Appendix 1: List of non-standard primers used

Construction of mouse OR gene vectors for *in situ* hybridisation [2.8]

MP1A	GTCTCATCCTGCCAACTAGACA	59.36	>
MP1.2.5B	TATGGATTGGACCAAACCTATA	56.11	<
MP1.2.5C	TATAGGTTTGGTCCAATCCATA	56.11	>
MP1D	CAAATTCATCTTAATTTGGAAA	53.88	<
MP2A	GTCTCATTCCTCCAACTAGATA	52.76	>
MP2D	AAGAAACTTTGTAATTCAGAAA	50.61	<
MP3A	GTCTCATCCCTCCAACTAGATA	55	>
MP3B	TACTGACTCCACTAGGCCAATC	58.34	<
MP3C	GATTGGCCTAGTGGAGTCAGTA	58.34	>
MP3D	TGCAGATTTCCCATTCAGCTTA	61.94	<
MP4A	CTTCCAGTTGGATTCATAGCAG	58.86	>
MP4B	TATGGATTGAAGCAGTCCCATC	61.2	<
MP4C	GATGGGACTGCTTCAATCCATA	61.2	>
MP4D	ATCTGGCCTTGAACTTACAGCA	61.16	<
MP5A	GTCTCATCCTGCCGACTAGACA	61.74	>
MP5D	TTCTAATTCATCTTAATTTGGA	51.28	<
MP6A	GGTTAGCTAGCGATAGACTTAG	52.48	>
MP6B	TGGCACCACACTAGTGGTGAAA	63.77	<
MP6C	TTTCACCACTAGTGTGGTGCCA	63.77	>
MP6D	AGAGAAAAGAATCTCGCTACTA	52.31	<
MP7A	TATTGCCCATAAGACTTATGGA	56.39	>
MP7B	ACGTCAATATGGAGTGAACCAG	58.98	<
MP7C	CTGGTTCACTCCATATTGACGT	58.98	>
MP7D	ACATTGATTTTCACAATAACTG	52.06	<
MP8A	CATGAAATTATTGCCTAGTAA	50.15	>
MP8B	ATTGAGACGTCAATATAGAGTG	51.25	<
MP8C	CACTCTATATTGACGTCTCAAT	51.25	>
MP8D	ATCATAAACAATTGCAGAATC	51.99	<
MP9A	TATTGACTTTAAAGATGATGGA	51.86	>
MP9B	ATTGAGATGTCAATATGGAGTG	54.17	<
MP9C	CACTCCATATTGACATCTCAAT	54.17	>
MP9D	ATCATACACAATTGCAGAATC	52.64	<
MP10A	GCTATCATAATTTCCATTTCTC	52.56	>
MP10B	CTCACTGAGGTGAGTGTCTGAG	57.64	<
MP10C	CTCAGACACTCACCTCAGTGAG	57.64	>
MP10D	ATGAGCCATGGTATTAACTGAC	55.78	<
MP11A	GCGTTCTAGTTTCTATTACTCA	51.14	>
MP11B	CTCATTGAGGTGAGTGTCAGAA	57.93	<
MP11C	TTCTGACACTCACCTCAATGAG	57.93	>
MP11D	CTGAGCAACAGTGTTAACTGAC	55.22	<
MP12.13B	ACAGAAGAAATGATGGACATGA	57.11	<
MP12.13C	TCATGTCCATCATTTCTTCTGT	57.11	>
MP12A	GTCAACATGCAGGGTAGTATT	54.27	>

MP12D	CTGGTTATTTGCACTGAAGAAG	56.76	<
MP13A	GTTAACATGCAGGGTAGTATT	51.22	>
MP13D	CTGGTTATTTACACCAAAGCAG	56.61	<
MP14A	CTGGTTGCATTCGACAAACTAT	59.16	>
MP14B	ACAGAAGAAATGATGAATATGA	51.54	<
MP14C	TCATATTCATCATTTCTTCTGT	51.54	>
MP14D	CTGGTTATTTACACCTCAGTAG	51.1	<
MP17A	TAGATAGTGTGAGGTCTATGCA	52.7	>
MP17B	TACACTTTCTGGGACTCATGAT	56.26	<
MP17C	ATCATGAGTCCCAGAAAGTGTA	56.26	>
MP17D	AAACCAATACATTGTCTTTCAT	53.04	<
MP18A	TTGATAGTGTGAGATGTATGGA	53.16	>
MP18.24B	TGCACTTTCTGGGACTCATGAT	61.97	<
MP18.24C	ATCATGAGTCCCAGAAAGTGCA	61.97	>
MP18D	ACATTAGACTTTGAAGTCATCC	52.67	<
MP19A	GACTGGATAGCTTGAGATCTAA	53.09	>
MP19B	TACACTTTCTGGGACACATGAT	57.05	<
MP19C	ATCATGTGTCCCAGAAAGTGTA	57.05	>
MP19D	ACATTAGAAATTGAAGGAATCT	51.64	<
MP20A	TGGATAGCTTGTGATCTATGTA	52.69	>
MP20B	TGATGACTTCATAGTGAAGTGG	55.32	<
MP20C	CCACTTCACTATGAAGTCATCA	55.32	>
MP20D	ACATTAGAGTTTGAAGGAAACC	54.3	<
MP21A	CACCTTTAGGGTCCATGTCTCT	59.5	>
MP21B	GATGAAGTGATAAATTTTGTGG	53.97	<
MP21C	CCACAAAATTTATCACTTCATC	53.97	>
MP21D	CATGTGGATCAACACCTTCATT	59.72	<
MP22A	ACATGGTAAGTCTTCTGCCTGA	58.86	>
MP22B	CATTCGGGTTGAGTAATGGAGT	60.24	<
MP22C	ACTCCATTACTCAACCCGAATG	60.24	>
MP22D	GGCAGACAAAAGAATGTACCAG	58.77	<
MP23A	CATTATTGGTACTAGCACTTGC	54.44	>
MP23B	GATTGTTACCACAGAAGGGCAG	60.9	<
MP23C	CTGCCCTTCTGTGGTAACAATC	60.9	>
MP23D	TTCTAATGTAAGATCATTGACC	51.23	<
MP24A	TATGAGGAAATAGTCTGAGATG	51.3	>
MP24D	ACATTACAGTTTGAAGGAATCC	54.93	<

Construction of ‘olfactory promoter region’ pGL3 luciferase reporter vectors [2.9]

OLFOP-F	CCTCGAGATCTTCCCTTACCTTTGCATGG	59.56	>
OLFOP-R	GATAGATCTTTCTATGGGCCAGCAATTC	60.04	<

Polymorphism analysis [2.11]

17C	TTGTCTTTCTGACAGGCTGG	59.01	>
17D	AGGGAGATCTAGTGCTGCGA	60.12	<
17E	CATAGAAGAGGGAGACCACGAT	59.6	>
17F	GTATCCTCGTTACCATCTTCCG	59.86	<
17G	AGAAGAACATCTGGAGAGCACA	59.09	>
17H	GATTGGCTATACGTCTGTCACG	59.66	<

20C	TTTTTCTTTTATCCAGTTGCCTC	58.85	>
20D	GGTGAGAAAATTCTGAGCCG	59.81	<
20E	TGATTAAGCTCAGTGTTCCCAC	59.24	>
20F	CCCCTTCTTTCTGACACTTCTC	59.36	<
20G	CAGATAGCCACAGAGAGGTCAA	59.5	>
20H	GGAAACCTGTCCTACCTGGATA	59.35	<

22C	CTATTGGGATTTTCTGACCGTC	59.84	>
22D	TATGCATTCAGTAGAACCCAGG	59.11	<
22E	GGTTGTGTTGCCCAACTCTATA	59.02	>
22F	TTCACAGAAGAAATGGTCCAAC	59.08	<
22G	TTGTCCAGTCCACTCTCACAGT	59.8	>
22H	ACAGGTATTGAAAGCCTTCCAG	59.65	<
22J	TCAGGCAGTAATGAGAATCTGC	59.48	>
22K	ACTGTTCCCTCATCAATTCCAT	59.7	<
22L	CCATGCATGTAGAGTTGATGGT	59.88	>
22M	ACCTCACGATAGTACTGGCATG	59.15	<
22N	CATTGATGATGTTGTCTCCACA	59.42	>
22P	ACCTTACAAAAGCATTTGAGGG	59.54	<
22Q	CCAAAGCATCTCACTGTTCATC	59.74	<
22R	CCACTGAGTTTGTGTCCCTAAA	59.14	<
22S	GCACATCCTCCTGTAGTCTTCA	59.36	>
22T	CTGTTGCCACCATTACATATCC	59.23	>

Mouse RT-PCR [2.12]

1a	CCTGACTTCCTCTCAGCCAC	59.99	>
1b	GGAGCACAAGACAACGACAA	59.88	<
2a	CTGAGAAGCACAAAAAGGGC	59.99	>
2b	TTTCCAGTTGTGGGTGTTCA	59.98	<
3a	CACCACACTCCTGGTTTCCT	60	>
3b	TTTCCAGTTGTGGGTGTTCA	59.98	<
4a	GGATGCACCTGTTCCTTGTT	59.97	>
4b	TCCAGTTGTGGGTGTTCAGA	60.13	<
6a	CAACGCCCACTTTTAAGGAA	60.1	>
6b	CCAGGTCTGTGTTCCTGGTT	60	<

Human RNA dot-blot hybridisations [2.13]

15A	GCAAGTAATTCCTAGGTCATGGA	59.51	>
15B	TGATCAGGGACATAAAGGAGC	59.14	<
16A	GCGAGCACTCAGGAGGTTAC	60.02	>
16B	CAGGTGTATGTGGGAAGGGT	59.7	<
20A	TAGCCATGGAACCCTAATGC	59.92	>
20B	CATGGGCAATATAGGCAAAA	58.51	<

PCR amplification of olfactory epithelium phage library [2.15]

16EST1	ACGAGGTTTCACCATGTTGA	>	52.4
16EST2a	ATCAGAAGGAACAGGGAACGA	>	55.07
16EST2b	TTTTCTGTCTCTGCTCACCCA	<	54.2
16EX3	TCAGACTCTCTTCACGGCCTT	>	55.14
16EX4a	TGTTTGTACCTGGAGGGATGA	<	54.6
16EX4b	AAGTCAGAGGCACCAATGTGA	>	54.01
16C	CAGGTGTATGTGGGAAGGGT	>	59.70
16D	CACCCAGGCTGAGTTTTGAT	<	60.11
16E	TAGAAGAGGGTGACCACAGTGA	>	59.76
16F	AGGCAGTGCTGAGGATTAACTC	<	59.91
16G	ACAGGAAGATGAAGAGCTGGAC	>	59.88
16H	CAAAGAAGACCATCAGCTTCCT	<	59.89

Appendix 2: Sequence feature files and output of ‘olfgrab’

“.genes” file

> *(or)* < 230 7898 genename CDS *(or)* mRNA

(tab) *(tab)* *(tab)* *(tab)*

230 560

(tab) *(tab)*

6560 6989 (*exons)*

(tab) *(tab)*

7854 7898

(tab) *(tab)*

etc. for all genes present in the clone.

The >/< determines whether the gene is present on the forward or reverse strand, whilst the gene name should ideally consist of one word (although - , _ ,* or numbers are acceptable).

“.rptm” file

68 145 repeatname +*(or)*C repeattype

(tab) *(tab)* *(tab)* *(tab)*

575 4260 repeatname +*(or)*C repeattype

(tab) *(tab)* *(tab)* *(tab)*

etc. for all repeats present in the clone.

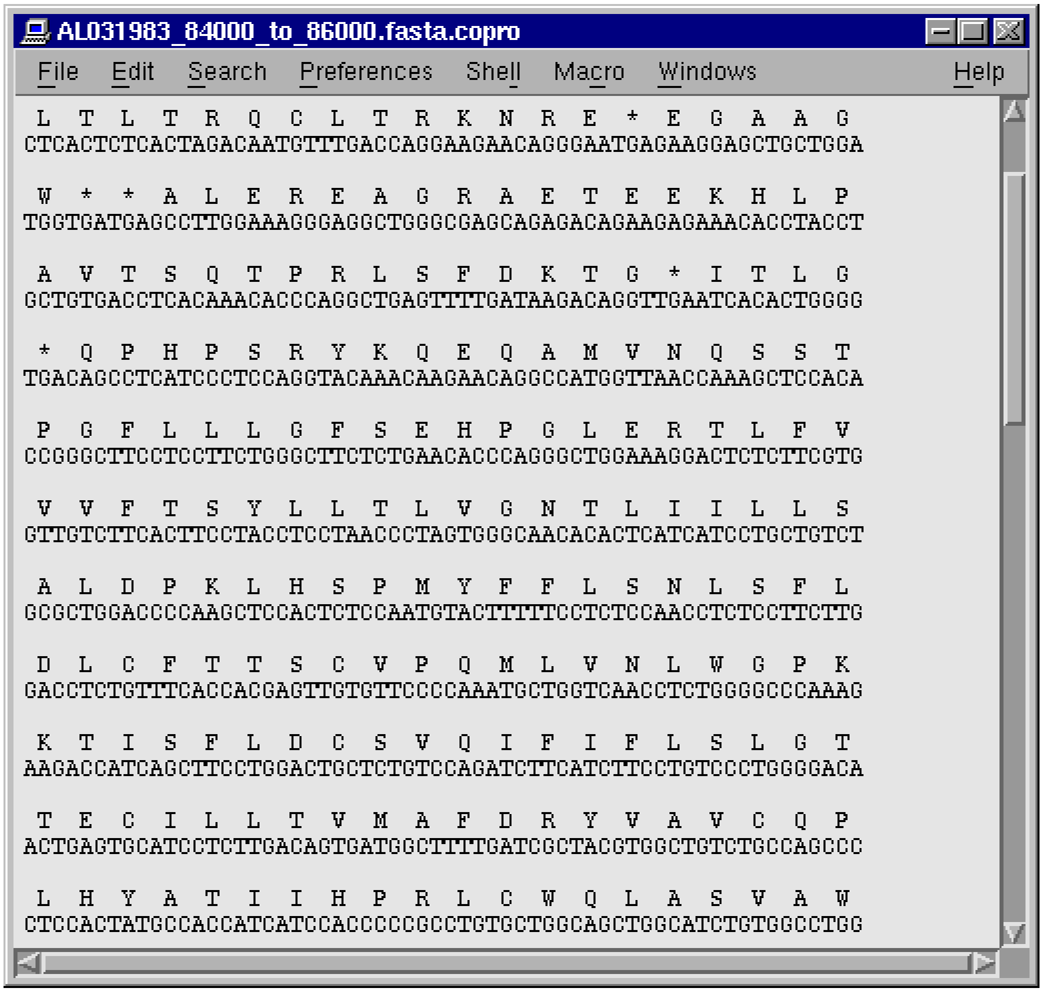
The +/C states whether the repeat match is on the forward or complement strand.

“.misf” file

12235 12325 CpG

(tab) *(tab)*

etc. for all CpG islands/other features in the clone.



```
 L  T  L  T  R  Q  C  L  T  R  K  N  R  E  *  E  G  A  A  G
CTCACTCTCACTAGACAATGTTTGACCAGGAAGAACAGGGAATGAGAAGGAGCTGCTGGA

 W  *  *  A  L  E  R  E  A  G  R  A  E  T  E  E  K  H  L  P
TGGTGATGAGCCTTGGAAAGGGAGGCTGGGCGAGCAGAGACAGAAGAGAAACACCTACCT

 A  V  T  S  Q  T  P  R  L  S  F  D  K  T  G  *  I  T  L  G
GCTGTGACCTCACAAACACCCAGGCTGAGTTTTGATAAGACAGGTTGAATCACACTGGGG

 *  Q  P  H  P  S  R  Y  K  Q  E  Q  A  M  V  N  Q  S  S  T
TGACAGCCTCATCCCTCCAGGTACAAACAAGAACAGGCCATGGTTAACCAAAGCTCCACA

 P  G  F  L  L  L  G  F  S  E  H  P  G  L  E  R  T  L  F  V
CCGGGCTTCCTCCTTCTGGGCTTCTCTGAACACCCAGGGCTGGAAAGGACTCTCTTCGTG

 V  V  F  T  S  Y  L  L  T  L  V  G  N  T  L  I  I  L  L  S
GTTGTCTTCACTTCCTACCTCCTAACCCTAGTGGGCAACACACTCATCATCCTGCTGTCT

 A  L  D  P  K  L  H  S  P  M  Y  F  F  L  S  N  L  S  F  L
GCGCTGGACCCCAAGCTCCACTCTCCAATGTACTTTTTCCTCTCCAACCTCTCCTTCTTG

 D  L  C  F  T  T  S  C  V  P  Q  M  L  V  N  L  W  G  P  K
GACCTCTGTTTCACCACGAGTTGTGTTCCCCAAATGCTGGTCAACCTCTGGGGCCCAAAG

 K  T  I  S  F  L  D  C  S  V  Q  I  F  I  F  L  S  L  G  T
AAGACCATCAGCTTCCTGGACTGCTCTGTCCAGATCTTCATCTTCCTGTCCCTGGGGACA

 T  E  C  I  L  L  T  V  M  A  F  D  R  Y  V  A  V  C  Q  P
ACTGAGTGCATCCTCTTGACAGTGATGGCTTTTGATCGCTACGTGGCTGTCTGCCAGCCC

 L  H  Y  A  T  I  I  H  P  R  L  C  W  Q  L  A  S  V  A  W
CTCCACTATGCCACCATCATCCACCCCCGCCTGTGCTGGCAGCTGGCATCTGTGGCCTGG
```

File created by the ‘olfgrab’ program. Key OR motifs can be identified (for example, MYFFL, MAFDRY), and the file can be edited and then run through the ‘olfproducer’ program to produce nucleotide and protein files.

Appendix 3: List of websites

HGMP	*http://www.hgmp.mrc.ac.uk*
RepeatMasker	*http://ftp.genome.washington.edu/cgi-bin/RepeatMasker*
Repbase	*http://www.geospiza.com/products/tools/repbase.htm*
GRAIL1	*http://compbio.ornl.gov/Grail-1.3*
Genfinder	*http://argon.cshl.org/genefinder/human.htm*
Genemark	*http://www.ebi.ac.uk/genemark/*
Fex	*http://genomic.sanger.ac.uk/gf/gf.shtml*
Fgene	*http://genomic.sanger.ac.uk/gf/gf.html*
Genscan	*http://genes.mit.edu/GENSCAN.html*
EBI	*http://www.ebi.ac.uk*
NCBI	*http://www.ncbi.nlm.nih.gov*
Pfam	*http://www.sanger.ac.uk/Software/Pfam/search.shtml*
BLOCKS	*http://blocks.fhcrc.org/blocks/blocks_search.html*
PRINTS	*http://bioinf.man.ac.uk/cgi-bin/dbbrowser/fingerPRINTScan/muppet/FPScan.cgi*
PROSITE	*http://www.expasy.ch/prosite/*
Psort	*http://psort.nibb.ac.jp/*
DSC	*http://www.hgmp.mrc.ac.uk/Registered/Option/dsc.html*
Simpa96	*http://npsa-pbil.ibcp.fr/cgibin/npsa_automat.pl?page=/NPSA/npsa_simpa96.html*
Phd	*http://www.public.iastate.edu/~pedro/pprotein_query.html*
Predator	*http://www.embl-heidelberg.de/cgi/predator_serv.pl*
COILS	*http://dot.imgen.bcm.tmc.edu:9331/seq-search/struc-predict.html*
Tmpred	*http://www.ch.embnet.org/software/TMPRED_form.html*
Tmap	*http://bioweb.pasteur.fr/seqanal/interfaces/tmap.html*
DAS	*http://www.sbc.su.se/~miklos/DAS/*

HTH	*http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hth.html*
Signal	*http://www.cbs.dtu.dk/services/SignalP/*
Sigcleave	*http://bioweb.pasteur.fr/seqanal/interfaces/sigcleave.html*
Antigenic	*http://bioweb.pasteur.fr/seqanal/interfaces/antigenic.html*
Digest	*http://www2.no.embnet.org/Pise/digest-simple.html*
Dotter	*http://www.cgr.ki.se/cgr/groups/sonnhammer/Dotter.html*
Ensembl	*http://www.ensembl.org*
UCSC	*http://genome.ucsc.edu*
PipMaker	*http://bio.cse.psu.edu/pipmaker*
Promoter Inspector	*http://www.genomatix.de/cgi-bin/promoterinspector/promoterinspector.pl*
TRANSFAC	*http://transfac.gbf.de/TRANSFAC/*

Appendix 4: List of scripts/programs

Script	Author	Description	Input	Additional information
AAbreak	rmy	Pulls out co-ordinates of olfactory domains, and also pulls out position of stops (*) and frameshifts (X)	1 protein file	Asks "Domain Information only ?", N/n -will show entire protein with amino acids numbered.
accmatch	rmy	Takes 2 lists of accession numbers and compares them to see if there are any numbers that match.	2 lists of accession numbers.	
addclone	rmy	Adds 2 clones together to produce new ".genes", ".rptm" and ".misf" files	2 ".genes" files 2 ".rptm" files 2 ".misf" files.	Asks for lengths of clones (defaults will be taken from EMBL) and for the overlap of the clone(default=100)
addclonegenes	rmy	Adds certain length to ".genes" file, produces ".ADD" file.	1 ".genes" file	Asks for length to be added to file.
addclonemisf	rmy	Adds certain length to ".misf" file, produces ".ADD" file.	1 ".misf" file	Asks for length to be added to file.
addclonerptm	rmy	Adds certain length to ".rptm" file, produces ".ADD" file.	1 ".rptm" file	Asks for length to be added to file.
blastcut	rmy	Given a blast results file and an accession number, chops out the first result associated with the accession number.	BLAST output file, accession number.	Output file: ".blast"
blastextract	rmy	Produces a list of BLAST hits, showing accession number and start and stop positions.	BLAST output file	Writes to the screen.
changename	rmy	Changes the name within an OR file to a new name.	Name of OR file, new name.	Produces a .bak file showing old contents of file.
chopseq	rmy	Takes a piece of sequence and produces files containing sequence of a specified length.	1 FASTA format file	Produces file.001, file.002 etc.
chopuprptm	rmy	Pulls out repeats from a given region in a ".rptm" file.	1 ".rptm" file, a start co-ordinate an a stop co-ordinate.	File is named .rptm_start_to_stop

clonebreak	rmy	Takes a sequence file and produces files of each piece of sequence divided from the other by 5 or more 'n''s	1 unfinished sequence file (containing 'n''s)	Files are given increasing numbers as they are produced.
cloneends	rmy	Given a sequence file pulls out the first and last 100 bases and puts these in ".start" and ".end" files.	1 finshed fasta file.	
count	rog	Counts the number of bases in a sequence file, and calculates GC content.	1 sequence file	
eponinehits	rmy	Takes hits from eponine output file and converts to ".misf" format.	1 eponine output file	
estlocate	rmy	Takes an EST accession number and BLAST searches available genomic databases.	1 EST accession number	
fembl	rog	Takes accession number and produces file of fasta sequence of accession number	1 accession number	Output file: ".fasta"
gc_table	rog	Calculates gc content across a specified number of bases in a fasta file.	1 sequence file	Prompts for length of bases over which content is to be calculated. Produces "gc_out" file.
getgeneSEQ	rmy	Gets sequence for a gene specified in a gene file. (Name of file corresponds to EMBL sequence file)	1 EMBL-based ".genes" file	
getspecies getspecieslots	rmy	Gets species an accession number is derived from.	1 accession number List of accession numbers	
hembl	rog	Takes an accession number and produces file showing descriptive part of EMBL file.	1 accession number	Output file: ".head"
listlook	rmy	Takes a file and pulls out a list of accession numbers featured in file.	File containing accession numbers.	
localareagrab	rmy	Takes list of accession numbers and stop and start positions and grabs sequence -500 to +500 upstream and downstream of these positions.	List of files containing accession numbers and start and stop positions.	Produces a number of ".geneseq" files.

olfcheck	rmy	Takes protein file, strips it of '*' and 'X' characters and compares edited protein against other protein files in the directory (using the Emboss program 'water'). Reports hits over 1500.	1 protein file (Other proteins in directory named ".pro")	Output file: ".hits"
olfchromosome	rmy	Takes accession number and looks in EMBL file to see if chromosome is attached to accession number.	1 accession number	
olfclonesift	rmy	Takes a list of accession numbers and compares list of OR/clones, pulling out ORs on particular clones.	List of accession numbers List of OR/clones (on same line).	
olfdba	rmy	Runs 'DBA' program on list of sequences.	List of genes (with sequences contained in files), and required match for 'DBA' program.	Produces a list.dba file. Match for DBA: MatchA 0.65 MatchB 0.75 MatchC 0.85 MatchD 0.95
olfdba2draw	rmy	Plots graphical representation of 'DBA' results.	1 ".dba" file, 1 ".rptm" file.	Produces postscript file as output.
olfgrab	rmy	Uses 'wombl' or 'pullout' to get a sequence, translates this in 6 frames and BLAST searches against the OR database. The highest hit(s) are then placed in a ".copro" file, created using the 'translator' program.	1 sequence file or accession number, start and stop positions.	Produces ".copro" file which can be edited and used in 'olfproducer' program.
olflocate	rmy	When given an OR accession number, prints where the number is in the genome (if known).	1 accession number	Location manually programmed in using information from Ensembl database.
or_nip	rmy	Looks in a fasta file for a sequence pattern specified in a PATTERNS/file.pat file.	1 fasta file, 1 pattern file (in correct directory)	Uses the 'NIP' program on the command line
olfproducer	rmy	From a ".copro" an OR protein and nucleotide file is produced.	1 ".copro" file	Prompts for name of OR and any additional comments.
PIPlocaldraw	rmy	Plots a graphical represenation of how 2 sequences are related to	1 edited PIP output file, position of start of region, position of end of region.	Produces a postscript file.

		each other, based on a PIP file.		
proteingrab	rmy	Creates a protein file of an accession number (if file found in protein databases)	1 EMBL accession number.	
pullout pulloutmany	rmy	Given a sequence file and start and stop positions pulls out the piece of sequence between these positions. Given a list start ands stop positions uses pullout to get required sequence	1 sequence file, start and stop positions. 1 sequence file, list file containing start and stop positions.	Produces a sequencefile_start_to_stop.fasta file.
rcin	rmy	Reverses a piece of sequence inputted on the command line	1 sequence on the command line	
reformatXLSclustal	rmy	Reformats to a clustalw fasta format file from an Excel format (".tab", generated by proalnXLS)	1 ".tab" file, created by proalnXLS.	
repembl	rmy	Calculates percentage type repeats from a EMBL file.	1 EMBL accession number	
relateolf	rmy	Takes OR and BLAST searches against the OR database, finding the top 5 hits, and the top hit on chromosome 6.	1 protein file	
reverseclone	rmy	Reverses ".genes", ".misf" and ".rptm" files.	1 filename (".genes", ".misf" and ".rptm" files corresponding to this name). Length of clone	If the size of the clone is not entered the length is given as the size in EMBL.
rolfp	rmy	Taking a protein file as input, BLAST searches against the OR database.	1 protein file	Produces a ".pblast" output file
rmpars	rog	Calculates percentage type repeats from a RepeatMasker output file.	1 RepeatMasker output file.	
rptcalc	rmy	Takes a ".rptm" file and calculates the percentage of repeats belonging to each class.	1 ".rptm" file	
rptmgenerate	rmy	Takes a RepeatMasker out file and produces a correctly formatted ".rptm" file.	File from repeatmasker containing co-ordinates and type of repeat.	
seeclone	rmy	Produces ".genes", ".rptm" and ".misf" files from an EMBL accession number, also produces ".ps" postscript file.	1 accession number	
SNPdefine	rmy	Takes gene file and list of SNPs and produces output	1 ".genes" file List of SNP position	Output file: list ".assn", SNP ,

		file showing whether a SNP is found in an exon		blank if SNP does not appear in exon
translator	rmy	Takes a DNA sequence file and produces an output file showing protein translation above (forward) and below (reverse) the DNA.	(i) 1 DNA file (ii) Optional – R or r will only plot reverse strand and reverse translation, F or f will only plot forward strand and forward translation	
transfachits	rmy	Takes hits from transfac output file and converts to ".misf" format.	1 transfac output file	
waterall	rmy	Uses the Emboss program 'water' to compare one sequence against a number of others	(i) 1 sequence file (ii) file containing a list of sequences to be compared against.	
wombl	rog	Takes an accession number and 2 positions, and pulls out sequence corresponding to that between these positions.	1 accession number 2 positions found within the accession number.	Output file: "file_pos1_to_pos2 .fasta"

Appendix 5: List of clones assembled to produce MHC extended class I region

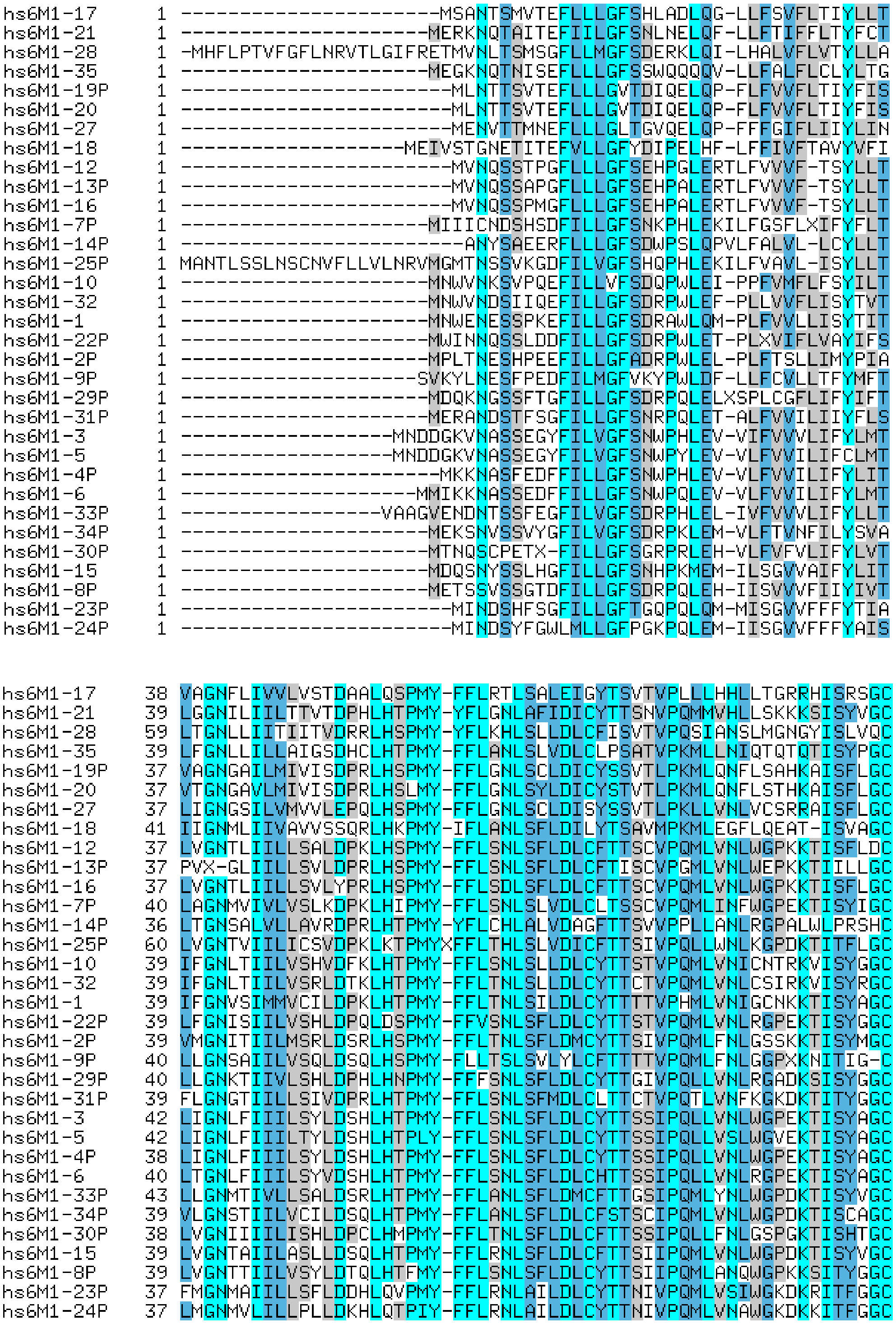
Accession number	Size	Position of start in consensus sequence	Overlap with neighbouring clone	Orientation
U91328	246282	1	-	rev
AL353759	101099	198877	47405	for
AL031777	89301	299877	100	for
AL021917	170001	389078	100	for
AL050330	9015	558979	100	for
AL121936	122979	567894	100	for
AL513348	109476	690773	100	for
AL591044	133561	800248	2000	for
AL596216	432	931809	255	for
AL133255	152782	932058	100	for
AL590062	65119	1084585	100	for
AL021807	89016	1149604	100	for
AL121934	49261	1238520	100	for
AL021808	154066	1287681	102	for
AL031118	82456	1441647	101	for
AL021918	159506	1524002	100	for
AL031229	61450	1683407	100	for
AL009179	139904	1744757	100	for
AL049822	24706	1884561	100	for
Z98744	100375	1909167	101	for
AL133267	44788	2009442	100	for
AL121944	130279	2054130	100	for
AL358933	55480	2184309	101	for
AL022393	85654	2239689	100	rev
AL390721	13994	2325243	101	for
AL021997	97847	2339137	100	rev
AL358785	17464	2436884	100	for
Z98745	128779	2454248	100	for
AL049543	95594	2582927	100	rev
AL121932	85952	2678421	100	for
AL390196	43042	2764273	101	for
AL133258	73666	2807215	100	for
Z84474	107527	2880781	101	for
AL139329	3812	2988207	100	for
Z84476	112659	2991918	100	for
AL035402	47216	3104477	100	for
AL022727	144868	3151593	100	for
AL050339	68105	3296361	100	for
AL096770	97392	3364366	100	for
AL035542	114868	3461658	100	for
AL031983	134292	3576426	100	for
AL050328	54106	3710618	100	for
AL022723	148834	3764624	100	for

Appendix 6: List of human MHC-linked ORs

Name	Clone		Position in clone	State	Length	Additional information
hs6M1-1	AL022727	<	7626..8567	C	942	
hs6M1-2	AL022727	<	59199..60135	P	937	1 bp insertion at 430 > frameshift
hs6M1-3	AL022727	>	33215..34150	C	936	
hs6M1-4	AL022727	>	22275..23207	P	933	Substitution at 574,575 or 576 > stop codon
hs6M1-5	AL022727	>	102837..103772	P	936	Substitution at 565,566 or 567 > stop codon
hs6M1-6	AL022727	>	94963..95901	C	939	
hs6M1-7	AL022727	>	136590..137533	P	944	1 bp deletion at 95 > frameshift
hs6M1-8	Z84476	>	4055..4982	P	928	1 bp deletion at 810 > frameshift
hs6M1-9	Z98745	<	74312..74816	PF		No starting methionine Lacking 2 end motifs 1bp deletion at 263 > frameshift Substitution at 282,283 or 284 > stop codon
hs6M1-10	Z98744	<	74988..76061	C	1074	
hs6M1-12	AL031983	>	84339..85277	C	939	
hs6M1-13	AL031983			P		1 bp deletion at 114 > frameshift 1 bp insertion at 618 > frameshift 1 bp insertion at 679 > frameshift
hs6M1-14	AL031983			P	942	No starting methionine
hs6M1-15	AL035402	<	12652..13614	C	963	
hs6M1-16	AL035542			C	951	
hs6M1-17	AL035542	>	51151..52089	C	939	
hs6M1-18	AL035542	<	37829..38776	C	948	
hs6M1-19	AL035542	>	28417..29363	P	947	16 bp deletion at 552 > frameshift
hs6M1-20	AL035542	>	7847..8770	C	924	
hs6M1-21	AL096770	>	32625..33590	C	966	
hs6M1-22	AL096770 AL050339	< <	97167..97392 1..817	P	943	1 bp insertion at 82 bp > frameshift
hs6M1-23	AL050339	>	22125..23084	P	960	Substitution at 550,551 or 552 > stop codon Substitution at 553, 554 or 555 > stop codon
hs6M1-24	AL050339	> >	27467..28053 28552..28912	P	956	1 LTR19A insertion at 587 1 bp deletion at 871 > frameshift
hs6M1-25	AL050339	>	61300..62293	P	994	1 bp deletion at 244 > frameshift

						1 bp deletion at 599 > frameshift
hs6M1-26	AL035402	<	40196..40877	P	682	Missing 1st 2 motifs 1 bp insertion at 160 > frameshift Substitution at 290, 291 or 292 > stop codon Substitution at 305, 306 or 307 > stop codon
hs6M1-27	AL096770	>	13518..14468	C	951	
hs6M1-28	AL096770	<	81166..82197	C	1032	3 potential starting methionines
hs6M1-29	AL121944	<	72007..72944	P	938	1 bp deletion at 76 > frameshift
hs6M1-30	AL121944	>	52652..53634	P	943	1 bp insertion at 28 > frameshift 1 bp deletion at 316 > frameshift 1 bp insertion at 448 > frameshift Substitution at 782,783 or 784 > stop codon
hs6M1-31	AL121944	>	65213..66146	P	982	1 bp insertion at 322 > frameshift
hs6M1-32	AL133267	>	20708..21649	C	942	
hs6M1-33	AL133267	>	920..1870	P	951	Substitution at 1 > Valine not methionine start
hs6M1-34	AL133267	>	40617..41542	P	926	1 bp deletion at 539 > frameshift
hs6M1-35	AL121944	>	92095..93108	C	1014	

Appendix 7: Alignment of human MHC-linked ORs



```
hs6M1-17     1 ----------------------MSANTSMVTEFLLLGFSHLADLQG-LLFSVFLTIYLLT
hs6M1-21     1 ---------------------MERKNQTAITEFIILGFSNLNELQF-LLFTIFFLTYFCT
hs6M1-28     1 -MHFLPTVFGFLNRVTLGIFRETMVNLTSMSGFLLMGFSDERKLQI-LHALVFLVTYLLA
hs6M1-35     1 ---------------------MEGKNQTNISEFLLLGFSSWQQQQV-LLFALFLCLYLTG
hs6M1-19P    1 -----------------------MLNTTSVTEFLLLGVTDIQELQP-FLFVVFLTIYFIS
hs6M1-20     1 -----------------------MLNTTSVTEFLLLGVTDIQELQP-FLFVVFLTIYFIS
hs6M1-27     1 -----------------------MENVTTMNEFLLLGLTGVQELQP-FFFGIFLIIYLIN
hs6M1-18     1 -------------------MEIVSTGNETITEFVLLGFYDIPELHF-LFFIVFTAVYVFI
hs6M1-12     1 -----------------------MVNQSSTPGFLLLGFSEHPGLERTLFVVVF-TSYLLT
hs6M1-13P    1 -----------------------MVNQSSAPGFLLLGFSEHPALERTLFVVVF-TSYLLT
hs6M1-16     1 -----------------------MVNQSSPMGFLLLGFSEHPALERTLFVVVF-TSYLLT
hs6M1-7P     1 ---------------------MIIICNDSHSDFILLGFSNKPHLEKILFGSFLXIFYFLT
hs6M1-14P    1 ------------------------ANYSAEERFLLLGFSDWPSLQPVLFALVL-LCYLLT
hs6M1-25P    1 MANTLSSLNSCNVFLLVLNRVMGMTNSSVKGDFILVGFSHQPHLEKILFVAVL-ISYLLT
hs6M1-10     1 ---------------------MNWVNKSVPQEFILLVFSDQPWLEI-PPFVMFLFSYILT
hs6M1-32     1 ---------------------MNWVNDSIIQEFILLGFSDRPWLEF-PLLVVFLISYTVT
hs6M1-1      1 ---------------------MNWENESSPKEFILLGFSDRAWLQM-PLFVVLLISYTIT
hs6M1-22P    1 ---------------------MWINNQSSLDDFILLGFSDRPWLET-PLXVIFLVAYIFS
hs6M1-2P     1 ---------------------MPLTNESHPEEFILLGFADRPWLEL-PLFTSLLIMYPIA
hs6M1-9P     1 --------------------SVKYLNESFPEDFILMGFVKYPWLDF-LLFCVLLTFYMFT
hs6M1-29P    1 ---------------------MDQKNGSSFTGFILLGFSDRPQLELXSPLCGFLIFYIFT
hs6M1-31P    1 ---------------------MERANDSTFSGFILLGFSNRPQLET-ALFVVILIIYFLS
hs6M1-3      1 ------------------MNDDGKVNASSEGYFILVGFSNWPHLEV-VIFVVVLIFYLMT
hs6M1-5      1 ------------------MNDDGKVNASSEGYFILVGFSNWPYLEV-VLFVVILIFCLMT
hs6M1-4P     1 ----------------------MKKNASFEDFFILLGFSNWPHLEV-VLFVVILIFYLIT
hs6M1-6      1 --------------------MMIKKNASSEDFFILLGFSNWPQLEV-VLFVVILIFYLMT
hs6M1-33P    1 -----------------VAAGVENDNTSSFEGFILVGFSDRPHLEL-IVFVVVLIFYLLT
hs6M1-34P    1 ---------------------MEKSNVSSVYGFILVGFSDRPKLEM-VLFTVNFILYSVA
hs6M1-30P    1 ---------------------MTNQSCPETX-FILLGFSGRPRLEH-VLFVFVLIFYLVT
hs6M1-15     1 ---------------------MDQSNYSSLHGFILLGFSNHPKMEM-ILSGVVAIFYLIT
hs6M1-8P     1 ---------------------METSSVSSGTDFILLGFSDRPQLEH-IISVVVFIIYIVT
hs6M1-23P    1 -----------------------MINDSHFSGFILLGFTGQPQLQM-MISGVVFFFYTIA
hs6M1-24P    1 -----------------------MINDSYFGWLMLLGFPGKPQLEM-IISGVVFFFYAIS


hs6M1-17    38 VAGNFLIVVLVSTDAALQSPMY-FFLRTLSALEIGYTSVTVPLLLHHLLTGRRHISRSGC
hs6M1-21    39 LGGNILIILTTVTDPHLHTPMY-YFLGNLAFIDICYTTSNVPQMMVHLLSKKKSISYVGC
hs6M1-28    59 LTGNLLIITIITVDRRLHSPMY-YFLKHLSLLDLCFISVTVPQSIANSLMGNGYISLVQC
hs6M1-35    39 LFGNLLILLAIGSDHCLHTPMY-FFLANLSLVDLCLPSATVPKMLLNIQTQTQTISYPGC
hs6M1-19P   37 VAGNGAILMIVISDPRLHSPMY-FFLGNLSCLDICYSSVTLPKMLQNFLSAHKAISFLGC
hs6M1-20    37 VTGNGAVLMIVISDPRLHSLMY-FFLGNLSYLDICYSTVTLPKMLQNFLSTHKAISFLGC
hs6M1-27    37 LIGNGSILVMVVLEPQLHSPMY-FFLGNLSCLDISYSSVTLPKLLVNLVCSRRAISFLGC
hs6M1-18    41 IIGNMLIIVAVVSSQRLHKPMY-IFLANLSFLDILYTSAVMPKMLEGFLQEAT-ISVAGC
hs6M1-12    37 LVGNTLIILLSALDPKLHSPMY-FFLSNLSFLDLCFTTSCVPQMLVNLWGPKKTISFLDC
hs6M1-13P   37 PVX-GLIILLSVLDPRLHSPMY-FFLSNLSFLDLCFTISCVPGMLVNLWEPKKTIILLGC
hs6M1-16    37 LVGNTLIILLSVLYPRLHSPMY-FFLSDLSFLDLCFTTSCVPQMLVNLWGPKKTISFLGC
hs6M1-7P    40 LAGNMVIVLVSLKDPKLHIPMY-FFLSNLSLVDLCLTSSCVPQMLINFWGPEKTISYIGC
hs6M1-14P   36 LTGNSALVLLAVRDPRLHTPMY-YFLCHLALVDAGFTTSVVPPLLANLRGPALWLPRSHC
hs6M1-25P   60 LVGNTVIILICSVDPKLKTPMYXFFLTHLSLVDICFTTSIVPQLLWNLKGPDKTITFLGC
hs6M1-10    39 IFGNLTIILVSHVDFKLHTPMY-FFLSNLSLLDLCYTTSTVPQMLVNICNTRKVISYGGC
hs6M1-32    39 IFGNLTIILVSRLDTKLHTPMY-FFLTNLSLLDLCYTTCTVPQMLVNLCSIRKVISYRGC
hs6M1-1     39 IFGNVSIMMVCILDPKLHTPMY-FFLTNLSILDLCYTTTTVPHMLVNIGCNKKTISYAGC
hs6M1-22P   39 LFGNISIILVSHLDPQLDSPMY-FFVSNLSFLDLCYTTSTVPQMLVNLRGPEKTISYGGC
hs6M1-2P    39 VMGNITIILMSRLDSRLHSPMY-FFLTNLSFLDMCYTTSIVPQMLFNLGSSKKTISYMGC
hs6M1-9P    40 LLGNSAIILVSQLDSQLHSPMY-FLLTSLSVLYLCFTTTTVPQMLFNLGGPXKNITIG-C
hs6M1-29P   40 LLGNKTIIVLSHLDPHLHNPMY-FFFSNLSFLDLCYTTGIVPQLLVNLRGADKSISYGGC
hs6M1-31P   39 FLGNGTIILLSIVDPRLHTPMY-FFLSNLSFMDLCLTTCTVPQTLVNFKGKDKTITYGGC
hs6M1-3     42 LIGNLFIIILSYLDSHLHTPMY-FFLSNLSFLDLCYTTSSIPQLLVNLWGPEKTISYAGC
hs6M1-5     42 LIGNLFIIILTYLDSHLHTPLY-FFLSNLSFLDLCYTTSSIPQLLVSLWGVEKTISYAGC
hs6M1-4P    38 LIGNLFIIILSYLDSHLHTPMY-FFLSNLSFLDLCYTTSSIPQLLVNLWGPEKTISYAGC
hs6M1-6     40 LTGNLFIIILSYVDSHLHTPMY-FFLSNLSFLDLCHTTSSIPQLLVNLRGPEKTISYAGC
hs6M1-33P   43 LLGNMTIVLLSALDSRLHTPMY-FFLANLSFLDMCFTTGSIPQMLYNLWGPDKTISYVGC
hs6M1-34P   39 VLGNSTIILVCILDSQLHTPMY-FFLANLSFLDLCFSTSCIPQMLVNLWGPDKTISCAGC
hs6M1-30P   38 LVGNIIIILISHLDPCLHMPMY-FFLTNLSFLDLCFTTSSIPQLLFNLGSPGKTISHTGC
hs6M1-15    39 LVGNTAIILASLLDSQLHTPMY-FFLRNLSFLDLCFTTSIIPQMLVNLWGPDKTISYVGC
hs6M1-8P    39 LVGNTTIILVSYLDTQLHTFMY-FFLSNLSFLDLCYTTSIIPQMLANQWGPKKSITYGGC
hs6M1-23P   37 FMGNMAIILLSFLDDHLQVPMY-FFLRNLAILDLCYTTNIVPQMLVSIWGKDKRITFGGC
hs6M1-24P   37 LMGNMVLILLPLLDKHLQTPIY-FFLRNLAILDLCYTTNIVPQMLVNAWGKDKKITFGGC
```

```
hs6M1-17     97 ALQMFFFLFFG-ATECCLLAAMAYDRYAAICEPLRYPLLLSHRVCLQLAGSAWAC-GVLV
hs6M1-21     98 VVQLFAFVFFV-GSECLLLAAMAYDRYIAICNPLRYSVILSKVLCNQLAASCWAA-GFLN
hs6M1-28    118 ILQVFFFIALA-SSEVAILTVMSYDRYAAICQPLHYETIMDPRACRHAVIAVWIA-GGLS
hs6M1-35     98 LAQMYFCMMFA-NMDNFLLTVMAYDRYVAICHPLHYSTIMALRLCASLVAAPWVI-AILN
hs6M1-19P    96 ISQLHFFHFLG-STEAMLLAVMAFDRFVAICKPLRYTVIMNPQLCTQMAITIWMI-GFFH
hs6M1-20     96 ISQLHFFHFLG-STESMLFAVMAFDLSVAICKPLRYTVIMNPQLCTQMAITIWVI-GFFH
hs6M1-27     96 ITQLHFFHFLG-STEAILLAIMAFDRFVAICNPLRYTVIMNPQVCILLAAAAWLI-SFFY
hs6M1-18     99 LLQFFIFGSLA-TAECLLLAVMAYDRYLAICYPLHYPLLMGPRRYMGLVVTTWLS-GFVV
hs6M1-12     96 SVQIFIFLSLG-TTECILLTVMAFDRYVAVCQPLHYATIIHPRLCWQLASVAWVI-GLVE
hs6M1-13P    95 SVQFFIFLSLG-TTECILLTVMAFDRYMAIFKPLRHATIVHLCLCWQLASVAWVI-GLVE
hs6M1-16     96 SVQLFIFLSLG-TTECILLTVMAFDRYVAVCQPLHYATIIHPRLCWQLASVAWVM-SLVQ
hs6M1-7P     99 AIQLYVFLWLG-ATEYVLLVVMAVDCYVAVCHPLQNTMIMHPKLCLQLAILAWGT-GLAQ
hs6M1-14P    95 TAQLCASLALG-SAECVLLAVMALDRAAAVCRPLRYAGLVSPRLCRTLASASWLS-GLTN
hs6M1-25P   120 VIQLYISLALG-STECVLLAVMAFDRYAAVCKPLHYTAVMNPQLCQALAGVAWLS-GVGN
hs6M1-10     98 VAQLFIFLAL-GSTECLLLAVMCFDRFVAICRPLHYSIIMHQRLCFQLAAASWIS-GFSN
hs6M1-32     98 VAQLFIFLAL-GATEYLLLAVMSFDRFVAICRPLHYSVIMHQRLCLQLAAASWVT-GFSN
hs6M1-1      98 VAHLIIFLAL-GATECLLLAVMSFDRYVAVCRPLHYVVIMNYWFCLRMAAFSWLI-GFGN
hs6M1-22P    98 VAQLYIFLAL-GSTECILLAIMAFDRYAAICKPLHYPVIMNHRRCIHMAAGTWIS-GFAN
hs6M1-2P     98 AVQLYFFHIM-GGTECLLLAIMSFDRYVAICRPLHYTLIMNQRVCILXVSTVWLI-GIIY
hs6M1-9P     98 MAQAYVFHWL-ACIECVLLGIVALDCYVAVCKPPRYTIIIDHKVCLHLSSTAWLI-GLAN
hs6M1-29P    99 VVQLYISLGL-GSTECVLLGVMAFDRYAAVCRPLHYTVVMHPCLYVLMASTSWVI-GFAN
hs6M1-31P    98 VTQLFIALGLXGGVECVLLSAMAYDRYAAVCRPLHYMVSMHPQLCLQLVVTTWLT-GFGN
hs6M1-3     101 MIQLYFVLAL-GTTECVLLVVMSYDRYAAVCRPLHYTVLMHPRFCHLLAVASWVS-GFTN
hs6M1-5     101 MVQLYFFLTL-GTTECVLLVVMSYDRYAAVCRPLHYTVLMHSRFCHLLAVASWVS-GFTN
hs6M1-4P     97 TVQLYFVLAL-GTAECVLLVVMSYDRYAAVCRPLHYTVLMHPRFCRLLAAASWVS-GFTT
hs6M1-6      99 MVQLYFVLAL-GIAECVLLVVMSYDRYVAVCRPLHYTVLMHPRFCHLLAAASWVI-GFTI
hs6M1-33P   102 AIQLYFVLALG-GVECVLLAVMAYDRYAAVCKPLHYTIIMHPRLCGQLASVAWLS-GFGN
hs6M1-34P    98 VVQLFSFLSVR-GIECILLAVMAYDSYAAVCKPLRYLVIMHLQLCLGLMAAAWGS-GLVN
hs6M1-30P    97 AIQLFMFLGLXGWQECILLAAVAYDRFIAICKPLHYSVIMHPQLCWKLVSVARGCXGLLS
hs6M1-15     98 IIQLYVYMWLG-SVECLLLAVMSYDRFTAICKPLHYFVVMNPHLCLKMIIMIWSI-SLAN
hs6M1-8P     98 VLQFFFVLDLG-ATECLLLAVMAYDRYAAVCQPLHY-----TLKCTLSFATAWLS-GLAS
hs6M1-23P    96 AFQLFIDVALY-SVECILLSMMSYDRLNAICKPLHHMTIMNLQLCQGLVVISWVV-GVIN
hs6M1-24P    96 AFQLFTNVTLC-TVECMLLAVMSYDPFNAVCKPLDYMTIMNPQLCQGLVAMTWLI-GVTN
hs6M1-26P     1 -----------GSTKCIILAVTSLDPYIAICKHLRYPAIMHQQLCVLLVAMAWLS-SLAN

hs6M1-17    155 GLGHTPFIFSLPFCGPNTIPQFFCEIQPVLQLVCG----DTSLN-ELQIILATALLILCP
hs6M1-21    156 SVVHTVLTFCLPFCGNNQINYFFCDIPPLLILSCG----NTSVN-ELALLSTGVFIGWTP
hs6M1-28    176 GLMHAAINFSIPLCGKRVIHQFFCDVPQMLKLACS----YEFIN-EIALAAFTTSAAFIC
hs6M1-35    156 PLLHTLMMAHLHFCSDNVIHHFFCDINSLLPLSCS----DTSLN-QLSVLATVGLIFVVP
hs6M1-19P   154 ALLHSLMTSRLNFCGSNRIYHFFCDVKPLLKLSX--------LN-QWLLSTVTGTIAMGP
hs6M1-20    154 ALLHSVMTSRLNFCGSNRIHHFLCDIKPLLKLACG----NTELN-QWLLSTVTGTIAMGP
hs6M1-27    154 ALMHSVMTAHLSFCGSQKLNHFFYDVKPLLELACS----DTLLN-QWLLSIVTGSISMGA
hs6M1-18    157 DGLVVALVAQLRFCGPNHIDQFYCDFMLFVGLACS----DPRVA-QVTTLILSVFCLTIP
hs6M1-12    154 SVVQTPSTLHLPFCPDRQVDDFV-CEVPALIRLS---CEDTSYN-EIQVAVASVFILVVP
hs6M1-13P   153 SVVQTPSTLRLPFCPHQQVDDFV-CEVPALIRLS---CEDTSYN-EIQMAVASVFILAVP
hs6M1-16    154 SIVQTPSTLHLPFCPHQQIDDFL-CEVPSLIRLS---CGDTSYN-EIQLAVSSVIFVVVP
hs6M1-7P    157 SLIQSPATLRLPFCSQRMVDDVV-CEVPALIQLS---STDTTYS-EIQMSIASVVLLVMP
hs6M1-14P   153 SVAQTALLAERPLCAPRLLDHFI-CELPALLKLA---CGGDGDTTENQMFAARVVILLLP
hs6M1-25P   178 TLIQALSPSGFLAVDTDCSNISSXREVPSMIKLA---CVDIHDN-EVQLFVASLVLLLLP
hs6M1-10    156 SVLQSTWTLKMPLCGHKEVDHFFCE-VPALLKLS---CVDTTAN-EAELFFISVLFLLIP
hs6M1-32    156 SVWLSTLTLQLPLCDPYVIDHFLCE-VPALLKLS---CVETTAN-EAELFLVSELFHLIP
hs6M1-1     156 SVLQSSLTLNMPRCGHQEVDHFFCE-VPALLKLS---CADTKPI-EAELFFFSVLILLIP
hs6M1-22P   156 SLVQSTLTVVAPRCGQRVLDHFFCE-VPALLKLA---CIDIRVN-EMELNVLGALLLLMP
hs6M1-2P    156 AVSEATATLQLPLCGLNKLDHLVCE-IPVLIKIA---CGEKGSN-ELTLSVVCIFMLAVP
hs6M1-9P    156 SLLQSTITIQLPL-----------------------------------------------
hs6M1-29P   157 SLLQTVLILLLTLCGRNKLEHFLCE-VPPLLKLA---CVDTTMN-ESELFFVSVIILLVP
hs6M1-31P   157 SVIQTALTMTLPLCDKNQVDHFFCE-VPVMLKLS---CTNTSIN-EAEIFAVSVFFLVVP
hs6M1-3     159 SALHSSFTFWVPLCGHRQVDHFFCE-VPALLRLS---CVDTHVN-ELTLMITSSIFVLIP
hs6M1-5     159 PALHSSFTFWVPLCGHRQIDHFFCE-VPALLLS----FVNTREN-KLTLMITSSIFVLLL
hs6M1-4P    155 SALHSSFTFWIPLCRHRLVDHFFCE-APALLRLS---CVDTAN--ELTLMVMSSIFVLIP
hs6M1-6     157 SALHSSFTFWVPLCGHRLVDHFFCE-VPALLRLS---CVDTHAN-ELTLMVMSSIFVLIP
hs6M1-33P   160 SLIMAPQTLMLPRCGHRRVDHFLCE-MPALIGMA---CVDTMML-EALAFALAIFIILAP
hs6M1-34P   156 AVVMSPLTMTLSRSGRRRVNHFLCEXKPALIKMA---CLDVRAV-EMLAFAFAVLIVLLP
hs6M1-30P   157 SLVMSPVTMKLPRCGRCKLKHFLCE-MPALIKIT---CVDTVAM-ESTVFTLSVVIVLMP
hs6M1-15    156 SVVLCTLTLNLPTCGNNILDHFLCE-LPALVKIA---CVDTTTV-EMSVFALGIIIVLTP
hs6M1-8P    151 ALIVCSLTLKLPRCGHREVDNFFCE-MPALIKMA---CVYSKVI-EIVVFAFGVVFLFVP
hs6M1-23P   154 CIIPSPYATSLPRCRNHHLDHFFVCVKCLQSRFK--IACVDTTAMEVTTFAMCLIIVLVP
hs6M1-24P   154 CMILSPCPVSLPRCGDHHLDHYFCEISAMVKIACGATTVMEETKPYLHCVVVVVFIFLAS
hs6M1-26P    49 STSVIXLAVQLPLGG-NKVDDFLCEVSAMIKISR----FDTTFN--V-SMLSIVRIFSLV
```

```
hs6M1-17    210 FGLILGSYGRILVTIFRIP-SVAGRRKAFSTCSS-HLIVVSLFYGTALFIYIRPKASYDP
hs6M1-21    211 FLCIVLSYICIISTILRIQ-SSEGRRKAFSTCAS-HLAIVFLFYGSAIFTYVRPISTYSL
hs6M1-28    231 LISIVLSYIRIFSTVLRIP-SAEGRTKVFSTCLP-HLFVATFFLSAAGFEFLRLPSDSSS
hs6M1-35    211 SVCILVSYILIVSAVMKVP-SAQGKLKAFSTCGS-HLALVILFYGAITGVYMSPLSNHST
hs6M1-19P   205 FFLTLLSYFYIITHLFFKTHSFSMLRKALSTCAS-HFMVVILLYAPVLFTYIHHASGTSM
hs6M1-20    209 FFLTLLSYFYIITYLFFKTRSCSMLCKALSTCAS-HFMVVILFYAPVLFTYIHPALESFM
hs6M1-27    209 FFLTLLSCFYVIGFLLFKNRSCRILHKALSTCAS-HFMVVCLFYGPVGFTYIRPASATSM
hs6M1-18    212 FGLILTSYARIVVAVLRVP-AGASRRRAFSTCSS-HLAVVTTFYGTLMIFYVAPSAVHSQ
hs6M1-12    209 LSLILVSYGAITWAVLRIN-SAKGRRKAFGTCSS-HLTVVTLFYSSVIAVYLQPKNPYAQ
hs6M1-13P   208 XSLILVSYGAIAWAVLRTN-CKXGQRKAFGTCSS-HLTVVTLFYSSVIAVYLQPKNPYAQ
hs6M1-16    209 LSLILASYGATAQAVLRIN-SATAWRKAFGTCSS-HLTVVTLFYSSVIAVYLQPKNPYAQ
hs6M1-7P    212 LIIILSSSGAIAKAVLRIK-STAGQKKAFGTCIS-HLLVVSLFYGTVTGVYLQPKNHYPH
hs6M1-14P   209 FAVILASYGAVARAVCCMR-FSGGRRRAVGTCGS-HLTAVCLFYGSAIYTYLQPAQRYNQ
hs6M1-25P   234 LVLILLSYGHIAKVVIRIK-SVQAWCKGLGTCGS-HLIVVSLFCGTITAVYIQSNSSYAH
hs6M1-10    211 VTLILISYAFIVQAVLRIQ-SAEGQRKAFGTCGS-HLIVVSLFYGTAISMYLQPPSPSSK
hs6M1-32    211 LTLILISYAFIVRAVLRIQ-SAEGRQKAFGTCGS-HLIVVSLFYSTAVSVYLQPPSPSSK
hs6M1-1     211 VTLILISYGFIAQAVLKIR-SAEGRQKAFGTCGS-HMIVVSLFYGTAIYMYLQPPSSTSK
hs6M1-22P   211 LTLILGTYVFIAQAVMRIC-SAESRWKAFNTCAS-HLLVVSLFYFTAISMYVQPPSSYSH
hs6M1-2P    211 LCLILASYASIGSAVFKIK-SSKGRKKAFGTCSS-HLIVVFLFYGPAISMYLQPPSSISR
hs6M1-29P   212 VALIIFSYSQIVRAVVRIK-SATGQRKVFGTCGSPHLTVVSLFYGTAIYAYLQPGNNYSQ
hs6M1-31P   212 LSLILASYGHITHAVLKIK-SAQGRQKAFGTCGS-HLLVVIIFFGTLISMYLQPPSSYSQ
hs6M1-3     214 LILILTSYGAIVRAVLRMQ-STTGLQKVFGTCGA-HLMAVSLFFIPAMCIYLQPPSGNSQ
hs6M1-5     213 LTLIFTSYGAIAQAVLRMQ-STTGLQKVFGTCGA-HHMVVSLFFIPAMCMYLQPPSGNSQ
hs6M1-4P    209 LILILTSYGAIARAVLSMQ-STTGLQKVLRTCGA-HLMVVSLFFIPVMCMYLQPPSENSQ
hs6M1-6     212 LILILTAYGAIARAVLSMQ-STTGLQKVFRTCGA-HLMVVSLFFIPVMCMYLQPPSENSP
hs6M1-33P   215 LILILISYGYVGGTVLRIK-SAAGRKKAFNTCSS-HLIVVSLFYGTIIYMYLQPANTYSQ
hs6M1-34P   212 LTLILVSYGYIAAAVLSIK-SAARQWKAFHTCSS-HLTVVSLFYGSIIYMYMQPGNSSSQ
hs6M1-30P   212 LCLILISYSYIALAVLRIK-SAAGRRKAFNMCGS-HLTVVSLFYGNIIYMYMQPNNS-SQ
hs6M1-15    211 LILILISYGYIAKAVLRTK-SKASQRKAMNTCGS-HLTVVSMFYGTIIYMYLQPGNRASK
hs6M1-8P    206 LSLILISYGVITQAVMRIK-SATRLQKILNTCGS-HLTVVILFYGTIIYIYMKPQNTISQ
hs6M1-23P   212 LLLILVSYGFIAVAVLKIK-SAAGRQKAFGTCSS-HLVVVSIFCGTVTYMYIQPGNSPNQ
hs6M1-24P   214 LLLILVSYGFIAVAVLKIK-SAAGRQKAFGTCFS-HLIVVSIFYGTVRYMYIEPGNSPSQ
hs6M1-26P   101 LSIIFAYCGFIVATVLRIQ-SSGGKKEVFNTCGS---HIVSLLYGPVISMYVQPSAN-SQ


hs6M1-17    268 ATDPLVS-LFYAVVTPILNPIIYSLRNTEVKAALKRTIQKTVPMEI--------------
hs6M1-21    269 KKDRLVS-VLYSVVTPMLNPIIYTLRNKDIKEAVKTIGSKWQPPISSLDSKLTY------
hs6M1-28    289 TVDLVFS-VFYTVIPPTLNPVIYSLRNDSMKAALRKMLSKEELPQRKMCLKAMFKL----
hs6M1-35    269 EKDSAAS-VIFMVVAPVLNPFIYSLRNNELKGTLKKTLSRPGAVAHACNPSTLGGRGGWI
hs6M1-19P   264 DQDRITA-IMYTVVTPVLNPLIYTLRNKEVKGAFNRAMKRWLWPKEILKNSSEA------
hs6M1-20    268 DQDRIVA-IMYTVVTPVLNPLIYTLRNKEVKGALGRVIRRL-------------------
hs6M1-27    268 IQDRIMA-IMYSAVTPVLNPLIYTLRNKEVMMALKKIFGRKLFKDWQQHH----------
hs6M1-18    270 LLSKVFS-LLYTVVTPLFNPVIYTMRNKEVHQALRKILCIKQTETLDRRX----------
hs6M1-12    267 ERGKFFG-LFYAVGTPSLNPLIYTLRNKEVTRAFRRLLGKEMGLTQS-------------
hs6M1-13P   266 ERGKFFG-LFYAVGTPSLNPLIYTLRNKEVTRAFRRLLAKEMGLIQS-------------
hs6M1-16    267 GRGKFFG-LFYAVGTPSLNPLVYTLRNKEIKRALRRLLGKERDSRESWRAA---------
hs6M1-7P    270 EWGKFLT-LFYTVVTPTLNPLIYTLRNKEVKGALIRLGRRTWDSQNN-------------
hs6M1-14P   267 ARGKFVS-LFYTVVTPALNPLIYTLRNKKVKGAARRLLRSLGRGQAGQ------------
hs6M1-25P   292 AHGKFIS-LFYTVVTPTLNPLIYTLRNNDVKGALRLFNRDLGT-----------------
hs6M1-10    269 DRGKMVS-LFCGIIAPMLNPLIYTLRNKEVKEAFKRLVAKSLLNQEIRNMQMISFAKDTV
hs6M1-32    269 DQGKMVS-LFYGIIAPMLNPLIYTLRNKEVKEGFKRLVARVFLIKK--------------
hs6M1-1     269 DWGKMVS-LFYGIITSMLNSLIYSLRNKDMKEAFKRLMPRIFFCKK--------------
hs6M1-22P   269 DRGKIMA-LFYGIVTPTLNPFIYTLRNKDVKAALRRSLTKEFWIKTR-------------
hs6M1-2P    269 DQPKFMA-LFYGVVTPSLNPFIYTLRNKNVKGALRNLVRSIFSFK---------------
hs6M1-29P   271 DQGKVIS-LFYTIITPMINPLIYTLRNKDVKGALKKVLWKNYDSR---------------
hs6M1-31P   270 DVNKSIA-LFYTLVTPLLNPLIYTLRNKEVKGATKKTSGEDHRCMRKLTQGLQFQTFVH-
hs6M1-3     272 DQGKFIA-LFYTVVTPSLNPLIYTLRNKVVRGAVKRLMGWE-------------------
hs6M1-5     271 DQGKFIA-LFYTVVTPSLNPLIYTLRNKDVRGVVKRLRGWE-------------------
hs6M1-4P    267 DQGKFIA-LFYTVVTPSLNPLIYTFRNKDVRGAVKRLMGWEWGM----------------
hs6M1-6     270 DQGKFIA-LFYTVVTPSLNPLIYTLRNKHVKGAAKRLLGWEWGK----------------
hs6M1-33P   273 DQGKFLT-LFYTIVTPSVNPLIYTLRNKDVKEAMKKVLGKGSAEI---------------
hs6M1-34P   270 DQGKFLT-LFYNLVTPMLNLLIYTLRNKEVKGALKKVLGRQ-------------------
hs6M1-30P   269 DQGKFLT-LFYNLMTPMLNPVIYTLRNKDVKGALKRLVSRKHSDSDCS------------
hs6M1-15    269 DQGKFLT-LFYTVITPSLNPLIYTLRNKDMKDALKKLMRFHHKSTKIKRNCKS-------
hs6M1-8P    264 DEGKFFTXLFYTIITPSLNLPIYTLRNKDVKSALKRILWMKKSSAES-------------
hs6M1-23P   270 NEGKLLS-IFYSIVTPSLNPLIYTVRNKEFKGAMKRLTGKEKDCMEKRGH----------
hs6M1-24P   272 DEGKLLH-IFYSIVTPTLNPXIP-LRNKEFKWAMKRLIGKEKGSGDTI------------
hs6M1-26P   156 DKNKFMS-LFYSLVTPMLNPFIYTLSNRDIKGAMRRLLVFLYHQEENKSNYFYTPHSSYT


hs6M1-35    328 MRSGDRDHPG-------------------- 337
hs6M1-10    328 LTYLTNFSASCPIFVITIENYCNLPQRKFP 357
hs6M1-26P   215 GQKISCSKITC------------------- 225
```

Appendix 8: List of mouse MHC-linked ORs

Name	Clone		Position in clone	State	Length	Comment
mm17M1-1	AL078630	<	65954..66892	C	939	
mm17M1-2	AL078630	<	105931..106869	C	939	
mm17M1-3	AL078630	<	40162..41094	C	933	
mm17M1-4	AL078630	<	47848..48786	C	939	
mm17M1-5	AL078630	<	124623..125561	P	939	Substitution at 292, 293 or 294 > stop codon
mm17M1-6	AL078630	<	150388..151317	C	1161	
mm17M1-7	AL133159	>	84417..85262	P	930	Substitution at 1 > Valine not methionine start
mm17M1-8	AL133159	>	63932..64765	P	930	Substitution at 1 > Valine not methionine start
mm17M1-9	AL133159	>	96478..97311	P	930	Substitution at 1 > Valine not methionine start
mm17M1-10	AL133159	>	107383..108312	C	930	
mm17M1-11	AL133159	>	89150..89942	C	963	
mm17M1-12	AL133159	<	1306..2232	C	927	
mm17M1-13	AL133159	<	15182..16108	C	927	
mm17M1-14	AL133159	<	38006..38947	C	942	
mm17M1-15	AL133159	<	109667..110311	P	595	Missing 1st 2 motifs Substitution at 55, 56, 57> stop codon Substitution at 109, 110, 111 > stop codon 1 bp insertion at 172 > frameshift
mm17M1-16	AL133159	<	148446..148910	P	523	Missing 1st 2 motifs Substitution at 13, 14, 15 > stop codon Substitution at 172, 173, 174 > stop codon Substitution at 205, 206, 207 > stop codon 1 bp insertion at 433 > frameshift
mm17M1-17	AL133159	<	154600..155102	P	523	As above
mm17M1-18	AL136158	<	111592..112530	C	1014	
mm17M1-19	AL136158	<	91202..92140	C	1014	
mm17M1-20	AL136158	<	76271..77185	C	1014	
mm17M1-21	AL136158	<	159485..160438	C	954	
mm17M1-22	AL136158	>	171735..172700	C	1002	
mm17M1-23	AL136158	>	27408..28361	C	954	
mm17M1-24	AL136158	<	60863..61900	C	1038	
mm17M1-25	AL359381	<	80755..81681	C	927	
mm17M1-26	AL359381	<	8047..8979	C	933	
mm17M1-27	AL359381	<	39327..40256	C	930	
mm17M1-28	AL359381	<	56079..56996	C	918	
mm17M1-29	AL359381	>	1739..2680	C	942	

mm17M1-30	AL359381	>	25985..27445	P	942	Substitution at 25, 26, 27 > stop codon LTR Insertion at 262 Substitution at 391, 392, 393 > stop codon Substitution at 577, 578, 579 > stop codon 1 bp insertion at 796 > frameshift
mm17M1-31	AL359381	>	53234..54676	P	917	Substitution at 25, 26, 27 > stop codon LTR Insertion at 262 Substitution at 391, 392, 393 > stop codon Substitution at 571, 572, 573 > stop codon 1 bp deletion at 630 > frameshift
mm17M1-32	AL136158	<	125473..126438	C	966	
mm17M1-33	AL450393	>	35674..36816	C	1143	
mm17M1-34	AL450393	>	3434..4378	C	945	
mm17M1-35	AL365336	>	11562..12527	C	966	
mm17M1-36	AL365336	>	36956..37921	C	966	
mm17M1-37	AL359352	>	13527..14492	C	966	
mm17M1-38	AL359352	>	33373..34338	C	966	
mm17M1-39	AL359352	>	73703..74650	C	948	
mm17M1-40	AL136158	>	89150..89942	P	793	1 bp deletion at 153 > frameshift Substitution at 273, 274, 275 > stop codon Substitution at 435, 436, 437 > stop codon 1 bp deletion at 461 > frameshift 1 bp insertion at 581 > frameshift 1 bp deletion at 722 > frameshift
mm17M1-41	AL359352	>	89991..90953	C	963	
mm17M1-42	AL359352	>	105573..106532	C	960	
mm17M1-43	AL359352	>	153054..154025	C	972	
mm17M1-44	AL359352	<	40711..41542	P	832	Substitution at 124, 125, 126 > stop codon Substitution at 133, 134, 135 > stop codon 1 bp deletion at 204 > frameshift 1 bp insertion at 270 > frameshift Substitution at 298, 299, 300 > stop codon Substitution at 349, 350, 351 > stop codon Substitution at 403, 404, 405 > stop codon Substitution at 421, 422, 423 > stop codon Substitution at 514, 515, 516 > stop codon 1 bp deletion at 730 > FS
mm17M1-45	AL590433	>	64004..64933	C	930	
mm17M1-46	AL590433	<	11889..12827	C	939	

Appendix 9: Alignment of mouse MHC-linked ORs

```
mm17M1-6       1 MWLCNKTKTWACAEFIPYWRFLFVVVSGKTGFYYVALAGLELTEISGLCLPAQGPQHCLA
mm17M1-33      1 -----------MRVCTYILRDRYGLCFAKKCFICGSVSNVVGGDIIGTSTCFLVLQTIIN
mm17M1-22      1 --------------------------------------------------------MDLM
mm17M1-19      1 -------------------------------------------------MPQFLSTAFVV
mm17M1-20      1 -------------------------------------------------MPQSHSTAFIV
mm17M1-24      1 -------------------------------------------------MPQCHSTAFIV
mm17M1-18      1 -------------------------------------------------MPQFLSTVFVV
```

```
mm17M1-29      1 ------------MGILSTGNQTVTEFVLLGFHEVPGLHLLFFSVFTILYASIITGNMLIA
mm17M1-30      1 -------------------MEMSPQKTFTESVPLGLSEVPPIYLCIMIYASIITKNMLVL
mm17M1-31P     1 -------------------MGMSPQETFTESVPHGLSEVPPIYLCIMIYASIITRSMLIV
mm17M1-10      1 -----------------MNCSKTPGFILLGLSSDPEKWQPLFNIFLCLYLLGLLGNLLLL
mm17M1-11      1 ------MQISSSIISPRMNCSQAPGFILLGLPREPEKWQHFFIIFLGLYLLGLLGNLLLL
mm17M1-26      1 ----------------MMNCSQAPGFILLGLSSNSEKWQPLFSIFLVLYLLGLLGNLLLL
mm17M1-27      1 -----------------MNCSQAPTFILLGLSSDAEKWQPLFSIFLVLYLLGLLGNLLLL
mm17M1-28      1 -----------------MNCSQAPTLILLGLSSDAEKWQPLFSIFLVLYLLGLLGNLLLL
mm17M1-1       1 ----------------MVNQSSPVVFFLLGFSEHPQLKKVLFVVVLCSYLLTLLGNTLIL
mm17M1-5P      1 ----------------MVNQSSPVVFFLLGFSEHPQLEKVLFVVVLCSYLLTLLGNTLIL
mm17M1-2       1 ----------------MVNQSTPVGFLLLGFSEHPQLEKVLFVVVLCSYLLTLLGNTLIL
mm17M1-4       1 ----------------MVNQSSPVGFLLLGFSEHPQLEKVLFVIVLCSYLLTLLGNTLIL
mm17M1-3       1 ----------------MVNQSSPVGFLLLGFSEHPQLEKVLIVVVLCSYLLTLLGNTLIL
mm17M1-21      1 ---------------MAINKSSGGDFILVGFSDQPQLEKILFVLVLISYLLTLVGNTAII
mm17M1-45      1 ----------------MINSSVSSDFILVGFSDQPQLERRLFIVVLISYLLTLVGNTIII
mm17M1-6      61 LNIFVSPSEPSWSFPPQANHSSAERFLLLGFSDWPSLQPVLFALVLLCYLLTLTGNAALV
mm17M1-39      1 --------------MWINNQSSVDDFILLGFSDRPWLETPLFVIFLVAYIFALFGNISII
mm17M1-13      1 ----------------MSNQTSVTEFLLLGVTDIQELNPILFVIFFTIYFVNITGNGAIL
mm17M1-25      1 ----------------MSNQTSVTEFLLLGVTDVQELNPILFVIFFTIYFVNITGNGAIL
mm17M1-12      1 ----------------MSNQTSVTEFLLLGVTDIQELNPILFVIFFTIYFINITGNGAIL
mm17M1-14      1 ----------------MLNQTSVTEFILLGVRDIQEPQPFLFAIFFTIYFVNITGNGAIL
mm17M1-8       1 ---------------VLLNHTLVTEFLLLGVTDIQELNPILFVTVLAMYFVNVAGNGAIL
mm17M1-9       1 ---------------VLLNQTLVTEFLLLGVTDIQELNPILFVTVLAMYFVNVAGNGAIL
mm17M1-7P      1 ---------------VLLNHTFITEFLLLGVTDIQELNPILFVMVLAMYFINVFGNGAIM
mm17M1-34      1 ----------------MENSTSVDEFLLLGLTSVQKLQPIIFVMFLTIYLLNLVGNGVIL
mm17M1-23      1 --------------MEGKNQTAPSEFIILGFDHLNELQYLLFTIFFLTYICTLGGNVFII
mm17M1-33     50 APVCQIHKGKPADIMEGKNQTAPSEFIILGFDHLNELQYLLFTIFFLTYICTLGGNVFII
mm17M1-22      5 ICCPFFQEMSVNCSLWQENKLSVKHFAFAKFSEVPEECFLLFTLILLMFLVSLTGNALIT
mm17M1-36      1 ---------MSVNCSLWQENKLSVKHFAFAKFSEVPEECFLLFTLILLMFLVSLTGNALIT
mm17M1-35      1 ---------MSINCSLWQENSLSVKRFAFAKFSEVPGECFLLFTLILLMFLVSLTGNALIA
mm17M1-37      1 ---------MSINCSLWQENSLSVKRFAFSKFSEVPGECFLLFTLILLMFLVSLTGNELIA
mm17M1-38      1 ---------MSINCSLWQENSLSVKRFAFAKFSEVPGECFLLFTLILLMFLVSLTGNSLIA
mm17M1-46      1 ---------------MGTNSSLVTEFVLVGFSRLVHLQGILFSLFLTVYLLTVAGNLLIV
mm17M1-19     12 LFYFPAATTISVFKMIMENITTMSGFLLMGFSDNHELQILQAVLFLVTYLVGSAGNVIII
mm17M1-20     12 LFYFPAATAISVFKMIVENITTMRGFLLMGFSDNRELQILHALFFLVAYLLGSAGNVIII
mm17M1-24     12 LFYFPAATVIYVFKMIVENITTMSGFLLMGFSDNHELQILQALLFLVTYLVGSAGNVIII
mm17M1-32      1 --------------MTVKNITTMSGFLLMGFSDNRELQILYALLFLLTYLLGSAGNFIII
mm17M1-18     12 LSYFLAATTISVAKMIMENITTMSGFLLMGFSDNRELQILQALLFLVTYLVGSAGNFIII
mm17M1-42      1 --------------MTARNMTTMSGFLLMGFSDNHELQILQALLFLLTYLLGSAGNFIII
mm17M1-43      1 --------------MTPRNMTTVSGFLLMGFSDNHELQILQALLFLVTYLLDSAGNFIII
mm17M1-41      1 -----------------MNVSFKTGFLLMGFSDERNLQILHAVLFLITYLLAIMGNLLII
mm17M1-40P     1 -------------MNVANFTAMTIFLLLMGFSRNSQVEIIFSTLALVVLIGTISIVAVTS
mm17M1-44P     1 ---------------------------------------------------MLTQNTMII
```

```
mm17M1-29    49 VVVVSSQRLHTPMYFFLVNLSFIEIVYTSTVVPKMLEG-----FLQEATISVAGCLLQFF
mm17M1-30    42 VVVNSSQKLPTPMYFFLVSQAFQAHVHSGDQNAGGFSVGKPHSCDBKSLATLAGCLLQFL
mm17M1-31P   42 VMLDSSQRLHTPTHFFLVSQAFQAHVHSGDQNAGGFPVGKHHSCEBKSLVTLPGCLLQFF
mm17M1-10    44 LAIGTDVHLHTPMYFFLSQLSLVDLCFITTTAPKMLEALW----TGDGSISFSGCLTQFY
mm17M1-11    55 LAIGSDVHLHTPMYFFLSQLSLVDLCFITTTAPKTLETWW----TGDGSISFSGCLTQLY
mm17M1-26    45 LAIGTDVHLHTPMYFFLSQLSLVDLCFITTTAPKMLETLW----TGDGSISFSGCLTQLY
mm17M1-27    44 LAIGTDVHLHTPMYFFLSQLSLVDLCFITTTAPKMLEALW----TGDGSISFSGCLTQLY
mm17M1-28    44 LAIGTDVHLHTPMYFFLSQLSFVDLCFITTTAPKMLEALW----TGDGSISFSGCLTQLY
mm17M1-1     45 LLSTLDPRLHSPMYFFLSNLSFLDLCFTTTCVPQMLFNLW----GPAKTISFLGCFVQLF
mm17M1-5P    45 LLSTLDPRLHTPMYFFLSNLSFLDLCFTTTCVPQMLFNLW----GPEKTISFLGCFV-LF
mm17M1-2     45 LLSTLDPRLHSPMYFFLSNLSFLDLCFTTTCVPQMLFNLW----GPTKTISFLGCSVQLF
mm17M1-4     45 LLSTLDPRLHSPMYFFLSNLSFLDLCFTTTCVPQMLFNLW----GPAKTISFLGCSVQLF
mm17M1-3     45 LLSTLDPRLHSPMYFFLSNLSFLDLCFTTTCVPQMLFNLW----GPAKTISFLGCFVQLF
mm17M1-21    46 LVSCLDSALQTPMYYFLTNLSFVDICFSTSIVPQLLWNLH----GPAKTITATGCAIQLY
mm17M1-45    45 LISSIDSKLKTPMYFFLTHLSFVDICFTTSIVPQLLWNLK----GPAKTITAVGCAVQLY
mm17M1-6    121 LLAIRDPRLHTPMYYFLCHLALVDVGFTTSVVPPLLASLR----GSMLQLPRAGCMAQLC
mm17M1-39    47 LVSRLDPQLDSPMYFFVSNLSLLDLCYTTSTVPQMLVNLR----GPEKTISYGGCVAQLY
mm17M1-13    45 MIVILDPRLHSPMYFFLGNLACLDICFSTVTLPKMLQNLL----STNKAISFLGCITQLH
mm17M1-25    45 MIVILDPRLHSPMYFFLGNLACLDICFSTVTLPKMLQNLL----STNKAISFLGCITQLH
mm17M1-12    45 MIVILDPRLHSPMYFFLGNLACLDISYSTVTVPKLLQNLL----STSKAISFLGCITQLH
mm17M1-14    45 MIVILDPRLHSPMYFFLGNLACLDISYSTVTVPKMLENLL----STNKAISLLGCITQLH
mm17M1-8     46 MIVISDPRLHLPMYFFLGNLACLDICFSTVTVPKMLENFF----STSKAISFLGCITQLH
mm17M1-9     46 LIVISDPRLHSPMYFFLGNLACLDICFSTVTVPKILDNFF----STSKAISFLGCITQLY
mm17M1-7P    46 MIVILDSRLYSPMYFFLGNLACLDICFSTVTVPKMLENFF----STSKAISFLGCITQLH
mm17M1-34    45 MIVTLERRLHSPMYFFLGNLSCLDICYSSVTLPKVLINLL----SRRRAISFLGCITQLY
mm17M1-23    47 VVTIADSHLHTPMYYFLGNLALIDICYTTTNVPQMMVHLL----SEKKIISYGGCVTQLF
mm17M1-33   110 VVTIADSHLHTPMYYFLGNLALIDICYTTTNVPQMMVHLL----SEKKIISYGGCVTQLF
mm17M1-22    65 LAICTSPALHTPMYFFLANLSLLEIGYTCSVIPKMLQNLV----SEIRGISREGCVTQMF
mm17M1-36    53 LAICTSPALHTPMYFFLANLSLLEIGYTCSVIPKMLKNLV----TEARGISREGCATQMF
mm17M1-35    53 LVICTNPSLHNPMYFFLANLSLLEIGYTCSVIPKMLQSLV----SEAREISREGCATQMF
mm17M1-37    53 IAICTSPALHTPMYFFLANLSLLEIGYTCSVIPKMLQSLV----SEAREISREGCATQMF
mm17M1-38    53 LAICTSPALHTPMYFFLANLSLLEIGYTCSVIPKMLQSLV----SEARGISREGCATQMF
mm17M1-46    46 ALVSTDAALQSPMYFFLRILSALEICYTSVTVPLLLHHLL----TGRRHISRSGCALQMF
mm17M1-19    72 TITTLDPQLQSPMYYFLKQLSILDLSSLSVTVPQYVDSS----LARSGYISYGQCMLQIF
mm17M1-20    72 TITTLDPQLQSPMYYFLKHLSILDLSSLSVTVPQYVDIC----LTQSGYISYAQCMLQIF
mm17M1-24    72 TITTLDPQLQSPMYYFLKHLSILDLSSLSVTVPQYVDSS----LAQSGYISYAQCMLQIF
mm17M1-32    47 TITTLDPQLQSPMYYFLKHLSILDLSSLSVTVPQYVDSS----LAGSGYISYGQCMLQIF
mm17M1-18    72 TITTLDPQLKSPMYYFLKHLSILDLSSLSVTVPQYVDSS----LARSGYISYEQCMLQIF
mm17M1-42    47 TITTLDPQLQSPMYYFLKQLSTLDLSSLSVTVPQYVASS----LARSGYISYGQCMLQIF
mm17M1-43    47 TITTIDKQLQSPMYYFLKHLSIMDFSSLSVTVPQYVDSS----LARSGYISYGQCMLQVF
mm17M1-41    44 TIITLDQRLHSPMYYFLKHLSFLDLCFISVTVPQSIANS----LMNNGFISLGQCMLQVF
mm17M1-40P   48 LSIVX---LCSLMPFLLIHLFCFDVCYISVMMPKSVCS------SFMYSAYISPNAHCQV
mm17M1-44P   10 LVSFLNSRLQTPMYFFLSNFFFLDLCFMTNVLIVTSKG------PEK---ITHAVQSMST

mm17M1-15P    1 ----------FLLIVMVYDHYLTICHHL-YPFLMGPLWGLGFGLTDLX--FVVDELIVAL
mm17M1-29   104 VFGSLATDECFLLAVMAYDRYLAICHPLRYPHLMGPQWCLGLVLTVWLSGFMVDGLVVAL
mm17M1-16P    1 ------------------DHYLICHPLH-YPLLMGHQWCLGFVLTLQLFGITVDGLVVIL
mm17M1-17P    1 ------------------DHYLICHPLH-YPLLMGHQWCLGFVLTLQLFGITVDGLVVIL
mm17M1-30   102 TFTSLDADEYFLLTLMAHDHCLAIFYSL-YPRLMRPQWCLGLVIIVWLSGFMEAGLVVAL
mm17M1-31P  102 TFSSLYIDEYFLLALMAYDHCLAICYSL-YPRLMRPQWCLGLVLTVWVSGFMEDGLVVAL
mm17M1-10   100 FFAVFADMDNLLLAVMAIDRYAAICHPLFYPFLMTPCRCEVLASGSWGIAHCVSLFYTLL
mm17M1-11   111 FFGVFADMDNLLLAVMAIDRYAAICHPLLYPLLMTPCRCEVLVSGSWGIAHCVSLMYTLL
mm17M1-26   101 FFAVFADMDNLLLAVMAIDRYAAICHPLLYPLLMTPCRCRVLVSGSWGVAHCVSLTHILL
mm17M1-27   100 FFAVFADMDNLLLAVMAIDRYAAICHPLLYPLLMTPCRCRVLVSGSWGVAHCVSLTHTLL
mm17M1-28   100 FFAVFADMDNLLLAVMAIDRYAAICHPLRYSALMTPFRCGVLVSGSWGVTNCVSLTHTLL
mm17M1-1    101 IFMSLGTTECILLTVMAFDRYVAVCQPLHYATKINPHLCRQLAGIAWAIGLVQSIVQTPP
mm17M1-5P   100 IFMSLGTTECILLTVMAFDRYVAVCQPLHYATVINPRLCQQLAGIAWAIGLVQSIVQTPP
mm17M1-2    101 IFMLLGTTECILLTVMAFDRYVAVCQPLHYATIIHPRLCRQLAGVAWAIGLVQSIVQIPP
mm17M1-4    101 IFLSLGTTECILLTVMSFDRYVAVCQPLHYATVIHPRLCWKLAAVAWMMGLLQSIVQTPP
mm17M1-3    101 IFLSLGTTECILLAVMSFDRYVAVCQPLHYATVIHPRLCCQLAAVACTIGLVESVVQTPS
mm17M1-21   102 VSLALGSTECVLLAVMAFDRYAAVCRPLHYATVMHPRLCQSLAGVAWLSGVGNTLIQGTI
mm17M1-45   101 VSLTLGSTECILLAVMAFDRYAAVCKPLHYVAVMNPQLCRALAGISWLSGIGNALIQGTI
mm17M1-6    177 SSLALGSAECVLLAVMALDRAAAVCNPLRYTSLASPLLCRTLAGVSWLGGLANSAAQTAL
mm17M1-39   103 IFLALGSTECILLAIMAFDRFAAICRPLHYPIIMNQKRCIHMATGTWISGFANSLVQSTL
mm17M1-13   101 FFHFLGSTEAMLLPVMAFDRFVAICRPLHYSVIMNHQLCIHMTVTIWTLGFFHALLHSVM
mm17M1-25   101 FFHFLGSTEAMLLPVMAFDRFVAICRPLHYSVIMNHQLCIHMTVTIWTLGFFHALLHSVM
mm17M1-12   101 FFHFLGSTETMLLPVMAFDRFVAICRPLHYSVIMNHQLCIHMTVTIWTLGFFHALLHSVM
mm17M1-14   101 FFHFLGTTESLLLAVMAFDRFVAICRPLHYSVIMNWQVCILMAVTIWTIAFLHALLHSVM
mm17M1-8    102 FFNFLGSTEALLLTVMAFDRFVAICRPLHYPAIMNSQVCIQVAISIWAIPFLHALVHSIL
mm17M1-9    102 FFHLLGSTEALLLAVMAFDRFVAICRPLHYPSIMNGQVCIQVAISIWAIPFVHALVHSIL
mm17M1-7P   102 FFHFLGCTDALLLTVMAFDRFVAICRPLHYPSIMNRQVCIQVAATIWAIPFLHALVHSIL
mm17M1-34   101 FFHFLGSTEAILLAVMAFDRFVAICSPLRYTAIMNPQLCILLAATAWFTSFFYALLHSVM
mm17M1-23   103 AFIFFVGSECLLLAAMAYDRYIAICKPLRYSFIMNKALCSWLAASCWTCGFLNSVLHTVL
mm17M1-33   166 AFIFFVGSECLLLAAMAYDRYIAICKPLRYSFIMNKALCSWLAASCWTGGFLNSVLHTVL
mm17M1-22   121 FFIFFGITECCLLAAMAFDCYMAICSPLHYSTRMSREVCAHLALVSWGMGCIVGLGQTNF
mm17M1-36   109 FFIFFGITECCLLAAMAFDRYMAICSPLHYATRMSREVCAHLAIVSWGMGCIVGLGQTNF
mm17M1-35   109 FFTFFGITECCLLAAMAYDRCMAICSPLHYPTRMSSGVCAHLAIVSWGMGCIVGLGQTNF
mm17M1-37   109 FFTFFGITECCLLAAMAYDRCMAICSPLHYATRMSHGVCAHLAIVSWGMGCIVGLGQTNF
mm17M1-38   109 FFIFFGITECCLLAAMAFDRYMAICSPLHYATRMSRGVCAHLAIVSWGMGCIVGLGQTNF
mm17M1-46   102 FFLFFGATECCLLAAMAYDRYAAICEPLRYQVLLSRRVCVQLAGAAWSCGALVGLGHTSF
mm17M1-19   128 FFTWFAWGEMAILTVMSYDRYIAVCLPLHYEIIMCPRKCRWAVTAVWLSSSIPGTLYLAT
mm17M1-20   128 FFTGFAWGEVAILTVMSYDRYVAVCLPLHYEVIMGPSKCRWAVTAVWLSSVIPGTLYIAS
mm17M1-24   128 FFTAFAWGELAILTVMSYDRYVAICLPLHYEVIMSPRKCTWAVATVWLSGGISGTLYITG
mm17M1-32   103 FFAAFAWGEVAILTVMSYDRYVAICLPLHYEVIMSPRKCTWAVTSVWLSSVIPGTLYIAS
mm17M1-18   128 FFTCFAWDEMAILTVMSYDRYVAVCLPLHYEVIMSPRKCTWALAAVWLSGGVSGTLYTAS
mm17M1-42   103 FFTGLAWSEMATLTVMSYDRYVAICLPLHYEVIMSPRKCTWAVAAVWLSGGISGTLFTAS
mm17M1-43   103 FFTGLAWSEVAILTVMSYDRYVAICLPLHYEVIMSPRKCTWAVAAVWLSGGISGTLFTAS
mm17M1-41   100 FFIALASSEVAILTVMSYDRYVAICRPLQYETIMDPHACKCAVIAVWMAGGLSGLLHTGV
mm17M1-40P   99 FYSQSSYTAMAILTVMSYDCYMAVWHKVITNVSTCIHGVLAVLVNVMN--YLWSYAHXQL
mm17M1-44P   61 LFCDWTXTKCVLLTMMAYNPVTPICWPLXCSPYYTPKICHTPKVSLEASCLGLDLFYGVH
```

```
mm17M1-15P   48 MAQLRFCVPK----QIDHFYYDFSPLVVLAYTDTGLVQVTTFVLFVVFLTVPFG-LVLIS
mm17M1-29   164 MAQLRFCGPN----LVDHFYCDFSPLMVLACSDTQVAQVTTFVLSVVFLTVPFG-LVLIS
mm17M1-16P   42 VAQMWFCGPN----LIDYFYNFSP--IMDLASDTQVFQVITFVLSVVFLTVPFG-LVLIS
mm17M1-17P   42 VAQMWFCGPN----LIDYFYNFSP--IMDLASDTQVFQVITFVLSVVFLTVPFG-LVLIS
mm17M1-30   161 TAQLRFCGPN----LIDHFYCDFSP-LMILACSDTVAQMTTFVLFVVFLPVLSG-LILMS
mm17M1-31P  161 IAQLRFCGPN----LIDHFYCDFS---PLMACFDTVAQMTTFVLSVIFLTVPFGXLVLIS
mm17M1-10   160 LSQFYYHTNQ----GIPHFFCDSRPLLLLSCS-DTHLSEGLMMALSGVLGMSSVLCLVSS
mm17M1-11   171 LSQLYFHTNQ----EIPRFFCDCRPLLLLSCS-DTHLNEVLMMALAGVLGVSAVLCIVSS
mm17M1-26   161 LSQLYFHTNQ----EIPHFFCDFGPLLLLSCS-DAHLNESLMMALAGVLGISALLCIVSS
mm17M1-27   160 FSKLYFHNNQ----EIPHFFCDFGPLLLLSCS-DTYLNESLMMALSGLLAISAFLCIVSS
mm17M1-28   160 LSKLYFHTNQ----EIPHFFCEFGPLLLLSCS-DTHLNKILVIILVGILGISAVLCIVSS
mm17M1-1    161 TLKLPFCSHR----QIDNFLCEVPSLIQLSCG-DTTYNEIQMAVASIFIVVVPLSLILVS
mm17M1-5P   160 TLKLPFCSHR----QIDNFVCEVPSLIQLSCG-DITYNEIQMAVASIFIVVVPLSLILVS
mm17M1-2    161 TLTLPFCSHR----QIDDFLCEVPSLIRLSCG-DTTFNEIQLSVAGVIFLLVPLSLIIVS
mm17M1-4    161 TLKLPFCPHR----QIDDFLCEVPSLIRLSCG-DTTFNEIQLAVSSVILVVVPLSLILVS
mm17M1-3    161 TLRLPFCPHH----QVDDFVCEVPALIRLSCG-DTTYNEIQMAVASVFILVVPLSLILVS
mm17M1-21   162 TLRLPRCGNH----KIYHFICEVPAMIKLACV-DIHANEVQLFMASLVLLLLPLTLILVS
mm17M1-45   161 TLWLPRCGHL----WLHHFFCEVPSMIKLACV-DIHANEVQLFVASLVLLLLPLALILTS
mm17M1-6    237 LAARPLCAPR----CLDHFICELPALLQLACRGGRSATERQMFAARVVILLVPSAVILAS
mm17M1-39   163 TVVAPRCGQR----VIDHFFCEVPALLKLACT-DTSVNEAELNVLGALLLLVPLSLILGT
mm17M1-13   161 TSRLSFCGPN----HVHHFFCDIKPLLDLACG-NTELNLWLLNTVTGTIALTPFFLTFLS
mm17M1-25   161 TSRLSFCGPN----HVHHFFCDIKPLLDLACG-NTELNLWLLNTVTGTIALTPFFLTFLS
mm17M1-12   161 TSRLSFCGPN----HVHHFFCDIKPLLDLACG-NTELNLWLLNTVTGTIALTSFFLIFLS
mm17M1-14   161 TSRLSFCGLN----HIHHFFCDVKPLLELACG-NTELNLWLLNTVTGTIASVPFFLTFLS
mm17M1-8    162 TSQLNFCGSN----HIHHFFCDVKPLLELACG-NTELNRWLLNTLTGTVAIGLFFLTFLS
mm17M1-9    162 TSQLNFCGSN----QIHHFFCDVKPLLELACG-NTELNRWLLNTFTGTFAIGLFFLTFLS
mm17M1-7P   162 TSQLNFCGSN----RIHHFFCDVKPLLELACG-NTELNRWLLNTLAGTIGIGLFFLTFLS
mm17M1-34   161 TAHLNFCHSH----KLSHFFCDVKPLLEVACG-NTVLNQWLLSVVTGSISMGAFLLILLS
mm17M1-23   163 TFHLPFCGNN----QINYFFCDIPPLLILSCG-DTSLNELALLSIGILIGWTPFLCIILS
mm17M1-33   226 TFHLPFCGNN----QINYFFCDIPPLLILSCG-DTSLNELALLSIGILIGWTPFLCIILS
mm17M1-22   181 IFSLNFCGPC----EIDHFFCDLPPVLALACG-DTSQNEAAIFVAVVLCISSPFLLIIYS
mm17M1-36   169 IFSLNFCGPC----EIDHFFCDLPPVLALACG-DTSQNEAAIFVTVVLCISSPFLLIIYS
mm17M1-35   169 IFSLEFCGPC----EIDHFFCDLPPVLALACG-DTSQNEAAIFVAAVLCISSPFLLIIYS
mm17M1-37   169 IFSLNFCGPC----EIDHFFCDLPPVLALACG-DTSQNEAAIFVAAILCISSPFLLIIYS
mm17M1-38   169 IFSLNFCGPC----EIDHFFCDLPPVLALACG-DTSQNEAAIFVAAVLCIFSPFLLIISS
mm17M1-46   162 IFSLPFCGPN----AVPHFFCEIQPVLQLVCG-DTSLNELQIILAAALIILCPFGLILSS
mm17M1-19   188 IFSIRICRAK----IIHQFFCDVPQLLKLSCS-NDYLVIMGVADFLSVIGFACFVGIVIS
mm17M1-20   188 IFSIRFCGDR----IIHQFFCDVPQVLKLSCS-DDYLVTVGVADFLSAVAFACFIGIVNS
mm17M1-24   188 TLFIRFCGDK----IIHQFFCDVPQLLKLSCS-NDHLVIMDMVSFLTAVSFACFTGIVIS
mm17M1-32   163 IFSIRFCRAK----IIHQLFCDVPQLLKLSCS-NDHLVVIGMVSFMTAVAFACFVGIVIS
mm17M1-18   188 TLSIRFCGDR----IIHQFFCDVPQVLKLSCS-NDYLLTIGVANILSAVAFACFIGIVIS
mm17M1-42   163 TLSIRFCGDK----IIHQFFCDIPQLLKLSCS-NDYFGVLEVSTFMSVMAFACFVGIAFS
mm17M1-43   163 TLSIRFCGHK----IIHQFFCDIPQLLKLSCS-NDDFGLLKVSTFIAVMGFACFVGIAFS
mm17M1-41   160 NFSIPLCGKR----IIHQFFCDIPQMLKLACS-YEFINEIAVAAFTTSTAFVCLIAIVFS
mm17M1-40P  157 LRLHLHCG------TSTIRFCDVLLVLKLSFT-NDHVNELESLAX---------------
mm17M1-44P  121 NSDHSGFSIISVPPQNEFFMCEESPLVKITFMDTTSLEKHISVFT--FLAVIPCGEYSII


mm17M1-15P  103 CAQIAVTVLR-VPSRTRRNKAFSTCSSHLDEVSTFYGSLMVWYTEPSAVHS--QILSKVI
mm17M1-29   219 YAQIVVTVLR-VPSGTRRTKAFSTCSSHLAVVSTFYGTLMVLYIVPSAVHS--QLLSKVI
mm17M1-16P   95 YIQIVVTVLR-VLSGDRRTKDFSTCSSHLAVVSTFYRSLMVLYTVPFAPX-----LSKVI
mm17M1-17P   95 YIQIVVTVLR-VLSGDRRTKDFSTCSSHLAVVSTFYRSLMVLYTVPFAPX-----LSKVI
mm17M1-30   215 YAQFVVIVLR-IPSGARRTKAFFTCSSHLAMMFTFYGSLMVWYTAPSAVXLVCTLLSKVI
mm17M1-31P  214 YAQLVVTVLR-ILSGARRTKAFVICSSHLAVVSTLYGTLMGLYTVPFVV--HSQLLTKVI
mm17M1-10   215 YGCIFYAVAR-VPSAQGKRKSLATCSSHLSVVLLFYSTVFATYLKPPST--SHSSAEVVA
mm17M1-11   226 YGCIFYAVAR-VPSAQGKRKALTTCSSHLSVVLLFYSTVFATYLKPPST--SHSSGEVVA
mm17M1-26   216 YGCIFYAVAK-VPSAQGKRKALATCSSHLSVVLLFYSTVFATYLKPPSS--SRSSGEVVA
mm17M1-27   215 YGCIFYAVAK-VPSAQGKRKALATCSSHLSVVLLFYSTVFATYLKPPSS--SHSSQEVVA
mm17M1-28   215 YGCIFYAVAK-VPSAQGKRKALSTCSSHLSVVLLFYSTVFATYLKPPSS--SRSSEEVVA
mm17M1-1    216 YGAIARAVLK-ISSAKGRRKAFGTCSSHLIVVTLFYSSVIAVYLQPKNP--YARERGKFF
mm17M1-5P   215 YGAIARAVLK-ISSAKGRRKAFGTCSSHLIVVTLFYSSVIAVYLQPKNP--YARERGKFF
mm17M1-2    216 YGVIARAVLK-TNSSKGRRKAFGTCSSHLIVVTLFYSSVIAVYLQPKNP--YAQERSKFF
mm17M1-4    216 YGAIARAVMR-INSTEAWKKALRTCSSHLIVVTLFYSSVIAVYLQPKNP--YAQERGKFF
mm17M1-3    216 YGAIARAVLR-ISSAKGRRKAFGTCSSHLIVVTLFYSSVIAVYLQPKNP--YARERGKFF
mm17M1-21   217 YGYIAQALMR-LRSALTWGKALGTCGSHLIVVVLFYGTSTAVYIHPNSS--YAQSQGKFI
mm17M1-45   216 YGHIAKAVIR-IKSSQAWRRALGTCGSHLMVVSLFYGSITAIYIQPNSS--YAHTHGKFI
mm17M1-6    293 YIAVGRAVWG-MHSSSGWRKAASTCGSHLTAVCLFYGSATYTYLQPTHS--YNQGRGKFV
mm17M1-39   218 YVFIAQAVLK-LRSAESRRKAFNTCASHLLVVSLFYFTAISMYVQPPSS--YSHERGKIM
mm17M1-13   216 YFYIITYLFLKTRSCSMLHKALSTCASHFMVVILLYVPVLFTYIRPASG--SSLDQDRII
mm17M1-25   216 YFYIIIYLLLKTRSCSMLHKALSTCASHFMVVFLFYAPVLFIYISPTSG--SSLDQDRII
mm17M1-12   216 YFYIITNLLLKTRSCSMLHKALSTCASHFMVVVLFYAPVLFTYIRPASG--SSLDQDTII
mm17M1-14   216 YFYIITYLFLKTRSCSMLHKALSTCASHFMVVVLFYAPVLFTYIRPTSG--SSLDQDRII
mm17M1-8    217 YFYIVTYLFLKTRSCSMLHKALSTCASHFMVVMIFYAPVLFIYINPDSG--SSLEKDRII
mm17M1-9    217 YFYIITYLFLKTRSCSMLHKALSTCASHFMVVMIFYAPVLFIYINPDSG--SSLEKDRII
mm17M1-7P   217 YFYIVTYLFLKTHSCSMLHKALSTCASHFMVVMIFYAPVLFIYINPDSG--SSLEKDRII
mm17M1-34   216 YFYIIAFLLFKNRSCRMLKKALSTCTSHFMVVCLFYGPVGFTYIRPATASASSMSEDRVV
mm17M1-23   218 YLYIISTILR-IRSSEGRQKAFSTCASHLLIVILYYGSAIFTYVRPISS--YSLEKDRLI
mm17M1-33   281 YLYIISTILR-IRSSEGRHKAFSTCASHLLIVILYYGSAIFTYVRPISS--YSLEKDRLI
mm17M1-22   236 YVRILVAVLV-MPSPEGRHKALSTCSSHLLVVTLFFGSGSITYLRPKSS--HLPGMDKLL
mm17M1-36   224 YVRILFAVLV-MPSPEGRHKALSTCSSHLLVVTLFYGSASITYLRPKSS--HSPGIDKLL
mm17M1-35   224 YVRILVAVLL-MPSPEGRHKALSTCSSHLLVVTMFYGSASITYLRPKSS--HSPGMDKLL
mm17M1-37   224 YVRILVAVLV-MPSPEGRHKALSTCSSHLLVVTLFFGSGSITYLRPKSS--HLPGMDKLL
mm17M1-38   224 YVRILIAVLV-MPSREGRHKALSTCSSHLLVVTLFYGSTSATYLRPKSD--HSPEVDKLL
mm17M1-46   217 YGRILVTIFR-IPSAAGRRKAFSTCSSHLVVVSLFYGTAIFIYIRPKAS--YDPTTDPLL
mm17M1-19   243 YVHIFSTVLR-MPSAESRSKVFSTCLPHLFVVSLFLSTGIFAYLNPTSD--FPTALEFLF
mm17M1-20   243 YVHIFSTVLR-MPSAESRSKVFSTCLPHLFVVLLFLSTGIFAYLNPTSD--SPTALQFLV
mm17M1-24   243 YVHIFSTVLR-MPSAESRSKVFSTCLPHLFVVSLFLSTGAFAYLNLTSD--SSTALEFLL
mm17M1-32   218 YVHIFSTVLR-MPSAESRSKVFSTCLPHLFVVSLFLSTGSCAYLNTSSD--SPTALEFLF
mm17M1-18   243 YVHIFSTVLR-MPSAESRYKVFSTCLPHLFVVSLFLSTSTFAYLNPTAD--SPTALEFLF
mm17M1-42   218 YGQIFSTVLR-MPSAEGRSKVFSTCLPHLFVVSFFLSTGICAYLKPTSD--SPTALDLML
mm17M1-43   218 YCQIFSTVLR-MPSAEGRSKVFSTCLPHLFVVSFFLSTGICAYLKPSSD--SPTALDLML
mm17M1-41   215 YTQIFSTVMR-IPSADSRTKVFSTCLPHLFVVMFFLSAAGFEFLRPPSD--SLSAMDLVF
mm17M1-40P  195 -----SHLWR-------AEPVFLTCLGHVSVGSLFNPPGVFEFLNPYSE---SPTSLDII
mm17M1-44P  179 YLLVLLKVWLKIKFTG-RMKTFGSCGFHLMAIVLFFGNESSVYMVYMYPRANACQYRKFL
```

```
mm17M1-15P   160  ALLYTVVTTIFDPGIYTLRNQEVQQSLRRHLYCKPTEM----------- 197
mm17M1-29    276  ALLYTVVTPIFNPVIYTLRNQEVQQALRRLLYCKPTEM----------- 313
mm17M1-16P   149  ALLYKVVIPIFNHVIYTLRNQEVP------------------------- 172
mm17M1-17P   149  ALLYKVVIPIFNHVIYTLRNQEVP------------------------- 172
mm17M1-30    274  ALLYTVFAPIFNSVIYTLRNLDMQKALRRLLYCKSTEM----------- 311
mm17M1-31P   271  ALLYTVFTPIFNPVIYTLKNQEVQQALRRLLYC---------------- 303
mm17M1-10    272  AVMYTLVTPTLNPFIYSLRNKDVKSSLRKILNMDKFQG----------- 309
mm17M1-11    283  AVMYTLVTPTLNPFIYSLRNKDVKSSLRRVLNIEKSQD----------- 320
mm17M1-26    273  AVMYTLVTPTLNPFIYSLRNKDVKSSLRRILNMVKSQD----------- 310
mm17M1-27    272  AVMYTLVTPTLNPFIYSLRNKDVKSSLRRILNMVKSQD----------- 309
mm17M1-28    272  AVMYSLVTPTLNPFIYSLRNKDVKSSLRRILNME--------------- 305
mm17M1-1     273  GLFYAVGTPTLNPLVYTLRNKEVKRAFWKLLRKDEDSEES--------- 312
mm17M1-5P    272  GLFYAVGTPILNPLVYTLRNKEVKRAFWKLLRKDEDSEES--------- 311
mm17M1-2     273  GLFYAVGTPTLNPLVYTLRNKEVKRAFWRLLGKDAASGRN--------- 312
mm17M1-4     273  GLFYAVGTPTLNPLIYTLRNKEVKRAFWRLLGKDGDSKNT--------- 312
mm17M1-3     273  GLFYAVGTPSLNPLIYTLRNKEVKRAFRRLLWKEVKPS----------- 310
mm17M1-21    274  TLLYTVVIPTLNPLIYTLRNKDVKGALKRLVRKDSSTGKKILSR----- 317
mm17M1-45    273  SLFYTVMTPTLNPLIYTLRNKEVKGALGRLFNRASGV------------ 309
mm17M1-6     350  SLFYTVVTPALNPLIYTLRNKEVKGAALRLLRSLGRP------------ 386
mm17M1-39    275  ALFYGIVTPTLNPFIYTLRNKDVKAALRRALTKEFWVKARQ-------- 315
mm17M1-13    274  AIMYSVVTPALNPLIYTLRNKEVRSALNRKVRRWL-------------- 308
mm17M1-25    274  AIMYSVVTPALNPLIYTLRNKEVRSALNRKLRRWL-------------- 308
mm17M1-12    274  AIMYSVVTPALNPLIYTLRNKEVRSALNRKVRRWL-------------- 308
mm17M1-14    274  AIMYSVVTPALNPLIYTLRNKEVRSALNRKVRRCLLLEEI--------- 313
mm17M1-8     275  AVMYTVVTPALNPLIYTLRNKEVRGALNRKIRILL-------------- 309
mm17M1-9     275  AVMYTVVTPALNPLIYTLRNKEVRGALNRKLRILL-------------- 309
mm17M1-7P    275  AVMYTVVTPALNPLIYALRNKEVRCALNRKLRILI-------------- 309
mm17M1-34    276  AIIYSAVTPVLNPLIYTLRNKEVMLALKKNFGKKLFKGN---------- 314
mm17M1-23    275  SVLYSVFTPMLNPIIYALRNKDIKEAVKAIGRKWQPPVFSSDM------ 317
mm17M1-33    338  SVLYSVVTPMLNPVIYTLRNKDIKEAVKAIGRKWQPPVFSSDI------ 380
mm17M1-22    293  ALFYTAVTSMLNPIIYSLRNKEVKAALRKTLSLKTSRAINR-------- 333
mm17M1-36    281  ALFYTAVTSMLNPIIYSLRNKEVKAALRRTLSLKKPLAINR-------- 321
mm17M1-35    281  ALFYTAVTSMLNPIIYSLRNKEVKAALRKTLSLKKPLAINR-------- 321
mm17M1-37    281  ALFYTAVTSMLNPIIYSLRNKEVKTALRKTLSLKTSRAINR-------- 321
mm17M1-38    281  ALFYTAVTSMLNPIIYSLRNKEVKAALRKTLSLKKVLIMNR-------- 321
mm17M1-46    274  SLFYAVITPILNPVIYSLRNADVKAALKRSIQKMGPSEI---------- 312
mm17M1-19    300  SVFYTVLPPTLNPVIYSLRNDAIKSVVRKLLLSRKFTS----------- 337
mm17M1-20    300  SIFYTVLPPTLNPVIYSLRNETIKSVIRKLLLSSKFTG----------- 337
mm17M1-24    300  SIFYTVLPPTLNPVIYSLRNETIKNVVRKLLLSTKFTVRIIFSCCF--- 345
mm17M1-32    275  SIFYTVLPPTLNPVIYSLRNETIKSVVRKLLLSSKFTVRIICPVATD-- 321
mm17M1-18    300  SILYTVLPPTINPVIYSLRNETIKSVVRKLLSSTKFTV----------- 337
mm17M1-42    275  SIFYTLLPPTLNPVIYSLRNESLKRALKKLLLSEEFIRKKCLFYF---- 319
mm17M1-43    275  SIFYTVLPPTLNPVIYSLRNESLKRAVKKLLLSEEFIGKNYVCSVFSAC 323
mm17M1-41    272  SIFYTVIPPTLNPLIYSLRNEAMKAALRKVLSKEEFSRRMVYVKAIFNL 320
mm17M1-40P   240  VKXVFILPQTLSVEIYSLSNEAIDTA----------------------- 265
mm17M1-44P   238  SXFYMIVTPSINPLIY-LRNKEFRWAVQRLVTRDPS------------- 272
```

Appendix 10: List of mouse ORs from the UCSC assembly

Name	Contig	Comments	Closest MHC-linked human OR	%age similarity to human MHC-linked OR.
mm13M1-1	13.20000001-13.25000000		hs6M1-22P	60
mm13M1-2P	13.20000001-13.25000000		hs6M1-15	85
mm13M1-3	13.20000001-13.25000000		hs6M1-15	86
mm13M1-4	13.20000001-13.25000000		hs6M1-8P	81
mm13M1-5	13.20000001-13.25000000		hs6M1-35	85
mm13M1-6	13.20000001-13.25000000		hs6M1-34P	85
mm13M1-7	13.20000001-13.25000000		hs6M1-31P	81
mm13M1-8	13.20000001-13.25000000		hs6M1-32	81
mm13M1-9	13.20000001-13.25000000		hs6M1-34P	94
mm13M1-10F	13.20000001-13.25000000		hs6M1-10	90
mm13M1-11	13.20000001-13.25000000		hs6M1-10	87
mm13M1-12	13.20000001-13.25000000		hs6M1-10	85
mm13M1-13	13.20000001-13.25000000		hs6M1-30P	83
mm17M1-47	17.40000001-17.45000000		hs6M1-7P	78
mm17M1-48	17.40000001-17.45000000		hs6M1-2P	83
mm17M1-49P	17.40000001-17.45000000	1 FS	hs6M1-2P	82
mm17M1-50P	17.40000001-17.45000000	1 FS	hs6M1-2P	82
mm17M1-51P	17.40000001-17.45000000	1 FS	hs6M1-3 hs6M1-4	81
mm17M1-52P	17.40000001-17.45000000	1 FS, 1 insertion, 4 stops, no methionine	hs6M1-3 hs6M1-6	49
mm17M1-53	17.40000001-17.45000000		hs6M1-2P	83
mm17M1-54	17.40000001-17.45000000		hs6M1-17	54
mm17M1-55	17.40000001-17.45000000		hs6M1-25P	72
mm17M1-56	17.40000001-17.45000000		hs6M1-32	61

Appendix 11: Human 'ROLF' database

Name	State	Accession number	Position in acc. No.	Comments
hs1M1-1	C	AL390860	65000..67000	
hs1M1-2	C	AL365440	43000..45000	
hs1M1-3	F	AL356607	59000..61000	goes into gap
hs1M1-4	C	AL356607	68000..70000	
hs1M1-5	P	AL365440	92000..94000	5 FS, 6 stops
hs1M1-6	C	AL513488	143000..145000	
hs1M1-7	C	AC024654	8000..10000	
hs1M1-8	C	AC024654	14000..16000	
hs1M1-9	C	AC024654	24000..26000	
hs1M1-10	P	AC024654	73500..75500	1 stop
hs1M1-11	P	AC024654	104500..106500	1 stop
hs1M1-12	C	AC024654	153500..155500	
hs1M1-13	C	AC025115	20500..22500	
hs1M1-14	P	AC025115	26000..28000	1 stop
hs1M1-15	P	AC025115	66500..68500	1 FS
hs1M1-16	C	AC025115	89000..91000	
hs1M1-17	C	AL356607	92500..94500	
hs1M1-18	C	AL356607	110500..112500	
hs1M1-19	P	AL354713	64000..66000	1 stop
hs1M1-20	P	AL357039	4000..6000	1 FS
hs1M1-21	C	AL357039	15000..17000	
hs1M1-22	C	AL357039	34000..36000	
hs1M1-23	P	AL357039	48000..50000	Missing start
hs1M1-24	C	AL357039	116000..118000	
hs1M1-25	P	AL357039	121000..123000	4 FS,stops
hs1M1-26	C	AL357039	149500..151500	
hs1M1-27	C	AL357039	171000..173000	
hs1M1-28	F	AC026038	126000..128000	goes into gap
hs1M1-29	P	AL358773	11500..13500	1 FS
hs1M1-30	P	AL513488	97500..99500	1 FS
hs1M1-31	C	AL513488	81500..83500	
hs1M1-32	C	AL358773	60500..62500	
hs1M1-33	C	AB045359	156000..158000	
hs1M1-34	P	AB045359	92000..94000	No Met, 5 FS,
hs1M1-34	C	AL513488	59000..61000	
hs1M1-35	C	AB045359	140000..142000	
hs1M1-36	C	AL358874	12000..14000	
hs1M1-38	C	AB045360	41500..43500	
hs1M1-39	P	AB045360	129000..131000	1 FS, 1 stop
hs1M1-40	C	AB045360	152000..154000	
hs1M1-41	P	AB045360	158500..160500	No Met
hs1M1-42	C	AB045360	134000..136000	
hs1M1-43	C	AB045361	14000..16000	2 stops, Alu insertion

hs1M1-44	C	AB045361	26000..28000	
hs1M1-45	C	AB045361	37000..39000	
hs1M1-46	C	AB045361	48000..50000	
hs1M1-47	P	AB045361	80000..82000	1 FS
hs1M1-48	P	AB045365	87500..89500	No Met, 1 FS
hs1M1-49	P	AB045365	122000..124000	1 FS
hs1M1-50	C	AL391534	101500..103500	
hs1M1-51	C	AL358874	84500..86500	
hs1M1-52	P	AL513323	144500..146500	1 FS
hs1M1-53	P	AL513323	159500..161500	1 stop, 1 FS
hs1M1-54	P	AL513323	198500..200500	4 stops, 2 FS
hs1M1-55	P	AC016787	13000..15000	1 stop
hs1M1-56	C	AC016787	134000..136000	
hs1M1-57	C	AL358874	102000..104000	
hs1M1-58	C	AL358874	122500..124500	
hs1M1-59	F	AL358874	133000..135000	goes into gap
hs1M1-60	C	AL358874	149000..151000	
hs1M1-61	C	AL391534	8000..10000	
hs1M1-62	C	AL513488	36000..38000	
hs1M1-63	C	AL513488	1500..3500	
hs1M1-64	C	AL391534	140000..142000	
hs1M1-65	C	AL391534	62000..64000	
hs1M1-66	C	AL450303	10500..12500	
hs1M1-67	C	AL450303	46500..48500	
hs1M1-68	C	AL450303	80000..82000	
hs1M1-69	C	AL450303	102000..104000	
hs1M1-70	C	AL450303	131000..133000	
hs1M1-71	PF	AL391904	40500..42500	No Met, 1 FS, missing motifs
hs1M1-72	P	AL391904	67000..69000	2 FS, 2 stops

hs2M1-1	P	AC009237	70000..72000	no Met
hs2M1-2	P	AC007881	21500..23500	3 FS,2 stop, no Met
hs2M1-3	P	AC007881	4000..6000	1 FS, 3 stops
hs2M1-4	F	AC069348	4000..6000	goes into gap
hs2M1-5	P	AC007040	177000..179000	1 FS,2 stop, no Met
hs2M1-6	P	AC064843	37500..39500	1 FS, 3 stops
hs2M1-7	P	AC064843	105000..107000	5 stops
hs2M1-8	P	AC064843	61500..62500	2 FS,goes into gap
hs2M1-9	P	AC013469	106000..108000	2 FS, stop
hs2M1-10	P	AC013469	21000..23000	3 FS
hs2M1-11	P	AC013469	45000..47000	4 FS
hs2M1-12	PF	AC013469	74000..76000	1 FS, no start

hs3M1-1	F	AC022049	12500..14500	goes into gap
hs3M1-2	P	AFO42089	47297..48332	
hs3M1-3	P	AFO42089	61204..62209	
hs3M1-4	P	AFO42089	69892..70886	
hs3M1-5	P	AC023058	70000..72000	stops,no met

hs3M1-6	P	AC024169	17500..19500	1 FS, 1 stop
hs3M1-7	P	AC024709	89500..91500	2 stop,3 FS
hs3M1-8	P	AC034187	21000..23000	2 stop,1 FS
hs3M1-9	P	AC067827	16000..18000	2 stop,1 FS
hs3M1-10	P	AC069518	41500..43500	3 FS, 1 stop
hs3M1-11	P	AC024892	56500..58500	1 FS
hs3M1-12	PF	AC024892	144000..146000	2 FS,missing start
hs3M1-13	C	AC025942	29500..31500	
hs3M1-14	P	AC025942	43500..45500	1 stop
hs3M1-15	C	AC025942	109000..111000	
hs3M1-16	P	AC025942	15000..17000	3 FS, 1 stop
hs3M1-17	P	AC025942	68000..70000	1 FS
hs3M1-18	P	AC025942	135500..137500	1 FS
hs3M1-19	P	AC025942	150000..152000	2 FS,short tail
hs3M1-20	C	AC074274	104500..106500	
hs3M1-21	C	AC074274	120500..124500	
hs3M1-22	P	AF186996	48000..50000	1 FS, 1 stop
hs3M1-24	P	AF186996	60000..62000	no met, stops
hs3M1-25	P	AF186996	66500..68500	no met, stops
hs3M1-26	P	AF186996	76500..78500	3 FS, stops
hs3M1-27	P	AC069047	92000..94000	1 FS, stops
hs3M1-28	P	AC069047	104500..106500	no met, stops
hs3M1-29	P	AC069047	174000..176000	no met,stops
hs3M1-30	P	AF165423	58000..60000	3 FS, 2 stops
hs3M1-31	P	AF165423	71000..73000	1 FS, no Met, stops

hs4M1-1	P	AC015709	82000..84000	3 FS, 1 stop
hs4M1-2	P	AC022674	33000..35000 stops	
hs4M1-3	P	AC022674	86500..88500	no met, 2 stop, 1 FS
hs4M1-4	P	AC022674	115500..117500	1 FS, no met, stops
hs4M1-5	P	AC068403	149500..151500	2 FS, 1 stop
hs4M1-6	P	AC007310	114000..116000	3 FS,stops, unfinished end
hs4M1-7	PF	AC013359	172000..174000	Alu insertion, no end, 4 stops
hs4M1-8	PF	AC013662	108000..110000	1 FS, 1 stop, missing end
hs4M1-9	P	AC011744	157000..159000	1 FS, 3 stop, no met
hs4M1-10	P	AC011744	175000..177000	2 FS, 2 stop
hs4M1-11	P	AC011744	96000..97000	no met, goes into gap
hs4M1-12	PF	AC008374	6000..8000	goes into gap, 1 FS(100% 16M1-1 no clone overlap)
hs4M1-13	P	AC022674	108000..110000	stops, 3 FS, no met

hs5M1-1	C	AC008620	87000..89000	
hs5M1-2	P	AC008620	119000..121000	1 FS
hs5M1-3	P	AC008454	47500..49500	Alu insertion,no met,4 stops
hs5M1-4	C	AC023255	180000..182000	
hs5M1-5	P	AC025336	160000..162000	1 FS,1 stop

For MHC-linked ORs, refer to Appendix 6

hs6M1-11	P	AL031259	8399..9336	1 FS
hs6M1-36	C	AL135904	21000..23000	
hs7M1-1	C	AC004853	17162..18115	
hs7M1-2	C	AC004853	41900..42853	
hs7M1-3	C	AC004853	85926..86861	
hs7M1-4	C	AC076959	40500..42500	
hs7M1-5	PF	AC076959	11000..12000	1 stop
hs7M1-6	C	AC076959	26000..28000	
hs7M1-7	C	AC076959	57000..59000	
hs7M1-8	P	AC076959	79500..81500	1 stop
hs7M1-9	C	AC073647	108000..110000	
hs7M1-10	P	AC004853	62834..63764	1 FS
hs7M1-11	P	AC004967	113880..114883	2 FS, 1 stop
hs7M1-12	P	AC004967	133062..134010	2 FS, 2 stop
hs7M1-13	P	AC004889	8842..9772	1 FS
hs7M1-14	C	AC004889	84457..83528	
hs7M1-15	P	AC004889	103217..102288	1 stop
hs7M1-16	P	AC076959	94000..96000	No Met, 2 FS
hs7M1-17	C	AC073647	164500..166500	
hs7M1-18	P	AC076959	123000..125000	4 FS, 3 stops
hs7M1-19	F	AC076959	136000..137000	
hs7M1-20	P	AC076959	140000..142000	1 FS
hs7M1-21	P	AC073647	140000..142000	1 stop
hs7M1-22	P	AC079804	89500..91500	3 FS, 1 stop
hs7M1-23	P	AC079804	77000..79000	No Met, 2 FS, 1 stop
hs7M1-24	P	AC079882	68000..70000	No Met, 2 FS, 1 stop
hs7M1-25	P	AC027522	39500..41500	1 FS, 1 stop
hs7M1-26	P	AC079882	207000..209000	3 FS, 1 stop
hs7M1-27	P	AC073264	60000..62000	1 FS
hs7M1-28	C	AC073079	9000..12000	
hs7M1-29	C	AC004977	101000..103000	
hs7M1-30	C	AC073079	69000..71000	
hs7M1-31	P	AC073264	37000..39000	1 FS
hs8M1-1	P	AC025126	94000..96000	4 stops, 1 FS
hs8M1-2	P	AC000385	125709..126761	2 stops, no met
hs8M1-3	P	AC000385	138500..140500	4 stops
hs8M1-4	P	AC000385	145952..147061	2 FS, 4 stops
hs8M1-5	P	AC015480	45000..47000	1 stop, goes into gap
hs9M1-1	C	AC006313	2200..3000	
hs9M1-2	C	AC006313	16000..17000	
hs9M1-3	C	AC006313	49000..50000	
hs9M1-4	C	AC006313	62500..63500	
hs9M1-5	C	AC006313	111000..112000	
hs9M1-6	C	AC006313	136000..137000	3 starts

hs9M1-7	C	AC006313	176000..177000	
hs9M1-8	C	AL135841		
hs9M1-9	C	AL135841		
hs9M1-10	P	AL135841		1 FS
hs9M1-11	P	AL138834		
hs9M1-12	P	AL138834		
hs9M1-13	P	AL138834		1 FS, 2 stops
hs9M1-14	F	AL138834		
hs9M1-15	C	AL353767	196500..198500	
hs9M1-16	C	AL162254	19000..21000	
hs9M1-17	C	AL162254	57500..59500	
hs9M1-18	C	AL162254	67000..69000	
hs9M1-19	P	AL162254	92500..94500	2 FS
hs9M1-20	C	AL162254	115000..117000	
hs9M1-21	C	AL162254	167500..169500	
hs9M1-22	C	AL359512	30000..32000	
hs9M1-23	C	AL354661	61000..63000	
hs9M1-24	P	AL354862	1..2000	1 FS,2 stops
hs9M1-25	P	AC009900	80500..82500	3 FS, 2 stop
hs9M1-26	P	AC009900	177500..179205	no met, stops
hs9M1-27	P	AC010612	118000..120000	3 FS, stops
hs9M1-28	C	AC072059	36500..38500	
hs9M1-29	C	AC072059	58500..60500	
hs9M1-30	C	AC072059	68000..70000	
hs9M1-31	C	AC072059	101500..103500	
hs9M1-32	P	AC072059	122500..124500	1 FS
hs9M1-33	C	AC072059	131000..133000	
hs9M1-34	C	AC072059	137000..139000	
hs9M1-35	C	AC072059	149500..151500	
hs9M1-36	P	AC072059	162000..164500	Alu insertion,1 FS, stops
hs9M1-37	P	AC072059	189000..191000	1 stop
hs9M1-38	C	AC009594	114000..116000	
hs9M1-39	P	AC009594	135932..136865	

hs10M1-1	P	AL358394	199500..201500	1 stop
hs10M1-2	C	AC011879	42500..44500	
hs10M1-3	PF	AC011879	10000..12500	1 FS,1 stop, goes into gap
hs10M1-4	P	AL157391	6000..8000	4 FS,2 stops
hs10M1-5	P	AL157391	14500..16500	stops
hs10M1-6	P	AL360083	173500..175500	1 FS,2 stops, missing last domains

hs11M1-1	P	AC016856	83000..85000	3 FS
hs11M1-2	C	AC016856	108000..110000	
hs11M1-3	P	AC068339	154000..155482	2 stops, goes into gap
hs11M1-4	P	AC022289	30500..32500	5 FS
hs11M1-5	C	AC027239	116500..118500	
hs11M1-6	C	AP001998	11500..13500	
hs11M1-7	P	AC027239	135000..137000	1 stop

hs11M1-8	P	AC023080	500..2500 1 FS, 1 stop	
hs11M1-9	C	AC023080	121000..123000	
hs11M1-10	C	AC023080	51000..53000	
hs11M1-11	PF	AC023080	141000..143000	1 FS, missing start
hs11M1-12	C	AC023080	161500..163500	
hs11M1-13	C	AC026975	38000..40000	
hs11M1-14	C	AC026975	95000..97000	
hs11M1-15	C	AC026975	140000..142000	
hs11M1-16	P	AC026975	158000..160000	3 FS, 1 stop
hs11M1-17	P	AC005729	81000..83000	2 FS, 1 stop
hs11M1-18	P	AC025730	135000..137000	1 FS
hs11M1-19	P	AC021427	11000..13000	1 FS
hs11M1-20	P	AC021427	55500..57500	1 FS
hs11M1-21	C	AC021427	78500..80500	
hs11M1-22	C	AC021427	107500..109500	
hs11M1-23	P	AC022998	73000..75000	1 FS, 1 stop
hs11M1-24	P	AC021427	27500..29500	
hs11M1-25	C	AC044810	53500..55500	
hs11M1-26	C	AC044810	86000..88000	
hs11M1-27	P	AC060812	175500..177500	2 stops, no met
hs11M1-28	P	AP001998	20500..24000	L1 insertion, 10 stops
hs11M1-29	C	AC002555	49657..50631	
hs11M1-30	C	AF137396	5387..6325	AAD29425.1
hs11M1-31	P	AF137396	13990..14962	-
hs11M1-32	C	AF137396	27728..28660	AAD29426.1
hs11M1-33	P	AC002555	26414..27316	
hs11M1-34	P	AC002555	12222..13091	
hs11M1-35	P	AC002555	10218..11147	
hs11M1-36	P	AC000378	55912..56936	
hs11M1-37	C	AC018700	1415..2374	
hs11M1-38	C	AC018700	23342..24409	
hs11M1-39	C	AC018700	31182..32114	
hs11M1-40	P	AC018700	88726..89666	
hs11M1-41	P	AC016856	139500..141500	
hs11M1-42	PF	AC018700	136-138	goes into a gap currently
hs11M1-43	C	AC018700	151916..152893	
hs11M1-44	C	AC018700	200871..201815	
hs11M1-45	P	AC017103	5500..7500	
hs11M1-46	C	AC017103	33500..35500	
hs11M1-47	C	AC017103	44500..46500	
hs11M1-48	C	AC017103	79000..81000	
hs11M1-49	C	AC017103	122500..124500	
hs11M1-50	C	AF321237	49500..51500	
hs11M1-51	P/C	AC019108	10500..12500	1 FS
hs11M1-52	F	AC019108	21000..23000	
hs11M1-53	PF	AC019108	52500..54500	
hs11M1-54	C	AC019108	56500..58500	
hs11M1-55	P	AC019108	108500..110500	

hs11M1-56	C	AC019108	128000..130000	
hs11M1-57	C	AP002407	155000..157000	
hs11M1-58	P	AC019108	152500..154500	1 FS
hs11M1-59	C	AP002407	60000..62000	
hs11M1-60	C	AP002407	57000..59000	
hs11M1-61	C	AP003034	40500..42500	
hs11M1-62	P/C	AC022882	62500..64500	
hs11M1-63	P/C	AC022802	93500..95500	1 FS
hs11M1-64	C	AC022882	100000..102000	
hs11M1-65	P	AC022882	106500..108500	1 FS, no met
hs11M1-66	C	AC022882	170500..172500	
hs11M1-67	C	AC022882	156500..158500	
hs11M1-68	C	AC022882	128500..130500	
hs11M1-69	P	AP003034	122000..124000	1 stop
hs11M1-70	P	AC022802	146500..148500	2 FS, 4 stop
hs11M1-71	P	AC027641	51000..53000	
hs11M1-72	P	AC027641	94000..96000	2 FS, stops
hs11M1-73	P	AC036111	7000..9000	no met
hs11M1-74	C	AC021530	130000..132000	
hs11M1-75	C	AC021530	73000..75000	
hs11M1-76	PF	AC036111	42000..44000	incomplete
hs11M1-77	P	AC036111	66500..68500	3 FS
hs11M1-78	F	AC036111	77000..79000	incomplete
hs11M1-79	C	AC021530	111000..113000	
hs11M1-80	C	AC036111	129000..131000	
hs11M1-81	P	AC021530	97500..99500	
hs11M1-82	C	AC021530	137500..139500	
hs11M1-83	C	AC021530	149000..151000	
hs11M1-84	PF	AC036111	181000..end	incomplete
hs11M1-85	P	AP002517	46500..48500	
hs11M1-86	P	AP001803	36000..38000	
hs11M1-87	P/C	AP001803	25500..27500	1 FS
hs11M1-88	C	AP001803	28500..30500	
hs11M1-89	C	AC037472	88500..90500	
hs11M1-90	P	AC037472	93500..95500	2 FS, stops
hs11M1-91	C	AP002517	62000..64000	
hs11M1-92	P	AP002517	78000..80000	1 FS
hs11M1-93	P	AP002517	82500..84500	1 FS
hs11M1-94	C	AP002517	168500..170500	
hs11M1-95	C	AC037472	170500..172500	
hs11M1-96	C	AP000818	10500..12500	
hs11M1-97	P	AP000818	22000..24000	3 FS
hs11M1-98	C	AP000818	57000..59000	
hs11M1-99	P/C	AP000818	86000..88000	no met
hs11M1-100	P/C	AP000818	120000..122000	
hs11M1-101	C	AP000825	6000..8000	
hs11M1-102	C	AP000868	21500..23500	
hs11M1-103	P	AP000825	45500..47500	

hs11M1-104	P	AP000868	128000..130000	1 FS, no met
hs11M1-105	C	AC083958	127000..129000	
hs11M1-106	C	AC083958	31500..33500	
hs11M1-107	PF	AP000916	16000..18000	start only, no met
hs11M1-108	C	AC083958	154500..156500	
hs11M1-109	P	AC083958	83500..85500	
hs11M1-110	C	AC083958	15000..17000	
hs11M1-111	P	AP000916	155000..157000	1 FS
hs11M1-112	P	AP000916	31000..33000	1 stop
hs11M1-113	C	AP002512	153000..155000	
hs11M1-114	C	AP001112	13500..15500	
hs11M1-115	P	AP002517	93000..95000	1 stop
hs11M1-116	C	AP002512	60000..62000	
hs11M1-117	P	AP001112	60000..62000	1 FS
hs11M1-118	P	AP001112	69500..71500	1 FS, 2 stop
hs11M1-119	P	AP001112	81500..83500	1 FS
hs11M1-120	C	AP001112	92000..94000	
hs11M1-121	P	AP001112	104000..106000	1 stop
hs11M1-122	P	AP001803	75500..77500	
hs11M1-123	P	AP001803	105000..107000	1 FS
hs11M1-124	P	AP001803	87000..89000	
hs11M1-125	C	AP001804	17500..19500	
hs11M1-126	C	AP001804	31500..33500	
hs11M1-127	P	AP001804	37000..39000	3 stops
hs11M1-128	P	AP001804	48000..50000	5 FS, disrputed MAYDRYVAIC
hs11M1-129	P	AP001803	60500..62500	1 stop
hs11M1-130	C	AP001804	132500..134500	
hs11M1-131	C	AP001804	142000..144000	
hs11M1-132	P	AP001804	152000..154000	1 stop
hs11M1-133	P	AP001804	67000..69000	1 FS, 3 stops
hs11M1-134	P	AP001804	127000..129000	2 FS, 2 stops
hs11M1-135	C	AP001884	11000..13000	similar to -57
hs11M1-136	PF	AP001884	90000..92000	1 FS, goes into gap
hs11M1-137	C	AP001884	129000..131000	
hs11M1-138	C	AP002407	132500..134500	
hs11M1-139	P	AC022289	22500..24500	FS, incomplete
hs11M1-140	P	AC027239	178000..180000	no met
hs11M1-141	C	AC022891	1..2000	
hs11M1-142	P	AC022891	16500..18500	5 stops, 1 FS, missing motifs
hs11M1-143	C	AC022891	22500..24500	
hs11M1-144	C	AP003034	17000..19000	
hs11M1-145	C	AC022891	57500..59500	
hs11M1-146	C	AC068339	40000..42000	
hs11M1-147	P	AP003033	140000..142000	1 FS
hs11M1-148	C	AP003033	86500..88500	
hs11M1-149	P	AC068339	79000..81000	no met
hs11M1-150	P	AP003033	1500..3500	1 stop
hs11M1-151	P	AC022891	156500..158500	1 stop

hs11M1-152	C	AC022891	167000..169000	
hs11M1-153	P	AP000723	79000..81000	
hs11M1-154	C	AC026076	42000..44000	
hs11M1-155	P	AP000723	25500..27500	2 FS, no met
hs11M1-156	P	AP000723	58000..60000	1 FS, 2 stops
hs11M1-157	P	AC026076	114500..116500	1 stop
hs11M1-158	C	AP000723	87000..89000	
hs11M1-159	C	AP001998	99000..101000	
hs11M1-160	P	AP001998	116500..118500	no met
hs11M1-161	C	AP001998	148000..150000	
hs11M1-162	C	AC027239	90000..92000	
hs11M1-163	C	AC027239	102500..104500	
hs11M1-164	P	AP001998	36000..38000	1 stop, no met
hs11M1-165	P	AP001998	54500..56500	1 FS
hs11M1-166	P	AP001998	81500..83500	3 stops, 1 FS
hs11M1-167	C	AP002512	86500..88500	
hs11M1-168	P	AP002512	91000..93000	1 FS
hs11M1-169	P	AP002512	110000..113000	2 FS
hs11M1-170	P	AP002512	122000..124000	1 FS, 2 stops
hs11M1-171	C	AP002512	128000..130000	
hs11M1-172	C	AP002517	26000..28000	95% sim to hs11M1-91
hs11M1-173	P	AP003033	30000..32000	3 stops, 1 FS, missing start motifs
hs11M1-174	C	AP003033	68000..70000	Q13606 U56420 OLF1
hs11M1-175	P	AP003033	105000..107000	3 stops
hs11M1-176	C	AP003033	134500..136500	
hs11M1-177	C	AP003033	158500..160500	
hs11M1-178	P	AP003034	70500..72500	1 stop
hs11M1-179	P	AP003034	141000..143000	missing motifs
hs11M1-180	C	AP003034	153000..155000	
hs11M1-181	P	AC019093	46000..48000	1 FS, no met, 1 stop
hs11M1-182	C	AC019093	63000..65000	
hs11M1-183	C	AC019093	119000..121000	
hs11M1-184	P	AC019093	128500..130500	1 stop
hs11M1-185	C	AC019093	154000..156000	
hs11M1-186	P	AC019093	21500..23500	
hs11M1-187	C	AC069371	51000..53000	
hs11M1-188	P	AC019093	57500..59500	1 stop, 1 FS
hs11M1-189	C	AC019093	109000..111000	
hs11M1-190	P/C	AC040925	56000..58000	1 stop
hs11M1-191.1	P	AC040925	120000..122000	1 FS, goes into gap
hs11M1-192	P	AC040925	151500..153500	3 FS, 4 stops
hs11M1-193	P	AP000435	53000..55000	1 FS, no met
hs11M1-194	C	AP000435	119500..121500	
hs11M1-195	P	AP000435	79000..81000	1 FS
hs11M1-196	P	AP000435	105000..107000	7 stops
hs11M1-197	P	AP000435	110000..112000	2 stops, 2 F
hs11M1-198	C	AP000435	28500..30500	
hs11M1-199	C	AP002345	10000..12000	

hs11M1-200	C	AP002345	30000..32000	
hs11M1-201	P	AP002345	45000..47000	3 stops, 1 FS
hs11M1-202	P	AP002345	66500..68500	1 FS
hs11M1-203	C	AP002345	114000..116000	
hs11M1-204	C	AC021809	19000..21000	
hs11M1-205	P	AC021809	35000..37000	4 stops, 2 FS, missing start
hs11M1-206	P	AC021809	117500..119500	1 stop, 1 FS, missing start
hs11M1-207	P	AP002780	97500..98500	1 stop, 1 FS, no end(-208P)
hs11M1-208	P	AP002780	97500..99500	2 FS, 3 stops
hs11M1-209	C	AP002780	7000..9000	
hs11M1-210	P	AP002780	143000..145000	
hs11M1-211	P	AC004923	54500..56500	1 FS, 3 stops, glv type
hs11M1-212	P	AC079973	120000..122000	1 FS
hs11M1-213	F	AC079973	1..1000	goes into gap
hs11M1-214	P	AC079973	184000..186000	no start, 1 FS, (TTC)n repeat insert
hs11M1-215	P	AP000719	55500..57500	7 FS, 2 stops
hs11M1-216	P	AP000719	45500..47500	1 FS, 2 stops
hs11M1-217	P	AP000867	17500..19500	2 stops, 1 FS
hs11M1-218	P	AP000867	79500..81500	3 stops
hs11M1-219	C	AP003175	59000..61000	
hs11M1-220	P	AP003385	15000..17000	5 FS, no met
hs11M1-221	P	AP003385	28000..30000	1 stop, no met
hs11M1-222	P	AC009867	16500..18500	6 FS
hs11M1-223	P	AC009867	36000..38000	2 FS
hs11M1-224	P	AP002965	5000..7000	4 FS, 2 stops
hs11M1-225	P	AP002965	18500..20500	1 FS
hs11M1-226	P	AP002965	36000..38000	4 FS
hs11M1-227	C	AP002965	49500..51500	
hs11M1-228	C	AP002965	101000..103000	
hs11M1-229	P	AP002965	159000..161000	1 stop, short last motif
hs11M1-230	P	AC009545	48000..52000	Disrupted by LTR5
hs11M1-231	P	AC009545	92000..94000	1 Fs, 1 stop
hs11M1-232	C	AC009545	112500..114500	8 potential met
hs11M1-233	C	AC009545	137000..139000	
hs11M1-234	P	AC009545	152500..154500	
hs11M1-235	P	AC009545	160000..162000	
hs11M1-236	C	AC009758	136000..138000	
hs11M1-237	C	AC009758	14500..16500	
hs11M1-238	P	AC009758	28500..30500	
hs11M1-239	C	AC009758	54000..56000	
hs11M1-240	P	AC009758	72000..74000	no met
hs11M1-241	P	AC009758	85500..87500	1 FS, 1 stop
hs11M1-242	P	AC009758	125000..127000	1 FS
hs11M1-243	C	AC009642	50500..52500	
hs11M1-244	P	AC009642	5500..7500	1 FS
hs11M1-245	C	AC009642	31500..33500	
hs11M1-246	PF	AC009642	67500..69500	missing 2 end motifs
hs11M1-247	P	AC009642	132000..134000	1 FS, 1 stop

hs11M1-248	C	AC009642	74000..76000	
hs11M1-249	C	AC009642	139500..141500	
hs11M1-250	C	AC010930	9000..11000	
hs11M1-251	P	AC010930	35000..37000	1 FS
hs11M1-252	C	AC010930	49000..51000	
hs11M1-253	P	AC010930	74500..76500	1 FS, small (simple repeat) insertion
hs11M1-254	C	AC010930	105000..107000	
hs11M1-255	C	AC010930	147000..148300	
hs11M1-256	C	AC011647	30000..32000	
hs11M1-257	P	AC011647	44500..46500	2 FS, 2 stops
hs11M1-258	P	AC011647	50000..52000	2 stops
hs11M1-259	P	AC011647	59500..61500	
hs11M1-260	C	AC011647	67000..69000	
hs11M1-261	P	AC011647	84000..86000	2 FS
hs11M1-262	C	AC011647	123000..125000	
hs11M1-263	P	AC011647	143500..145500	1 FS
hs11M1-264	P	AC026083	1000..3000	2 FS
hs11M1-265	P	AC026083	12500..14500	1 FS
hs11M1-266	C	AC026083	74500..76500	
hs11M1-267	C	AC026083	87000..89000	
hs11M1-268	C	AC026083	121500..123500	
hs11M1-269	C	AC026083	138000..140000	
hs11M1-270	P	AC026083	157500..159500	1 FS
hs11M1-271	F	AC020597	3500..5500	goes into gap
hs11M1-272	P	AC020597	13500..15500	1 FS
hs11M1-273	C	AC020597	25500..27500	
hs11M1-274	P	AC020597	31500..33500	2 FS
hs11M1-275	P	AC020597	55500..57500	1 FS, 2 stops
hs11M1-276	C	AC020597	63000..65000	
hs11M1-277	C	AC020597	76000..78000	
hs11M1-278	C	AC020597	90000..92000	
hs11M1-279	P	AC020597	108000..110000	1 stop
hs11M1-280	P	AC020597	132000..134000	3 stops
hs11M1-281	C	AC020597	141000..143000	
hs11M1-282	P	AC020597	151000..153000	3 FS, 4 stops
hs11M1-283	C	AC020597	160500..162500	
hs11M1-284	P	AC020597	177000..179000	1 stop
hs11M1-285	C	AC011711	2000..4000	
hs11M1-286	C	AC011711	9000..11000	
hs11M1-287	P	AC011711	27000..29000	3 FS
hs11M1-288	C	AC011711	34500..36500	
hs11M1-289	P	AC011711	40000..42000	2 FS
hs11M1-290	C	AC011711	68000..70000	
hs11M1-291	P	AC011711	84000..86000	2 FS
hs11M1-292	C	AC011711	95000..97000	
hs11M1-293	P	AC011711	129500..131500	1 FS
hs11M1-294	C	AC011711	147000..149000	
hs11M1-295	C	AC011711	113000..115000	

hs11M1-296	C	AP002826	86500..88500	
hs11M1-297	P	AP002826	115500..117500	1 FS, 2 STOPS
hs11M1-298	C	AP002826	68500..70500	
hs11M1-299	P	AP002509	18000..20000	1 FS
hs11M1-300	P	AP002509	98000..100000	1 FS
hs11M1-301	P	AP001521	34500..36500	1 FS, 1 STOP
hs11M1-302	C	AP001521	5500..7500	
hs11M1-303	P	AP001521	54000..56000	
hs11M1-304	P	AP000629	51000..53000	4 FS
hs11M1-305	C	AC025249	16500..18500	
hs11M1-306	C	AC025249	25500..27500	
hs11M1-307	C	AC025249	38000..40000	
hs11M1-308	C	AC025249	84000..86000	
hs11M1-309	P	AC025249	139500..141500	1 STOP
hs11M1-310	P	AC025249	157000..159000	1 FS
hs11M1-311	P	AC021935	62000..64000	1 FS
hs11M1-312	P	AC021935	39000..41000	1 FS
hs11M1-313	C	AC021935	80000..82000	
hs11M1-314	C	AC021935	109500..111500	
hs11M1-315	P	AC019088	10000..12000	4 FS
hs11M1-316	C	AC019088	16500..18500	
hs11M1-317	C	AC019088	1500..3500	
hs11M1-318	C	AC019088	34000..36000	
hs11M1-319	C	AC019088	47000..49000	
hs11M1-320	C	AC019088	70000..72000	
hs11M1-321	C	AC019088	77500..79500	
hs11M1-322	C	AC019088	98000..100000	
hs11M1-323	P	AC019088	114500..116500	2 FS
hs11M1-324	P	AC019088	124000..126000	2 FS, 1 STOP, MISSING START DOMAINS
hs11M1-325	C	AC019088	148500..150500	
hs12M1-1	P	AC022207	41500..43500	1 stop
hs12M1-2	C	AC022207	55000..57000	
hs12M1-3	P	AC022207	105500..107500	1 FS,stop
hs12M1-4	P	AC022207	22000..24000	2 FS
hs12M1-5	C	AC022207	163500..165500	
hs12M1-6	P	AC068994	382500..384500	4 FS, 1 stop
hs12M1-7	P	AC009779	124000..126000	1 FS
hs12M1-8	C	AC009779	184500..186500	
hs12M1-9	P	AC008035	142000..144000	fragment, 1 FS, 1 stop
hs12M1-10	P	AC009775	73000..75000	3 FS,2 stops
hs12M1-11	P	AC009775	5500..9500	2 stops, large insertion unique DNA, SINEs)
hs12M1-12	C	AC024257	6500..8500	
hs12M1-13	P	AC083933	15500..17500	1 FS, goes into gap, 2 stops
hs12M1-14	P	AC090115	67500..69500	1 FS,1 stop
hs12M1-15	P	AC090115	76500..78500	no met, 6 stops
hs12M1-16	P	AC090115	106000..108000	3 stops, 2 FS, no final motif
hs12M1-17	P	AC090115	20000..22000	1 stop

hs12M1-18	P	AC078864	128000..130000	3 FS, 1 stop
hs12M1-19	C	AC009779	161000..163000	
hs13M1-1	P	AL354833	1500..3500	2 FS, 2 stops
hs13M1-2	PF	AL138686	64000..66000	1 FS, no start, 2 stops,LINE repeat disrupts
hs13M1-3	PF	AC024458	101000..103000	2 FS, 1 stop, no start
hs13M1-4	P	AL353580	32500..34500	2 stops, 1 FS
hs13M1-5	P	AL353580	41000..43000	2 stops, no met
hs14M1-1	C	AE000658	75000..77000	
hs14M1-2	P	AE000658	107000..109000	1 stop
hs14M1-3	C	AE000658	170500..172500	
hs14M1-4	P	AE000658	175500..177500	1 FS
hs14M1-5	C	AE000658	139000..141000	
hs14M1-6	C	AL157687	145500..147500	
hs14M1-7	C	AL160314	102000..104000	
hs14M1-8	P	AL160314	167000..169000	1 FS, 1 stop
hs14M1-9	P	AC024399	16000..18000	1 FS, no starting motifs
hs14M1-10	C	AC024399	32000..34000	
hs14M1-11	P	AC024399	64000..66000	2 FS
hs14M1-12	C	AC024399	114500..116500	
hs14M1-13	C	AC024399	122500..124500	
hs14M1-14	C	AC024399	136500..138500	
hs14M1-15	P	AC024399	152500..154500	2 FS, 1 stop
hs14M1-16	C	AC024399	67000..69000	
hs14M1-17	C	AC024399	183500..185322	
hs14M1-18	C	AL359218	26500..28500	
hs14M1-19	P	AL163152	10500..12500	3 FS,6 stops
hs14M1-20	C	AL359218	53000..55000	
hs14M1-21	C	AL359218	72500..74500	
hs14M1-22	P	AL359218	84500..86500	1 FS
hs14M1-23	C	AL359218	111000..113000	
hs14M1-24	P	AL359218	130000..132000	1 FS,1 stop
hs14M1-25	C	AL359218	151000..153000	
hs14M1-26	PF	AL163152	60000..62000	no start, 2 FS, 7 stops
hs14M1-27	C	AL163152	84000..86000	
hs14M1-28	C	AL163152	110500..112500	
hs14M1-29	P	AL163152	146000..148000	3 FS
hs14M1-30	P	AL356019	69000..71000	2 stops, 1 FS
hs14M1-31	C	AL356019	83500..85500	
hs14M1-32	P	AL356019	89000..91000	1 stop
hs14M1-33	C	AL356019	102500..104500	
hs14M1-34	C	AL356019	57000..59000	
hs14M1-35	C	AL163636	72500..74500	
hs14M1-36	P	AL132827	58000..60000	3 FS, 5 stops
hs14M1-37	P	AL079307	91500..93500	3 stops
hs14M1-38	P	AL079307	97500..99500	2 stops, 2 FS
hs14M1-39	PF	AL079307	106000..108000	1 FS, only end 2 motifs

hs15M1-1	P	AC010760	25000..27000	1 FS, no Met
hs15M1-2	P	AC010760	37500..39500	1 FS, 2 stops
hs15M1-3	C	AC010760	61500..63500	
hs15M1-4	C	AC010760	75000..77000	
hs15M1-5	P	AC010760	106000..108000	1 stop
hs15M1-6	P	AC010760	150500..152500	1 stop, 1 FS
hs15M1-7	P	AC010760	172000..174000	1 FS
hs15M1-8	P	AC020679	33000..35000	2 FS
hs15M1-9	P	AC020679	41500..43500	no met, 1 FS
hs15M1-10	C	AC025234	70000..72000	
hs15M1-11	C	AC005143	2500..4500	
hs16M1-1	C	AJ003147	163476..164414	
hs16M1-2	P	AJ003147	174780..175778	No Met
hs16M1-3	C	AC068380	147000..149000	
hs17M1-1	P	AC007194	142000..143000	OR17-25
hs17M1-2	C	AC007194	142000..143000	OR17-24
hs17M1-3	PF	AC023106	1000..2000	1 FS, 1 stop, goes into gap
hs17M1-4	C	AC007194	56000..57000	OR17-2
hs17M1-5	C	AC007194	162000..163000	OR17-40
hs17M1-6	C	AC007194	406000..407000	OR17-31
hs17M1-7	C	AC007194	376000..377000	OR17-31
hs17M1-8	P	AC007194	352000..353000	OR17-210
hs17M1-9	C	AC007194	342000..343000	OR17-209
hs17M1-10	P	AC007194	314000..315000	OR17-208
hs17M1-11	C	AC007194	255000..256000	OR17-6
hs17M1-12	C	AC007194	238000..239000	OR17-7
hs17M1-13	C	AC007194	212000..214000	OR17-30
hs17M1-14	P	AC007194	187000..189000	2 FS, OR17-23
hs17M1-15	C	AC007194	175000..176000	OR17-228
hs17M1-16	P	AC007194	67000..69000	2 FS, OR17-1
hs17M1-17	C	AC007194	32000..35000	OR17-201
hs17M1-18	C	AC007194	20000..22000	OR17-93
hs17M1-19	C	AC005962	76500..78500	
hs17M1-20	C	AC005962	91000..93000	
hs18M1-1	P	AC025953	66500..68500	1 stop
hs18M1-2	P	AC025953	79500..81500	2 FS
hs19M1-1	C	AC005255	40752..41771	
hs19M1-2	C	AC011517	16000..18000	
hs19M1-3	C	AC010322	1..2000	
hs19M1-4	C	AC002988	93896..94343	
hs19M1-5	P	AC002988	38600..37677	1 FS
hs19M1-6	P	AC002988	9581..10510	no met
hs19M1-7	C	AC005255	12644..13606	

hs19M1-8	P	AC005255	65502..66414	1 FS
hs19M1-9	C	AC005255	54418..55347	
hs19M1-10	P	AC004659	23355..24339	2 stops,1FS
hs19M1-11	F	AC004659	1..476	goes into gap
hs19M1-12	C	AC004659	37792..38751	
hs19M1-13	C	L78442	4856..5773	
hs19M1-14	P	AC003956	29110..30022	no met, 2 FS
hs19M1-15	C	AC004510	1852..2808	
hs19M1-16	C	AC004597	2990..3937	
hs19M1-17	C	AC004597	16239..17289	
hs19M1-18	C	AC004794	6034..7101	
hs19M1-20	P	AC006271	14410..15367	no met, 1 FS, 1 stop
hs19M1-21	C	AC006271	24021..24957	
hs19M1-22	P	AC006271	45263..46252	
hs19M1-23	C	AC006271	61163..62179	
hs19M1-24	P	AC006271	88552..89488	1 FS, 1 stop
hs19M1-25	P	L78442	10000..12000	no met, 1 FS, 1 stop
hs19M1-26	P	L78442	21000..23000	7 FS, 4 stops
hs19M1-27	C	AC011464	39000..41000	
hs19M1-28	C	AC011464	98000..100000	
hs19M1-29	F	AC011464	103000..105000	missing 2 end motifs
hs19M1-30	C	AC011464	110000..112000	
hs19M1-31	C	AC011464	122000..124000	
hs19M1-32	C	AC011537	7516..8463	
hs19M1-33	C	AC011464	131000..133000	
hs19M1-34	PF	AC011464	142000..144000	1 FS, missing end motifs
hs19M1-35	PF	AC006271	26302..26526	(1 stop)
hs19M1-36	P	AC006271	76494..75515	no met, 1 FS
hs19M1-37	PF	AC006271	71835..71098	missing 2 start motifs, 2 stops
hs21M1-1	P	AP000181	48500..50500	4 FS, 1 stop
hs21M1-2	P	AP001465	1500..3500	2 FS
hs21M1-3	P	AP001465	38500..40500	2 FS
hs22M1-1	C	AP000534	32582..33529	
hsXM1-1	P	AL049734	500..2500	
hsXM1-2	C	AL049734	120000..121000	
hsXM1-3	P	AL109853	26000..28000	
hsXM1-4	P	AL109853	4000..6000	fragment, repeat insert(nnn)
hsXM1-5	P	AL135784	66000..68000	3 FS
hsXM1-6	P	AL135784	10000..12000	
hsXM1-7	P	AL355366	12000..14000	1 L1PA7 insert, 1 FS, 1 stop
hsXM1-8	P	AF277315	139000..141000	1 FS, no met, stops