNA sequence variability in Brazilian (blue-egg Caipira) chickens. Anim Genet 35:278–284

Livant EJ, Brigati JR, Ewald SJ (2004) Diversity and locus specificity of chicken MHC B class I sequences. Anim Genet 35:18–27

Maccari G, Robinson J, Ballingall K, Guethlein L, Grimholt U, Kaufmann J, Ho C-K, de Groot N, Flicek P, Bontrop R, Hammond J, Marsh S (2017) IPD-MHC 2.0: an improved inter-species database for the study of the major histocompatibility complex. Nucleic Acids Res 45(Database issue):D860–D864. https://doi.org/10.1093/nar/gkw1050

Maccari G, Robinson J, Bontrop RE, Otting N, de Groot NG, Ho C-S, Ballingall KT, Marsh SGE, Hammond JA (2018) IPD-MHC: nomenclature requirements for the non-human major histocompatibility complex in the next-generation sequencing era. Immunogenetics. https://doi.org/10.1007/s00251-018-1072-4

Marsh SGE, Albert ED, Bodmer WF, Bontrop RE, Dupont B, Erlich HA, Fernández-Viña M, Geraghty DE, Holdsworth R, Hurley CK, Lau M, Lee KW, Mach B, Maiers M, Mayr WR, Müller CR, Parham P, Petersdorf EW, Sasazuki T, Strominger JL, Svejgaard A, Terasaki PI, Tiercy JM, Trowsdale J (2010) Nomenclature for factors of the HLA system, 2010. Tissue Antigens 75:291–455

Miller MM, Bacon LD, Hala K, Hunt HD, Ewald SJ, Kaufman J, Zoorob R, Briles WE (2004) Nomenclature for the chicken major histocompatibility (B and Y) complex. Immunogenetics 56:261–279

Miller D, Tallmadge RL, Binns M, Zhu B, Mohamoud YA, Ahmed A, Brooks SA, Antczak DF (2016) Polymorphism at expressed DQ and DR loci in five common equine MHC haplotypes. Immunogenetics 69:145–156

Miltiadou D, Ballingall KT, Ellis SA, Russell GC, McKeever DJ (2005) Haplotype characterization of transcribed ovine major histocompatibility complex (MHC) class I genes. Immunogenetics 57:499–509

Nonaka MI, Aizawa K, Mitani H, Bannai HP, Nonaka M (2011) Retained orthologous relationships of the MHC Class I genes during euteleost evolution. Mol Biol Evol 28:3099–3112

Robinson J, Waller MJ, Stoehr P, Marsh SGE (2005) IPD - the Immuno Polymorphism Database. Nucleic Acids Res 33:D523–D526

Robinson J, Halliwell JA, Hayhurst JH, Flicek P, Parham P, Marsh SGE (2015) The IPD and IPD-IMGT/HLA database: allele variant databases. Nucleic Acids Res 43:D423–D431

Shiina T, Dijkstra JM, Shimizu S, Watanabe A, Yanagiya K, Kiryu I, Fujiwara A, Nishida-Umehara C, Kaba Y, Hirono I et al (2005) Interchromosomal duplication of major histocompatibility complex class I regions in rainbow trout (Oncorhynchus mykiss), a species with a presumably recent tetraploid ancestry. Immunogenetics 56: 878–893

Simonsen M, Hála K, Nicolaisen EM (1980) Linkage disequilibrium of MHC genes in the chicken. I. The B-F and B-G loci. Immunogenetics 10:103–112

Simonsen M, Crone M, Koch C, Hála K (1982) The MHC haplotypes of the chicken. Immunogenetics 16:513–532

Smith DM, Lunney JK, Martens GW, Ando A, Lee JH, Ho CS, Schook L, Renard C, Chardon P (2005a) Nomenclature for factors of the SLA class-I system, 2004. Tissue Antigens 65:136–149

Smith DM, Lunney JK, Ho CS, Martens GW, Ando A, Lee JH, Schook L, Renard C, Chardon P (2005b) Nomenclature for factors of the swine leukocyte antigen class II system, 2005. Tissue Antigens 66:623–639

Tallmadge RL, Lear TL, Antczak DF (2005) Genomic characterization of MHC class I genes of the horse. Immunogenetics 57:763–774

Tallmadge RL, Campbell JA, Miller DC, Antczak DF (2010) Analysis of MHC class I genes across horse MHC haplotypes. Immunogenetics 62:159–172

Wade CM, Giulotto E, Sigurdsson S, Zoli M et al (2009) Genome sequence, comparative analysis, and population genetics of the domestic horse. Science 326:865–867

WHO Nomenclature Committee (1968) Nomenclature for factors of the HL-a system. Bull World Health Organ 39:483–486

Worley K, Gillingham M, Jensen P, Kennedy LJ, Pizzari T, Kaufman J, Richardson DS (2008) Single locus typing of MHC class I and class II B loci in a population of red jungle fowl. Immunogenetics 60: 233–247