Online Supplemental Tables:

Supplemental Table 1

Characteristics of *Monodelphis* genes and their *Homo* orthologues lying on each chromosome. *Monodelphis* genes that lie on MDO4, MDO7 and MDOX and have synteny to human X chromosome are also displayed separately. Shown are median values and, in parentheses, lower and upper quartiles.

										MDO4	**MDO7**	**MDOX**
	MDO1	**MDO2**	**MDO3**	**MDO4**	**MDO5**	**MDO6**	**MDO7**	**MDO8**	**MDOX**	**With synteny to HSAX**		
1-to-1 orthologues	2988	2535	1453	1590	984	856	619	1066	169	54	55	161
Gene counts	4056	3353	2234	2443	1509	1166	848	1426	283	62	68	178
Chromosome length (Mbp)	748.1	541.6	526.1	430.1	308.9	244.8	262.6	308.5	60.7	n/a	n/a	n/a
d_N / d_S *	0.084 (0.044 - 0.147)	0.092 (0.046 - 0.162)	0.080 (0.041 - 0.142)	0.088 (0.047 - 0.157)	0.085 (0.045 - 0.153)	0.084 (0.044 - 0.148)	0.095 (0.048 - 0.166)	0.090 (0.044 - 0.159)	0.095 (0.047 - 0.151)	0.094 (0.042 - 0.156)	0.101 (0.061 - 0.162)	0.089 (0.047 - 0.151)
d_N *	0.092 (0.042 - 0.172)	0.095 (0.044 - 0.175)	0.090 (0.041 - 0.169)	0.097 (0.044 - 0.176)	0.094 (0.047 - 0.172)	0.106 (0.052 - 0.189)	0.096 (0.043 - 0.172)	0.089 (0.040 - 0.163)	0.119 (0.057 - 0.203)	0.069 (0.029 - 0.144)	0.115 (0.061 - 0.196)	0.117 (0.053 - 0.192)
d_S *	1.038 (0.761 - 1.456)	0.997 (0.755 - 1.352)	1.011 (0.739 - 1.605)	0.990 (0.727 - 1.410)	1.007 (0.762 - 1.380)	1.161 (0.839 - 1.602)	0.937 (0.742 - 1.255)	0.946 (0.716 - 1.223)	1.213 (0.907 - 1.714)	0.839 (0.619 - 0.983)	1.062 (0.814 - 1.433)	1.206 (0.909 - 1.594)
Amino acid sequence identity*	81.3% (70.1% - 90.4%)	80.9% (69.7% - 89.6%)	81.6% (70.0% - 91.0%)	81.0% (69.0% - 90.0%)	81.6% (70.3% - 89.9%)	79.4% (68.4% - 89.0%)	81.0% (70.8% - 89.5%)	82.0% (71.8% - 90.5%)	76.1% (65.7% - 87.1%)	84.5% (71.1% - 94.0%)	79.9% (69.7% - 88.1%)	77.3% (66.7% - 87.1%)
Pairwise alignment coverage of the longer sequence*	94.5% (80.9% - 98.8%)	94.2% (78.7% - 98.7%)	94.7% (81.8% - 98.8%)	94.4% (83.0% - 98.7%)	95.0% (83.3% - 98.8%)	93.3% (77.8% - 98.3%)	94.7% (81.4% - 99.1%)	94.0% (79.3% - 98.7%)	91.4% (76.4% - 97.4%)	92.4% (84.0% - 98.9%)	95.1% (87.5% - 99.2%)	90.8% (73.6% - 97.4%)

Number of exons	6 (3 - 12)	6 (3 - 11)	6 (3 - 11)	6 (3 - 11)	5 (2 - 11)	7 (3 - 12)	6 (3 - 11)	6 (3 - 12)	5 (2 - 10)	7 (3 - 14)	9 (6 - 13)	7 (4 - 12)
Sequence length (codons)	380 (236 - 599)	378 (241 - 581)	378 (234 - 612)	354 (244 - 589)	328 (267 - 546)	364 (225 - 601)	380 (247 - 597)	379 (241 - 614)	373 (280 - 589)	417 (198 - 644)	484 (299 - 725)	431 (266 - 633)
Unspliced transcript length (bp)	16409 (3581 - 48194)	13763 (3482 - 40123)	16367 (3440 - 47710)	12772 (2231 - 42862)	11463 (1200 - 45516)	19279 (3993 - 55080)	21314 (3705 - 54981)	19025 (3546 - 55564)	6312 (1601 - 28200)	43309 (16713 - 87226)	41579 (25605 - 99229)	15510 (5894 - 54303)
G+C content at 4D sites	48.1% (36.8% - 58.2%)	50.0% (39.8% - 60.1%)	46.6% (34.3% - 59.5%)	48.4% (36.3% - 61.0%)	46.4% (38.5% - 58.4%)	48.9% (38.3% - 59.1%)	43.6% (33.4% - 55.7%)	46.2% (35.4% - 58.6%)	59.5% (49.7% - 67.7%)	37.1% (30.9% - 44.9%)	38.8% (30.6% - 46.7%)	59.6% (50.0% - 66.3%)

* Denotes values calculated for *Homo-Monodelphis* 1-to-1 orthologues that lie on a particular *Monodelphis* chromosome.

Supplemental Table 2

Paralogous clusters of gene families with functions related to reproduction in *Monodelphis* and that have experienced at least two lineage-specific duplications. Where orthology relationships to human genes have been predicted, the gene count and chromosomal location of the corresponding human orthologues, and the median d_S between orthologues are included.

Gene Count			**Chromosomes**		
Monodelphis	***Homo***	d_S	***Monodelphis***	***Homo***	**Description**
110			3, 4, 5, Un		Vomeronasal type-2 receptors
14			4,Un		Vomeronasal type-1 receptors
13			5		Vomeronasal type-1 receptors
12	1	3.0	1	10	beta-microseminoprotein, prostate secreted seminal plasma protein
11	1	2.4	2, Un	6	Vomeronasal type-1 receptors
8	1	1.2	4,8, Un	19	Vomeronasal type-1 receptors
6	5	3.0	1	9	lipocalin-1s, odorant-binding protein 2a/2b
5	2	1.0	2	6	Trace amine-associated receptors
5	1	2.6	1	9	Epididymal-specific lipocalin-9 precursor (mup-like lipocalin).
4			4		Vomeronasal type-1 receptors
3	1	2.2	1	9	MSFL2541-like lipocalin
3	1	1.0	2	6	Trace amine-associated receptor 9
3			X		CCR4-NOT transcription complex subunit 7 (CAF1), an essential factor for spermatogenesis

Supplemental Table 3

Paralogous gene clusters of olfactory receptor families in *Monodelphis* that have experienced at least two lineage-specific duplications. Where orthology relationships to human genes have been predicted, the gene count and chromosomal location of the corresponding human orthologues, and the median d_S between orthologues are included.

Gene Count			**Chromosomes**		
Monodelphis	***Homo***	d_S	***Monodelphis***	***Homo***	**Description**
30	1	2.1	5	11	Olfactory receptor 10AG1
21	4	1.1	5	11	Olfactory receptor 5B family
18	1	1.3	4	1	Olfactory receptor 5BF1
17			3		Olfactory receptor 7A family
15			3,5		Olfactory receptor 5AV1 family
13	3	1.1	4	11	Olfactory receptor 8G family
11	4	1.1	5	11	Olfactory receptor 5P family
11	2	1.1	8,Un	12	Olfactory receptor 6C family
11			2		Olfactory receptor 5AY1 family
8	2	1.6	Un,2	1	Olfactory receptor 10R2
8	1	2.5	3	19	Olfactory receptor 1I1
8			2,5		Olfactory receptor 4X / 4K family
7	4	1.6	4	11	Olfactory receptor 10G family
7	2	1.5	6	9	Olfactory receptor 13D1
7			1		Olfactory receptor 4C family
7			2		Olfactory receptor 2W5
7			5		Olfactory receptors 4S2 / 1S1
7			5		Olfactory receptor 10Q1
6	1	1.2	1	14	Olfactory receptor 11H4.
6	1	1.5	4,Un	11	Olfactory receptor OR11-293
6	1	1.1	7	3	Olfactory receptor 5AC2
6	1	1.2	5	11	Olfactory receptor Olfr1432-like
6	1	1.0	8	12	Olfactory receptor 6C76
6	1	2.5	1,4	10	Olfactory receptor 13A1
5	5	1.9	5	11	Olfactory receptor 5AK / 5T families
5	2	1.3	5	11	Olfactory receptor 5J2
5	1	1.0	4	11	Olfactory receptor 8B4
5	1	1.5	6	9	Olfactory receptor 13F1
5	1	1.3	6	9	Olfactory receptor 13C8
5			4,Un		Olfactory receptor 4C / 4N / 5AC families
5			6		Olfactory receptor 13F1 / 13C3 families
5			6,1		Olfactory receptor 13F1
5			5		Olfactory receptor 8K / 5T families
5			5		Olfactory receptor 4A4
4	5	1.5	2	6, c6_COX, c6_QBL	Olfactory receptor 10C1 / 12D2 families
4	4	1.1	5	11	Olfactory receptor 4D family
4	2	1.6	Un	11	Olfactory receptor 52E family

4	2	1.6	5	11	Olfactory receptor 4C family
4	1	1.4	1	14	Olfactory receptor 11G2
4	1	1.6	4	11	Olfactory receptor Olfr560-like
4	1	2.0	Un	11	Olfactory receptor 52E2
4			4		Olfactory receptor Olfr926-like
4			8,Un		Olfactory receptor 52J3
4			4		Olfactory receptor 12D3 / 10w1
4			1		Olfactory receptor 1F1 / 8G2
4			1,3		Olfactory receptor Olfr1377-like
4			1		Olfactory receptor 4F family
4			Un		Olfactory receptor 5D14 / 10K2 families
4			Un		Olfactory receptor 56A3
4			Un		Olfactory receptor 56A family
4			4		Olfactory receptors 5AN1 / 5C1 / 2AK2 families
4			5		Olfactory receptor 9I1
4			4		Olfactory receptor 8D2 / 8B4 families
3	5	1.8	3	19	Olfactory receptor 10H family
3	3	1.7	Un	11	Olfactory receptor 52A family
3	3	1.3	5	11	Olfactory receptor 8J family
3	2	1.2	Un	11	Olfactory receptor 56B4
3	2	1.3	1	11	Olfactory receptor 2AG family
3	1	1.1	8	7	Olfactory receptor 2F1
3	1	1.2	8	12	Olfactory receptor 9K2
3	1	1.5	4,Un	11	Olfactory receptor 51A7
3	1	1.3	1	11	Olfactory receptor 2D2
3	1	1.9	4	11	Olfactory receptor 6T1
3	1	1.3	6	16	Olfactory receptor 2C1
3	1	1.3	5	11	Olfactory receptor 5W2
3	1	1.8	5	11	Olfactory receptor 9G4
3	1	1.1	5	11	Olfactory receptor 9G1
3	1	1.4	5	11	Olfactory receptor 8U1
3	1	1.4	4	11	Olfactory receptor 51F1
3			1		Olfactory receptor Olfr49-like
3			5		Olfactory receptor 4A family
3			4		Olfactory receptor 51L1 / 2AP1
3			8		Olfactory receptor 6C family
3			Un		Olfactory receptor 51E1 / 4K15.
3			5		Olfactory receptor 4S family
3			5		Olfactory receptor 4C46
3			5		Olfactory receptor 6C76 / 5B3
3			5		Olfactory receptor 6C2
3			5		Olfactory receptor Olfr1009-like
3			8		Olfactory receptor 5M8 / 8D4.
3			1,Un		Olfactory receptor 4K15.
3			Un		Olfactory receptor 51L1.
3			4		Olfactory receptor 5AT1.
3			5		Olfactory receptor 6C6

Supplemental Table 4

Paralogous clusters of gene families with immunity-related functions in *Monodelphis* and that have experienced at least two lineage-specific duplications. Where orthology relationships to human genes have been predicted, the gene count and chromosomal location of the corresponding human orthologues, and the median d_S between orthologues are included.

Gene Count			**Chromosomes**		
Monodelphis	***Homo***	d_S	***Monodelphis***	***Homo***	**Description**
63	16	1.0	1,2,4,Un	19	Carcinoembryonic antigen-related cell adhesion molecule family
39	2	0.7	3,Un	22	Immunoglobulin domain variable regions
39			1,3,4,5,6,7,8, Un		Tripartite motif (TRIM) proteins
25			4,Un		Leukocyte immunoglobulin-like receptors
21	11	0.4	1,Un	14	Ig heavy chain V-III region
21	1	0.8	1	2	Ig kappa chain V region
19			1		T-cell receptor alpha chain C region.
16			4,Un		Leukocyte immunoglobulin-like receptors
11			1		Ig kappa chain V region
8	6	0.5	1	2	Ig kappa chain V region
8	4	1.1	1	2	Lithostathine 1 alpha/beta, Reg III-alpha
7	26	1.1	2	6, c6_COX, c6_QBL	HLA class I histocompatibility antigen, alpha chain
6	6	0.7	3	22	Ig lambda chain V regions
5	3	0.9	2	6, c6_COX, c6_QBL	Butyrophilin-like protein 2
5	1	1.3	8	7	T-cell receptor beta chain V regions
5	1	1.1	2	17	CCL4, small inducible cytokine A4
5			2,4		Carcinoembryonic antigen-related cell adhesion molecule family
4	15	1.3	6	9	Interferon alphas
4	2	0.9	2	1	Intelectins 1 & 2
4	1	1.5	8	12	Lysozymes C (1,4-beta-N-acetylmuramidases C)
4			1		HLA class I histocompatibility antigen chains
4			8		T-cell receptor beta chain C regions
4			3		T-cell receptor beta chain C regions
3	12	1.2	4	19	Sialic acid-binding Ig-like lectin (Siglec) family
3	4	1.1	2	1	Interferon-induced guanylate-binding proteins
3	3	1.4	Un	1	Obscurin, cytoskeletal calmodulin and titin-interacting RhoGEF
3	3	1.1	2	6, c6_COX, c6_QBL	MHC class II, DR alpha
3	2	1.7	1,2	5	Butyrophilin-like proteins 3 and 8
3	2	1.9	5	11	Eosinophil granule major basic protein / proteoglycan 3
3	1	1.1	4	11	GTPase, very large interferon inducible 1

					(GVIN1)
3	1	1.7	1	14	Ribonuclease pancreatic (RNase 1)
3	1	1.7	8	7	GTPase, Immunity-associated nucleotide 7 protein
3	1	2.1	1	2	Interleukin-1 family member 7 precursor (IL-1F7)
3	1	1.2	4,Un	19	Immunoglobulin alpha Fc receptor
3	1	1.4	3	19	EGF-like module-containing mucin-like hormone receptor-like 4 (G-protein coupled receptor 127)
3			3		Butyrophilin-like protein 2
3			1		Zinc-alpha-2-glycoproteins lipid mobilizing factors
3			2,8		GTPase, Immunity-associated protein 4
3			5		Membrane-spanning 4-domains subfamily A member 4A (CD20 antigen-like 1
3			3,Un		Ig kappa chain V-III region

Supplemental Table 5

Paralogous gene clusters of KRAB ZnF families in *Monodelphis* that have experienced at least two lineage-specific duplications. Where orthology relationships to human genes have been predicted, the gene count and chromosomal location of the corresponding human orthologues, and the median d_S between orthologues are included.

Gene Count			**Chromosomes**		**Exemplar gene identifiers in the *Monodelphis***
Monodelphis	***Homo***	d_S	***Monodelphis***	***Homo***	
119			1,2,3,4,7,8,Un		ENSMODG00000025186
117			1,2,3,4,5,7,8,Un		ENSMODG00000014123
23	21	1.4	4,Un	19	ENSMODG00000023659
15			2,4,8,Un		ENSMODG00000008473
10			1,3,4,Un		ENSMODG00000013749
9	1	0.9	1,2,4	2	ENSMODG00000019817
7			1,2,3,6		ENSMODG00000007417
7			3		ENSMODG00000000349
7			3, Un		ENSMODG00000007565
6			4, Un		ENSMODG00000023579
5			3,4,5, Un		ENSMODG00000013242
5			X, Un		ENSMODG00000025234
5			2, Un		ENSMODG00000008457
5			4		ENSMODG00000010852
5			1		ENSMODG00000019812
4			1		ENSMODG00000023049
4			4		ENSMODG00000014208
4			4		ENSMODG00000023509
4			3		ENSMODG00000023709
3			8,Un		ENSMODG00000004173
3			2,3		ENSMODG00000003057
3			3,6		ENSMODG00000024086
3			4,Un		ENSMODG00000022838

Supplemental Table 6

Paralogous clusters of gene families with detoxification-related functions in *Monodelphis* and that have experienced at least two lineage-specific duplications. Where orthology relationships to human genes have been predicted, the gene count and chromosomal location of the corresponding human orthologues, and the median d_S between orthologues are included.

Gene Count			**Chromosomes**		
Monodelphis	***Homo***	d_S	***Monodelphis***	***Homo***	**Description**
9			1		Cytochrome p450 2C18/2C9 (drug & steroid metabolism)
8	1	1.7	2	1	Cytochrome p450 2J family (drug & steroid metabolism)
6	5	1.9	3	19	Cytochrome p450 4F family (fatty acid metabolism)
6	4	1.1	6	7	Cytochrome p450 3A family (drug & steroid metabolism) CYP3A4 is involved in the oxidation of the largest range of substrates of all the CYPs
5	4	0.9	1	10	Cytochrome p450 2C family (drug & steroid metabolism)
4	1	2.9	7	X	Arylsulfatase E
4	1	0.7	7	2	Sulfotransferase 1C2
4	1	2.3	1,Un	4	Estrogen sulfotransferase EST-1
4	1	0.8	4,5	21	Carbonyl reductase (Prostaglandin-E(2) 9-reductase)
3	3	1.2	5	11	Glycine-n-acyltransferase-like 1
3	2	3.5	1	2	CML2 protein. N-acetyltransferase 8.
3	2	1.2	1	16	Liver carboxylesterase 1 precursor (egasyn).
3	1	1.5	1	9	Prostaglandin G/H synthase 1 (COX-1)
3	1	1.4	2	7	A/B hydrolase domain containing 11
3	1	1.8	4	2	Aldehyde oxidase
3	1	2.5	1	16	Carboxylesterase 2

Supplemental Methods

Augmenting Ensembl gene predictions in Monodelphis

Monodelphis transcripts were predicted from the *Monodelphis* genome assembly (version 3) using Exonerate (version 0.9;(Slater and Birney 2005)). Substantial sequence similarity (homology) to *Homo* protein-coding transcripts (Ensembl release version 36 based on NCBI assembly 35) provided the necessary evidence to initiate prediction of a transcript.

Transcript prediction proceeded in three stages. Firstly, putative gene locations were predicted using Exonerate (Slater and Birney 2005) in heuristic mode. Adjacent matches to a single human transcript were combined whereas overlapping matches to different transcripts were collated into priority lists sorted by the alignment score between template and predicted sequence. Secondly, transcripts were then predicted in descending order for each priority list. The prediction process was terminated if a transcript was predicted that matched the intron-exon structure of the template. For computing exon matches, locations of introns were allowed to lie up to 10 nucleotides away in aligned sequence; in addition, for multi-exon genes at most one missing exon was permitted. For each such successful prediction, genomic segments (150kb) up- and down-stream of the predicted transcript were checked for duplications. Finally, predictions were classified by their conservation of intron-exon structure, coverage of template and the presence of either in-frame stop-codons or frameshifts. Redundant transcripts and transcripts spanning the lengths of other transcripts were removed by applying a set of heuristic rules that preferentially retain full-length transcripts possessing conserved gene structure (as above) with no disruptions. Multiple predictions of equivalent quality were all retained. Transcripts with overlapping exons were finally collated into single genes using a connected components algorithm. We discarded all gene predictions that overlapped an existing Ensembl gene.

This additional gene set was found to include 467 predicted transcripts from 367 genes that did not overlap with any Ensembl exon, and that possessed a human orthologue, as predicted by PhyOP (see below). The additional gene set included predictions with frame-shifts and in-frame stop codons only if these were located in exons of low sequence similarity, but the set excluded single exon transcripts and transcripts without conserved gene structure. 27 transcripts containing exons from apparently distinct, but neighbouring, gene loci were removed manually.

Predicting additional Monodelphis paralogues

We sought to predict previously unidentified *Monodelphis* genes from 701 gene families that contained *Monodelphis*-specific gene duplicates. These had been identified using the PhyOP orthology and paralogy prediction pipeline on the augmented Ensembl gene set.

We used both *Homo* and *Monodelphis* sequences from these duplicated gene families as templates and generated gene predictions using Exonerate, as described above. Predictions of fewer than 30 codons and those overlapping ones from Ensembl or our previous predictions were discarded. This resulted in the prediction of 330 additional transcripts from 290 genes. Subsequently, we derived the final set of orthology and paralogy relationships using PhyOP, including genes from Ensembl and all our additional *Monodelphis* predictions.

Supplemental Figure 1

G+C content along the Monodelphis *autosomes MDO1-8*

This plot shows the clear increase in G+C content at the telomeric ends of the *Monodelphis* autosomes. (These include both ends of the metacentric MDO1 and MDO2, and the distal ends of the long arms of MDO3-8). The thin blue lines show the G+C content for adjacent 50 kbp windows while the thick red line is a running average of 50 such windows. The average G+C content for each chromosome is shown by the horizontal green line.