Cellular Immunology 347 (2020) 103999

Contents lists available at ScienceDirect

Cellular Immunology

journal homepage: www.elsevier.com/locate/ycimm

Research paper

NHP-immunome: A translational research-oriented database of non-human primate immune system proteins

David F. Plaza[a,b], Manuel F. Gómez[a], Manuel A. Patarroyo[a,b,*]

[a] Molecular Biology and Immunology Department, Fundación Instituto de Inmunología de Colombia (FIDIC), Bogotá, Colombia
[b] School of Medicine and Health Sciences, Universidad del Rosario, Bogotá, Colombia

ARTICLE INFO

Keywords:
Pan troglodytes
Macaca mulatta
Callithrix jacchus
Aotus nancymaae
Immune system orthology
Splicing/protein isoforms

ABSTRACT

We are currently living the advent of a new age for medicine in which basic research is being quickly translated into marketable drugs, and the widespread access to genomics data is allowing the design and implementation of personalized solutions to medical conditions. Non-human primates (NHP) have gained an essential role in drug discovery and safety testing due to their close phylogenetic relationship to humans. In this study, a collection of well characterized genes of the human immune system was used to define the orthology-based immunome in four NHP species, with carefully curated annotations available based on multi-tissue RNA-seq datasets. A broad variation in the frequency of expressed protein isoforms was observed between species. Finally, this analysis also revealed the lack of expression of at least four different chemokines in new-world primates. In addition, transcripts corresponding to four genes including interleukin 12 subunit alpha were expressed in humans but no other primate species analyzed. Access to the non-human primate immunome is available in http://www.fidic.org.co:90/proyecto/.

1. Introduction

Worldwide, the mouse is the most commonly used model in biomedical research; however, ≈80 million years of evolution [1] had resulted in unsurpassable molecular [2], anatomical [3] immunological [4] and neurological [5] differences between rodents and humans. These differences significantly limit the use of this model in the study and understanding of human disease, as well as the design of translational therapeutics for conditions with no homology in mice.

As new *in silico* technologies and *in vitro* assays had reduced the necessity of animals for research experimentation, guidelines from regulatory agencies in the United States and Europe still consider the results obtained in non-human primates (NHPs) as the gold standard to assess the safety and efficacy of therapies to be later used in humans. Also, data from the US Department of Agriculture in 2017 show that a record number of NHPs is being used in biomedical research (https://www.aphis.usda.gov/aphis/home/) making the design and implementation of tools allowing the scientific community to use NHPs in the most rational, relevant, efficient and humane way a crucial necessity.

Due to their evolutionary relatedness to human, non-human primates are currently used as animal models in a wide range of medical fields including infection biology [6], development of medical prosthetics [7], neuroscience [8], safety testing of pharmaceutical molecules [9], ophthalmology [10], vaccinology [11] and xenotransplantation [12], among others.

The molecular characterization of the NHP immunome is critical since most diseases where these animals are used as models have an important immune component. Efforts have been undertaken by us and other groups to define the coding sequence and allelic variation of proteins involved in antigen presentation and recognition in owl monkeys [13–22] and rhesus macaques [23,24]; however, a detailed description of immunome genes in different NHP species and a web-based platform to easily access this pivotal information for the biomedical community are urgently needed.

Our group has been using *Aotus nancymaae* for the past 28 years to model the infection course of the malaria parasite *Plasmodium falciparum* in humans and to test vaccine candidates against this disease that killed 445,000 people in 2016 only [25]. We had previously characterized Toll like receptor 9 [26], CD1b [27], CD45 [28], IL-2, IL-4, IL-6, IL-10, IFN-γ and TNF-α [29] from *A. nancymaae*; nevertheless, the sequencing of the owl monkey genome allows us now to take a deeper look into the homology of immune system components of this animal model in the fight against malaria.

* Corresponding author at: Molecular Biology and Immunology Department, Fundación Instituto de Inmunología de Colombia (FIDIC), Carrera 50 # 26-20, Bogotá, Colombia.
E-mail address: mapatarr.fidic@gmail.com (M.A. Patarroyo).

https://doi.org/10.1016/j.cellimm.2019.103999
Received 1 February 2019; Received in revised form 21 October 2019; Accepted 29 October 2019
Available online 31 October 2019

Previous computational efforts led to the molecular characterization of the set of genes playing a role in the human immune response excluding those loci involved in antigen recognition (immunoglobulins and T cell receptors) and presentation (major histocompatibility complex) [30–33]. In this study, we annotated and compiled the immunome of the chimpanzee (*Pan troglodytes*) and the rhesus macaque (*Macaca mulatta*), as well as the new-world common marmoset (*Callithrix jacchus*) and Nancy Ma's night (*Aotus nancymaae*) monkeys based on their orthology (defined as the best hit resulting from a reciprocal Blast search) to the curated human immunome. We also took a step further and included a comprehensive analysis on the different immune protein isoforms expressed on these species based on the latest RNA-seq-assisted annotations available [34,35]. Moreover, we looked at the human immune system components which are not expressed in NHPs and those human proteins with expressed orthologs in individual NHP species, as well as the human diseases associated to the genes encoding these proteins. Data access for the NHP research community is available in the NHP-immunome web resource (http://www.fidic.org.co:90/proyecto/).

2. Materials and methods

2.1. Human and NHP protein datasets

UniProtIDs corresponding to 831 proteins of the *Homo sapiens* immunome were retrieved from the human Immunome database [32]. FASTA sequences for all the known protein isoforms (2082 isoforms) were downloaded from UniProt [36]. Complete predicted proteomes based on tissue-specific RNA-seq datasets (Table 1) from *Pan troglodytes* (80,807 sequences), *Macaca mulatta* (55,312 sequences), *Callithrix jacchus* (45,251 sequences) and *Aotus nancymaae* (47,568 sequences) were obtained from the NCBI release reports 105 (March 21st 2018), 102 (December 22nd 2015), 104 (August 31st 2016) and 101 (June 28th 2017), respectively. All the annotations used were supported by RNA-seq datasets from the immune tissues shown in table 1 [34,35]. Other organs/tissues were also used to inform the annotation of the NHP genomes (Table S1) [34,35].

2.2. Search of immunome orthologues by reciprocal protein Blast

A reciprocal protein Blast [37,38] was performed between the protein isoforms of the human immunome and each of the complete annotated proteomes from *P. troglodytes*, *M. mulatta*, *C. jacchus* and *A. nancymaae* using 10%, 40%, 50%, 60% and 70% as the minimum identity values and 50 as the minimum percentage of query coverage for Blast matches. Identical sequences were collapsed in the five datasets to prevent the underrepresentation of genes with highly similar paralogs. As the number of protein isoforms and proteins identified was found to plateau at the 40% identity threshold (Fig. S1), we decided to use this percentage as the minimum identity required for a hit to be considered a true ortholog in NHPs. For comparison, a reciprocal protein Blast was computed between the complete annotated RefSeq proteome of human (113628 sequences, release report 109 from March 26th 2018) and the proteome of each of the four species of NHPs with the same settings previously used for the immunome alone and orthology groups were assembled (Table S2). The mean identity percentages shown in the last column of Table 2 correspond to the values calculated from these whole proteome comparisons. Finally, lists of matches were combined into a single dataset to create orthology groups including five species (human and each of the four NHPs). FASTA sequences corresponding to the immunome from the four NHP species were retrieved using the Filter FASTA tool from the Galaxy package and the list of ortholog IDs obtained from the reciprocal Blast [39].

2.3. Analysis of similarity in the splicing isoform expression pattern between NHP species

Average linkage hierarchical clustering was computed with Gene Cluster 3.0 (using the default weight options: cutoff 0.1; and Euclidian distance as the similarity metric) to visualize and group together isoform orthology groups according to their distribution in the different NHP species. Whereas bit score values/1000 were used for all the human protein isoforms with at least a match in one of the NHP species included, values of 0 were assigned to those isoforms showing no expressed orthologs in NHPs. The resulting heatmap was visualized with TreeView 3 [40]. A phylogenetic tree for the species compared was created with Phylot (https://phylot.biobyte.de/) and visualized with iTOL using the corresponding NCBI taxonomy IDs (*H. sapiens*: 9606, *P. troglodytes*: 9598, *M. mulatta*: 9544, *C. jacchus*: 9483 and *A. nancymaae*: 37293) [41]. Disease association scores of the genes differing between species were retrieved from the integrative database DisGeNET (DGN). DGN integrates data from genome-wide association studies (GWAS), literature automated searches and functional evidence derived from experiments in rodents. Scores above or equal to 0.5 were highlighted in the data since at most 8% (approx. 52,798 out of 628,685) of the gene-disease associations in the database have a DGN score in this value range. Furthermore, scores above or equal to 0.5 are supported by two or more information sources [42]. Enrichment of gene ontology (GO) terms for highly conserved immunome proteins resulting from the hierarchical clustering analysis was computed with PANTHER 11 using the immunome (not the entire human proteome) as a reference [43].

Table 1
Tissues with RNA-seq datasets used for annotation of immune genes in *P. troglodytes*, *M. mulatta*, *C. jacchus* and *A. nancymaae* [34,35].

Species	Immune organs with RNA-seq datasets	Number of datasets	Total number of sequenced reads	BioProject IDs
P. troglodytes	Blood	1	244,621,062	PRJNA261948 (Peng X, Thierry-Mieg J)
	Bone marrow	1	117,895,252	PRJNA271912 (Peng X, Thierry-Mieg J)
	Lymph node	1	96,439,946	PRJNA271912 (Peng X, Thierry-Mieg J)
	Spleen	1	112,324,566	PRJNA271912 (Peng X, Thierry-Mieg J)
	Thymus	1	138,058,532	PRJNA271912 (Peng X, Thierry-Mieg J)
M. mulatta	Bone marrow	1	79,877,686	PRJNA261940 (Peng X, Thierry-Mieg J),
	Spleen	2	183,030,222	PRJNA183644 (Chen JY, Peng Z)
	Thymus	2	121,650,454	PRJNA261940 (Peng X, Thierry-Mieg J),
C. jacchus	Bone marrow	1	158,879,166	PRJNA271912 (Peng X, Thierry-Mieg J)
	Lymph node	1	109,162,936	PRJNA271912 (Peng X, Thierry-Mieg J)
	Spleen	1	115,423,250	PRJNA271912 (Peng X, Thierry-Mieg J)
A. nancymaae	Blood	1	76,305,936	PRJNA280454
	Bone marrow	1	95,736,656	PRJNA280454
	Lymph node	1	90,486,978	PRJNA280454
	Spleen	1	100,318,156	PRJNA280454

Table 2
Summary of the NHP immunome showing the number of orthologous proteins and protein isoforms found in the different species, as well as the median and mean sequence identity to humans for the four NHP immunomes. For comparison, the last column on the right shows the mean proteome identity to humans for the four NHPs.

	Orthologous proteins found	Orthologous isoforms found	Immunome median identity to human (%)	Immunome mean identity to human (%)	Proteome mean identity to human RefSeq (%)
P. troglodytes	800	1135	99.19	98.71	99.10
M. mulatta	790	999	95.53	94.15	96.77
C. jacchus	759	907	90.86	89.03	94.11
A. nancymaae	761	909	91.30	89.66	94.46

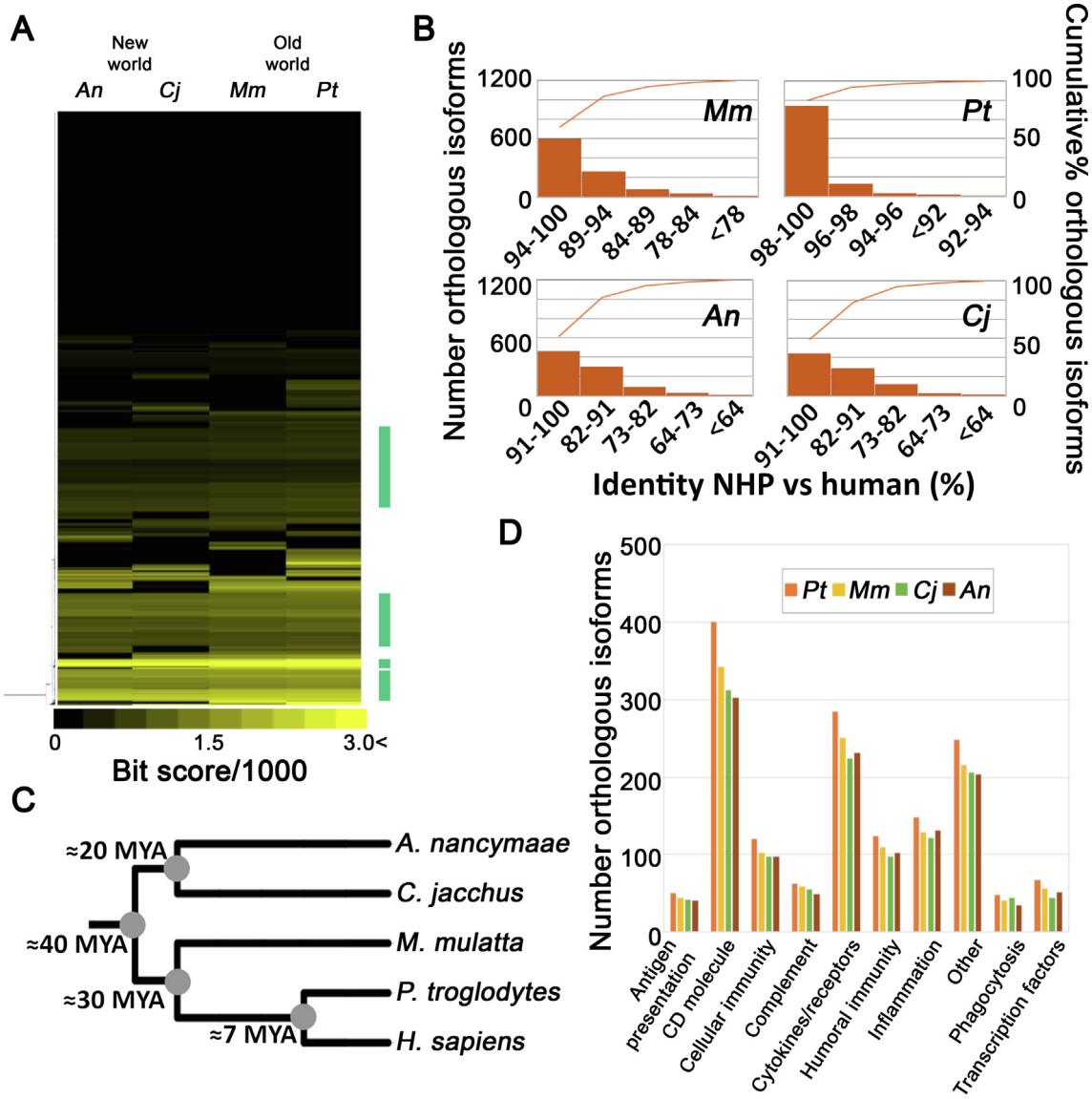

Fig. 1. Immunome comparison between human, chimpanzee (*Pt*) and rhesus macaque (*Mm*), as well as the common marmoset (*Cj*) and Nancy Ma's night (*An*) monkeys. (A) An average linkage hierarchical clustering analysis was performed using the bit-score/1000 as this value integrates identity and coverage percentages between two sequences into a single metric. Euclidian distance was the similarity metric used. Conserved isoform clusters in the four NHP species are marked with green bars on the right-hand side of the heatmap. (B) Identity distribution for the isoform orthologs found in NHPs. Left and right Y axes show the number of orthologous isoforms in each of the identity bins and the cumulative percentage of orthologous isoforms, respectively. (C) PhyloT phylogenetic tree for human and the four NHP species studied. The tree was visualized with iTOL. Approximate clade divergence times were obtained from the latest primate molecular phylogeny available [44]. (D) Functional classification and number of orthologous isoforms per category in the individual NHP immunomes. Functional categories correspond to the ones previously defined for humans [32].

3. Results

3.1. The NHP immunome reflects the evolutionary history of NHPs and human

From the 831 genes of the human immunome analyzed for this study, 827 genes showed at least one expressed ortholog in one of the four NHP species. In comparison, from 2082 protein isoforms of the human immunome, 1313 had at least one expressed ortholog in an NHP (Fig. 1A, Table S3). Identity distribution analysis for orthologous isoforms found in the four NHPs showed that in chimpanzees, at least 50% of the immunome isoforms identified showed more than 99.2% identity to their corresponding human orthologs. For the other NHPs, the immunome median identities to their human orthologs were 95.5% for *M. mulatta*, 90.9% for *C. jacchus* and 91.3% for *A. nancymaae* (Table 2 and Fig. 1B). These identity percentages accurately reflect the evolutionary history of the four NHPs and humans (Fig. 1C). The last common ancestor of human and chimpanzee lived ≈7 MYA. Old-world superfamilies Hominoidea (human and chimpanzee) and Cercopithecoidea (rhesus macaque) shared a common ancestor ≈30 MYA (Fig. 1C). Finally, the last common ancestor of the parvorders Catarrhini (old-world primates) and Platyrrhini (new-world primates) lived approximately 40 MYA [44]. It is important to note that the set of immunome proteins in the four individual NHPs showed a lower mean identity percentage to humans than the NHP proteome as a whole (Table 2). This observation suggests that environmental forces might be positively selecting sequence or functional diversification in components of the immune system which are not directly involved in antigen presentation or recognition.

It was also found that all reported splicing isoforms of the protein tyrosine kinase Fyn, a key T cell regulator that experiences alternative splicing in a protein-coding region resulting in the production of isoforms FynB (expressed in the brain) and FynT (expressed in T cells) [45], are present in the four NHPs. The same conservation of previously described splicing isoforms [46] was observed for the negative T cell regulator CTLA-4. Intriguingly, described splicing isoforms for immune proteins such as membrane cofactor protein (Mcp) [47], CD45 [48], CD6 [49], CD8 [50] and IL-4 [51], among many others, display a patchwork distribution in the different primates. One explanation for this is that alternative splicing might be a critical factor determining interspecies differences in the immune response; however, the isoform distribution could also be explained by tissue and organ coverage differences between species leading to a higher number of isoforms being identified for widely studied species such as human, chimpanzee and rhesus macaque in comparison to emergent NHP models from the new-world.

3.2. Proteins playing a role in cell-to-cell contact and communication are the most abundant in the NHP immunome

As previously seen in humans, membrane cluster of differentiation (CD) molecules as well as cytokines and their receptors are the most abundant proteins of the immunome in NHPs (Fig. 1D). These proteins are crucial for the activation of the immune response upon challenge and allow the communication between different cellular components of the immune system.

3.3. Four human immunome genes are not expressed in NHP immune tissues

Interestingly, the expression for four genes was undetectable in any of the immune RNA-seq datasets [34,35] from the four NHPs (Table 3).

These genes showed high DGN scores associated to a number of human diseases (Table 3), such as primary biliary cirrhosis (DGN score: 0.630) and kappa chain deficiency (DGN score: 0.600). These conditions might not be properly modelled by any of the four NHPs analyzed in this study.

3.4. The core primate immunome comprises 671 protein isoforms

Analysis of the NHP orthologs to the known splicing isoforms for immune proteins in human revealed that only 671 isoforms (32.22% of the total number of human isoforms), representing 629 genes, are conserved among the five species compared (green sidebars in Fig. 1A, Table S4). Furthermore, a GO term enrichment analysis for the cluster comprising the 22 most highly conserved immunome components in NHPs (correlation coefficient: 0.9868, second last side bar from top to bottom in Fig. 1A) revealed that this set of proteins has an 12.20 fold enrichment (false discovery rate = 0.0454) for the term “peptidase inhibitor” when the entire human immunome is used as a reference. Interestingly, two out of five proteins classified under this GO term are components of the complement system: C5 and C3. The other three proteins classified as peptidase inhibitors are Alpha-2-macroglobulin-like protein 1 (A2ML1), CD109 and Alpha-2-macroglobulin (A2MG).

3.5. Seven orthologs to human immunome components are found to be expressed in only one NHP

The orthology analysis of individual NHP species revealed that *P. troglodytes* expresses five unique human orthologs (only shared with humans but no other NHP). Similarly, *M. mulatta* and *C. jacchus* were found to transcribe one unique human ortholog each (Fig. 1A). According to this, the chimpanzee could be a uniquely suitable disease model for Crohn’s disease (DGN score: 0.900), ulcerative colitis (DGN score: 0.900) and granulomatous synovitis (DGN score: 0.800) due to the presence of an ortholog for *nod2* in its genome. Similarly, the rhesus macaque is the only NHP with an ortholog for *cfhr4*, making it a potentially good model organism for the study of atypical hemolytic uremic syndrome (DGN score: 0.300) (Table 4). Finally, *C. jacchus* expresses an ortholog for *mica* that makes this monkey a more human-like model for the study of the immune response against liver carcinomas (DGN score: 0.400).

Additionally, 18 genes (2.17% of the human immunome) were found to be present in old-world primates only, suggesting that a limited number of immune genes either significantly diverged following the speciation event that separated new- and old- world primates approx. 40 million years ago or were present in the last common ancestor of these two clades but were later lost in new-world primates (Table 5) [44].

4. Discussion

Non-human primates are widely used in biomedical research and, during the past decade, have been regaining preponderance over other animal models (e.g. mice and rats) due to their phylogenetic relatedness to humans and the need to translate the findings of basic research into druggable targets and marketable therapies. In this study, we computationally characterized the protein components of the immune system of four species of NHPs (*P. troglodytes*, *M. mulatta*, *C. jacchus* and *A. nancymaae*) excluding those proteins directly involved in antigen presentation and recognition based on the orthology relationship to previously described components of the human immunome [30–33]. Interestingly, the mean identity percentage was lower for the immunome compared to the genome-wide proteome of the four species compared, suggesting that there is evolutionary pressure positively selecting for sequence variability in immune components which can be hijacked (mimicry or functional blockage) by infectious microorganisms [52].

The analysis of different protein isoforms expressed in several NHP tissues including blood, bone marrow, lymph nodes, spleen and thymus, showed that the expression pattern of these immunome variants differs between NHP species. This expression variability of protein isoforms has been previously reported for human, chimpanzee and rhesus macaque [53,54] and might be explained by interspecific differences in the sequence preferences of the splicing machinery responsible for the

Table 3
Four immunome genes are expressed exclusively in humans. The genes in the table do not show an expressed ortholog (using the 40% identity and 50% coverage thresholds) in any of the four NHP primates they were compared against. The top three associated diseases according to their DGN score (ranging from 0 for a complete lack of gene-disease association and 1 for diseases which are specifically caused by alterations in the associated gene) are shown for every gene. DGN scores greater than or equal to 0.5 are shown in bold.

UniprotID	Protein name	Gene names	Top associated human disease	DisGeNET score
Q7LDP2	Interleukin-12 p35 subunit (Fragment)		**Primary biliary cirrhosis**	**0.630**
			Biliary cirrhosis	0.400
			Sjogren's syndrome	0.320
P01834	Immunoglobulin kappa constant	IGKC	**Kappa-Chain Deficiency**	**0.600**
			Focal myositis	0.300
			Idiopathic inflammatory myopathies	0.300
Q9UBD3	Cytokine SCM-1 beta	XCL2 SCYC2	Ovarian carcinoma	0.010
			Malignant ascites	0.010
			Osteoarthritis	0.010
P06028	Glycophorin-B	GYPB GPB	Malaria	0.320
			Chronic alcoholic intoxication	0.020
			Menkes Kinky hair syndrome	0.010

processing of primary transcripts. Alternative splicing has also been shown to be an important mechanism of gene regulation during the immune response [55] and there is a patchwork distribution of splicing isoforms in NHPs involving crucial genes such as Mcp, CD45, CD6, CD8 and IL4. It is important to note here that the set of tissues from which RNA sequencing libraries are available varies between species and a more complete coverage of tissue-specific transcriptomes is available for the most widely studied species (human, chimpanzee and rhesus macaque); nevertheless, the transcriptome of several immune tissues had been sequenced for the five species we analyzed as shown in Table 1.

The global monoclonal antibody market was worth USD89 billion in 2016 [56]. New antibody therapies targeting modulatory molecules hit the market every year and must be tested in NHPs for efficacy and safety. The precise knowledge of sequence in NHP orthologs to therapeutic targets in humans is a crucial need for the pharmaceutical industry, since minor differences in the epitopes recognized by monoclonal antibodies in NHPs and humans can lead to major changes in binding affinities with disastrous consequences for clinical volunteers. The comprehensive catalog of immune-related protein isoforms here presented provides a valuable tool for those groups testing antibody therapies in NHPs, allowing them to interpret results that might be explained by differences in ortholog structure.

Functional classification of immunome isoforms showed a similar distribution in the different NHPs and no species-specific enrichment for any single category was observed. In addition, seven orthologs were shown to be expressed either in chimpanzee (5 genes), macaque (1 gene) or common marmoset monkey (1 gene), exclusively. These genes are associated to diseases such as Crohn's disease, ulcerative colitis, granulomatous synovitis or pneumonia, indicating that chimpanzees, macaques or common marmoset monkeys might be the only suitable NHP species to model these medical conditions.

Table 4
Human genes with expressed orthologs in a single species of NHP. Disease association scores (DGN scores) are shown. Values greater than or equal to 0.5 (the maximum possible association score is 1) are show in bold.

Species /UniProtID in human	Protein name	Gene names	Identity to human (%)	Top associated human disease	DisGeNET score
P. troglodytes					
P43626	Killer cell immunoglobulin-like receptor 2DL1	KIR2DL1 CD158A, NKAT1	88.82	Acute graft versus host disease	0.030
				Rheumatoid arthritis	0.020
				Spontaneous abortion	0.010
Q9HC29	Nucleotide-binding oligomerization domain-containing protein 2	NOD2 CARD15, IBD1	98.94	**Crohn disease**	**0.900**
				Ulcerative colitis	**0.900**
				Granulomatous synovitis	**0.800**
P19876	C-X-C motif chemokine 3	CXCL3 GRO3, GROG, SCYB3	96.26	Malignant neoplasm of breast	0.310
				Mammary neoplasm	0.310
				Breast carcinoma	0.310
P10321	MHC class I antigen Cw*7	HLA-C HLAC	92.08	Psoriasis	0.400
				Psoriatic arthritis	0.400
				HIV infections	0.400
P09341	Growth-regulated alpha protein	CXCL1 GRO, GRO1, GROA, MGSA, SCYB1	100.00	**Pneumonia**	**0.530**
				Pneumonitis	0.320
				Chronic obstructive airway disease	0.310
M. mulatta					
Q92496	Complement factor H-related protein 4	CFHR4 CFHL4, FHR4	87.48	Atypical hemolytic uremic syndrome	0.300
				Age-related macular degeneration	0.100
				Maculopapular eruption	0.100
C. jacchus					
Q29983	MHC class I polypeptide-related sequence A	MICA PERB11.1	66.85	Liver carcinoma	0.400
				Oral submucous fibrosis	0.310
				Coronary aneurysm	0.300

Table 5
Eighteen immunome proteins are present in old-world primates exclusively. This set of immune components either diverged significantly in old-world primates over the past ≈40 million years or was present in the common ancestor of new- and old-world primates but was later lost in the former.

UniProtID	Protein name	Gene names
P40198	Carcinoembryonic antigen-related cell adhesion molecule 3	CEACAM3 CD66D, CGM1
P55774	C−C motif chemokine 18	CCL18 AMAC1, DCCK1, MIP4, PARC, SCYA18
Q9Y258	C−C motif chemokine 26	CCL26 SCYA26, UNQ216/PRO242
Q8WTT0	C-type lectin domain family 4 member C	CLEC4C BDCA2, CLECSF11, CLECSF7, DLEC, HECL
P10147	C−C motif chemokine 3	CCL3 G0S19-1, MIP1A, SCYA3
Q01524	Defensin-6	DEFA6 DEF6
O43699	Sialic acid-binding Ig-like lectin 6	SIGLEC6 CD33L, CD33L1, OBBP1
Q8NHJ6	Leukocyte immunoglobulin-like receptor subfamily B member 4	LILRB4 ILT3, LIR5
P30511	HLA class I histocompatibility antigen, alpha chain F	HLA-F HLA-5.4, HLAF
P55773	C−C motif chemokine 23	CCL23 MIP3, MPIF1, SCYA23
O14798	Tumor necrosis factor receptor superfamily member 10C	TNFRSF10C DCR1, LIT, TRAILR3, TRID, UNQ321/PRO366
Q99788	Chemokine-like receptor 1	CMKLR1 CHEMR23, DEZ
Q6UWJ8	CD164 sialomucin-like 2 protein	CD164L2 UNQ6122/PRO20044
A6NI73	Leukocyte immunoglobulin-like receptor subfamily A member 5	LILRA5 ILT11, LILRB7, LIR9
Q9HC73	Cytokine receptor-like factor 2	CRLF2 CRL2, ILXR, TSLPR
P36980	Complement factor H-related protein 2	CFHR2 CFHL2, FHR2, HFL3
Q6Q8B3	Cell surface glycoprotein CD200 receptor 2	CD200R1L CD200R2
P13765	MHC class II antigen DOB	HLA-DOB

Four chemokines (CCL3, CCL18, CCL23 and CCL26) were found to be absent in new-world primates. These chemokines mediate the mobilization of B and T lymphocytes (CCL3, CCL18 and CCL23), eosinophils and basophils (CCL26) and neutrophils (CCL23), binding the receptors CCR1, CCR3, CCR4, CCR5 and CCR8. Besides CCL3 and CCL23, CCR1 can also be bound by CCL2, CCL4, CCL5, CCL7 and CCL15. Similarly, CCR3 is bound not only by CCL26, but also by CCL4, CCL5, CCL7, CCL11, CCL13 and CCL15. CCR1 is present in new- and old-world primates while CCR3 isoforms are absent in *C. jacchus*. These broad functional redundancy and promiscuity observed for several chemotactic agents and their receptors [57] explain at least in part why some of them seem to be dispensable in the immune system of new-world primates.

In conclusion, non-human primates are a valuable tool in the dawn of translational medicine due to their anatomical and molecular similarity to human. We characterized the immunome of *P. troglodytes*, *M. mulatta*, *C. jacchus* and *A. nancymaae* finding that at least 827 proteins of the human immune system have a transcribed ortholog in at least one of the species compared. Considering that tissue and organ coverage for RNA-libraries differs between NHP species, widespread differences in splicing isoform usage were observed, suggesting that alternative splicing could be a major driving force in the diversification and regulation of immune proteins. The NHP immunome can be accessed in http://www.fidic.org.co:90/proyecto/.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cellimm.2019.103999.

References

[1] N.M. Foley, M.S. Springer, E.C. Teeling, Mammal madness: is the mammal tree of life not yet resolved? Philos. Trans. R. Soc. Lond. B Biol. Sci. 371 (2016).

[2] D.M. Church, L. Goodstadt, L.W. Hillier, M.C. Zody, S. Goldstein, X. She, C.J. Bult, R. Agarwala, J.L. Cherry, M. DiCuccio, W. Hlavina, Y. Kapustin, P. Meric, D. Maglott, Z. Birtle, A.C. Marques, T. Graves, S. Zhou, B. Teague, K. Potamousis, C. Churas, M. Place, J. Herschleb, R. Runnheim, D. Forrest, J. Amos-Landgraf, D.C. Schwartz, Z. Cheng, K. Lindblad-Toh, E.E. Eichler, C.P. Ponting, C. Mouse Genome Sequencing, Lineage-specific biology revealed by a finished genome assembly of the mouse, PLoS Biol 7 (2009) e1000112.

[3] O. Bodenreider, T.F. Hayamizu, M. Ringwald, S. De Coronado, S. Zhang, Of mice and men: aligning mouse and human anatomies, AMIA Annu. Symp. Proc. (2005) 61–65.

[4] J. Mestas, C.C. Hughes, Of mice and not men: differences between mouse and human immunology, J. Immunol. 172 (2004) 2731–2738.

[5] J.A. Miller, S. Horvath, D.H. Geschwind, Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathways, Proc. Natl. Acad. Sci. U.S.A. 107 (2010) 12698–12703.

[6] J.D. Estes, S.W. Wong, J.M. Brenchley, Nonhuman primate models of human viral infections, Nat. Rev. Immunol. 18 (2018) 390–404.

[7] N.J. Killian, M. Vurro, S.B. Keith, M.J. Kyada, J.S. Pezaris, Perceptual learning in a non-human primate model of artificial vision, Sci. Rep. 6 (2016) 36329.

[8] P.R. Roelfsema, S. Treue, Basic neuroscience research with nonhuman primates: a small but indispensable component of biomedical research, Neuron 82 (2014) 1200–1204.

[9] F.R. Brennan, J. Cavagnaro, K. McKeever, P.C. Ryan, M.M. Schutten, J. Vahle, G.F. Weinbauer, E. Marrer-Berger, L.E. Black, Safety testing of monoclonal antibodies in non-human primates: Case studies highlighting their impact on human risk assessment, MAbs 10 (2018) 1–17.

[10] V. Raghunathan, J.S. Eaton, B.J. Christian, J.T. Morgan, J.N. Ver Hoeve, C.C. Yang, H. Gong, C.A. Rasmussen, P.E. Miller, P. Russell, T.M. Nork, C.J. Murphy, Biomechanical, ultrastructural, and electrophysiological characterization of the non-human primate experimental glaucoma model, Sci. Rep. 7 (2017) 14329.

[11] T. Rivera-Hernandez, D.G. Carnathan, P.M. Moyle, I. Toth, N.P. West, P.R. Young, G. Silvestri, M.J. Walker, The contribution of non-human primate models to the development of human vaccines, Discov. Med. 18 (2014) 313–322.

[12] D. Dufrane, P. Gianello, Pig islet xenotransplantation into non-human primate model, Transplantation 86 (2008) 753–760.

[13] C.F. Suarez, L. Pabon, A. Barrera, J. Aza-Conde, M.A. Patarroyo, M.E. Patarroyo, Structural analysis of owl monkey MHC-DR shows that fully-protective malaria vaccine components can be readily used in humans, Biochem. Biophys. Res. Commun. 491 (2017) 1062–1069.

[14] C.A. Moncada, E. Guerrero, P. Cardenas, C.F. Suarez, M.E. Patarroyo, M.A. Patarroyo, The T-cell receptor in primates: identifying and sequencing new owl monkey TRBV gene sub-groups, Immunogenetics 57 (2005) 42–52.

[15] P.P. Cardenas, C.F. Suarez, P. Martinez, M.E. Patarroyo, M.A. Patarroyo, MHC class I genes in the owl monkey: mosaic organisation, convergence and loci diversity, Immunogenetics 56 (2005) 818–832.

[16] D. Garzon-Ospina, C. Lopez, L.F. Cadavid, M.E. Patarroyo, M.A. Patarroyo, Identification and diversity of killer cell Ig-like receptors in Aotus vociferans, a New World monkey, PLoS One 8 (2013) e79731.

[17] J.E. Guerrero, D.P. Pacheco, C.F. Suarez, P. Martinez, F. Aristizabal, C.A. Moncada, M.E. Patarroyo, M.A. Patarroyo, Characterizing T-cell receptor gamma-variable gene in Aotus nancymaae owl monkey peripheral blood, Tissue Antigens 62 (2003) 472–482.

[18] E.C. Hernandez, C.F. Suarez, C.A. Parra, M.A. Patarroyo, M.E. Patarroyo, Identification of five different IGHV gene families in owl monkeys (Aotus nancymaae), Tissue Antig. 66 (2005) 640–649.

[19] O.L. Diaz, C.A. Daubenberger, R. Rodriguez, M. Naegeli, A. Moreno, M.E. Patarroyo, G. Pluschke, Immunoglobulin kappa light-chain V, J, and C gene sequences of the owl monkey Aotus nancymaae, Immunogenetics 51 (2000) 212–218.

[20] M.C. Suarez, M.A. Patarroyo, M.E. Patarroyo, Characterisation and comparative analysis of MHC-DPA1 exon 2 in the owl monkey (Aotus nancymaae), Gene 470 (2011) 37–45.

[21] C.F. Suarez, M.E. Patarroyo, E. Trujillo, M. Estupinan, J.E. Baquero, C. Parra, R. Rodriguez, Owl monkey MHC-DRB exon 2 reveals high similarity with several HLA-DRB lineages, Immunogenetics 58 (2006) 542–558.

[22] D. Diaz, M. Naegeli, R. Rodriguez, J.J. Nino-Vasquez, A. Moreno, M.E. Patarroyo, G. Pluschke, C.A. Daubenberger, Sequence and diversity of MHC DQA and DQB genes of the owl monkey Aotus nancymaae, Immunogenetics 51 (2000) 528–537.

[23] S. Shen, C.W. Pyo, Q. Vu, R. Wang, D.E. Geraghty, The essential detail: the genetics and genomics of the primate immune response, ILAR J. 54 (2013) 181–195.

[24] T. Shiina, A. Blancher, H. Inoko, J.K. Kulski, Comparative genomics of the human, macaque and mouse major histocompatibility complex, Immunology 150 (2017) 127–138.

[25] G.W.H. Organization, World malaria report 2017, (2017).
[26] R. Spirig, E. Peduzzi, M.E. Patarroyo, G. Pluschke, C.A. Daubenberger, Structural and functional characterisation of the Toll like receptor 9 of Aotus nancymaae, a non-human primate model for malaria vaccine development, Immunogenetics 57 (2005) 283–288.
[27] F. Castillo, C. Guerrero, E. Trujillo, G. Delgado, P. Martinez, L.M. Salazar, P. Barato, M.E. Patarroyo, C. Parra-Lopez, Identifying and structurally characterizing CD1b in Aotus nancymaae owl monkeys, Immunogenetics 56 (2004) 480–489.
[28] G.E. Montoya, J.P. Vernot, M.E. Patarroyo, Partial characterization of the CD45 phosphatase cDNA in the owl monkey (Aotus vociferans), Am. J. Primatol. 57 (2002) 1–11.
[29] E.C. Hernandez, C.F. Suarez, J.A. Mendez, S.J. Echeverry, L.A. Murillo, M.E. Patarroyo, Identification, cloning, and sequencing of different cytokine genes in four species of owl monkey, Immunogenetics 54 (2002) 645–653.
[30] C. Ortutay, M. Siermala, M. Vihinen, Molecular characterization of the immune system: emergence of proteins, processes, and domains, Immunogenetics 59 (2007) 333–348.
[31] C. Ortutay, M. Siermala, M. Vihinen, ImmTree: database of evolutionary relationships of genes and proteins in the human immune system, Immun. Res. 3 (2007) 4.
[32] C. Ortutay, M. Vihinen, Immunome: a reference set of genes and proteins for systems biology of the human immune system, Cell Immunol. 244 (2006) 87–89.
[33] K. Rannikko, C. Ortutay, M. Vihinen, Immunity genes and their orthologs: a multispecies database, Int. Immunol. 19 (2007) 1361–1370.
[34] J.Y. Chen, Z. Peng, R. Zhang, X.Z. Yang, B.C. Tan, H. Fang, C.J. Liu, M. Shi, Z.Q. Ye, Y.E. Zhang, M. Deng, X. Zhang, C.Y. Li, RNA editome in rhesus macaque shaped by purifying selection, PLoS Genet 10 (2014) e1004274.
[35] X. Peng, J. Thierry-Mieg, D. Thierry-Mieg, A. Nishida, L. Pipes, M. Bozinoski, M.J. Thomas, S. Kelly, J.M. Weiss, M. Raveendran, D. Muzny, R.A. Gibbs, J. Rogers, G.P. Schroth, M.G. Katze, C.E. Mason, Tissue-specific transcriptome sequencing analysis expands the non-human primate reference transcriptome resource (NHPRTR), Nucl. Acids Res. 43 (2015) D737–742.
[36] C. The UniProt, UniProt: the universal protein knowledgebase, Nucl. Acids Res. 45 (2017) D158–D169.
[37] C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, T.L. Madden, BLAST+: architecture and applications, BMC Bioinform. 10 (2009) 421.
[38] P.J. Cock, J.M. Chilton, B. Gruning, J.E. Johnson, N. Soranzo, NCBI BLAST+ integrated into Galaxy, Gigascience 4 (2015) 39.
[39] E. Afgan, D. Baker, B. Batut, M. van den Beek, D. Bouvier, M. Cech, J. Chilton, D. Clements, N. Coraor, B.A. Gruning, A. Guerler, J. Hillman-Jackson, S. Hiltemann, V. Jalili, H. Rasche, N. Soranzo, J. Goecks, J. Taylor, A. Nekrutenko, D. Blankenberg, The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update, Nucl. Acids Res. 46 (2018) W537–W544.
[40] M.B. Eisen, P.T. Spellman, P.O. Brown, D. Botstein, Cluster analysis and display of genome-wide expression patterns, Proc. Natl. Acad. Sci. U.S.A. 95 (1998) 14863–14868.
[41] I. Letunic, P. Bork, Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees, Nucl. Acid. Res. 44 (2016) W242–245.
[42] J. Pinero, A. Bravo, N. Queralt-Rosinach, A. Gutierrez-Sacristan, J. Deu-Pons, E. Centeno, J. Garcia-Garcia, F. Sanz, L.I. Furlong, DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants, Nucl. Acids Res. 45 (2017) D833–D839.
[43] H. Mi, X. Huang, A. Muruganujan, H. Tang, C. Mills, D. Kang, P.D. Thomas, PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements, Nucl. Acids Res. 45 (2017) D183–D189.
[44] P. Perelman, W.E. Johnson, C. Roos, H.N. Seuanez, J.E. Horvath, M.A. Moreira, B. Kessing, J. Pontius, M. Roelke, Y. Rumpler, M.P. Schneider, A. Silva, S.J. O'Brien, J. Pecon-Slattery, A molecular phylogeny of living primates, PLoS Genet. 7 (2011) e1001342.
[45] E.H. Palacios, A. Weiss, Function of the Src-family kinases, Lck and Fyn, in T-cell development and activation, Oncogene 23 (2004) 7990–8000.
[46] G. Magistrelli, P. Jeannin, N. Herbault, A. Benoit De Coignac, J.F. Gauchat, J.Y. Bonnefoy, Y. Delneste, A soluble form of CTLA-4 generated by alternative splicing is expressed by nonstimulated human T cells, Eur. J. Immunol. 29 (1999) 3596–3602.
[47] A.L. Astier, G. Meiffren, S. Freeman, D.A. Hafler, Alterations in CD46-mediated Tr1 regulatory T cells in patients with multiple sclerosis, J. Clin. Invest. 116 (2006) 3252–3257.
[48] K.W. Lynch, A. Weiss, A model system for activation-induced alternative splicing of CD45 pre-mRNA in T cells implicates protein kinase C and Ras, Mol. Cell. Biol. 20 (2000) 70–80.
[49] M.A. Castro, M.I. Oliveira, R.J. Nunes, S. Fabre, R. Barbosa, A. Peixoto, M.H. Brown, J.R. Parnes, G. Bismuth, A. Moreira, B. Rocha, A.M. Carmo, Extracellular isoforms of CD6 generated by alternative splicing regulate targeting of CD6 to the immunological synapse, J. Immunol. 178 (2007) 4351–4361.
[50] D. Thakral, J. Dobbins, L. Devine, P.B. Kavathas, Differential expression of the human CD8beta splice variants and regulation of the M-2 isoform by ubiquitination, J. Immunol. 180 (2008) 7431–7442.
[51] W.J. Alms, S.P. Atamas, V.V. Yurovsky, B. White, Generation of a variant of human interleukin-4 by alternative splicing, Mol. Immunol. 33 (1996) 361–370.
[52] G. Hajishengallis, J.D. Lambris, Microbial manipulation of receptor crosstalk in innate immunity, Nat. Rev. Immunol. 11 (2011) 187–200.
[53] J.A. Calarco, Y. Xing, M. Caceres, J.P. Calarco, X. Xiao, Q. Pan, C. Lee, T.M. Preuss, B.J. Blencowe, Global analysis of alternative splicing differences between humans and chimpanzees, Genes Dev. 21 (2007) 2963–2975.
[54] R. Blekhman, J.C. Marioni, P. Zumbo, M. Stephens, Y. Gilad, Sex-specific and lineage-specific alternative splicing in primates, Genome Res. 20 (2010) 180–189.
[55] N.M. Martinez, K.W. Lynch, Control of alternative splicing in immune responses: many regulators, many predictions, much still to learn, Immunol. Rev. 253 (2013) 216–236.
[56] P.J. Carter, G.A. Lazar, Next generation antibody drugs: pursuit of the 'high-hanging fruit', Nat. Rev. Drug Discov. 17 (2018) 197–223.
[57] A. Mantovani, The chemokine system: redundancy for robust outputs, Immunol. Today 20 (1999) 254–257.