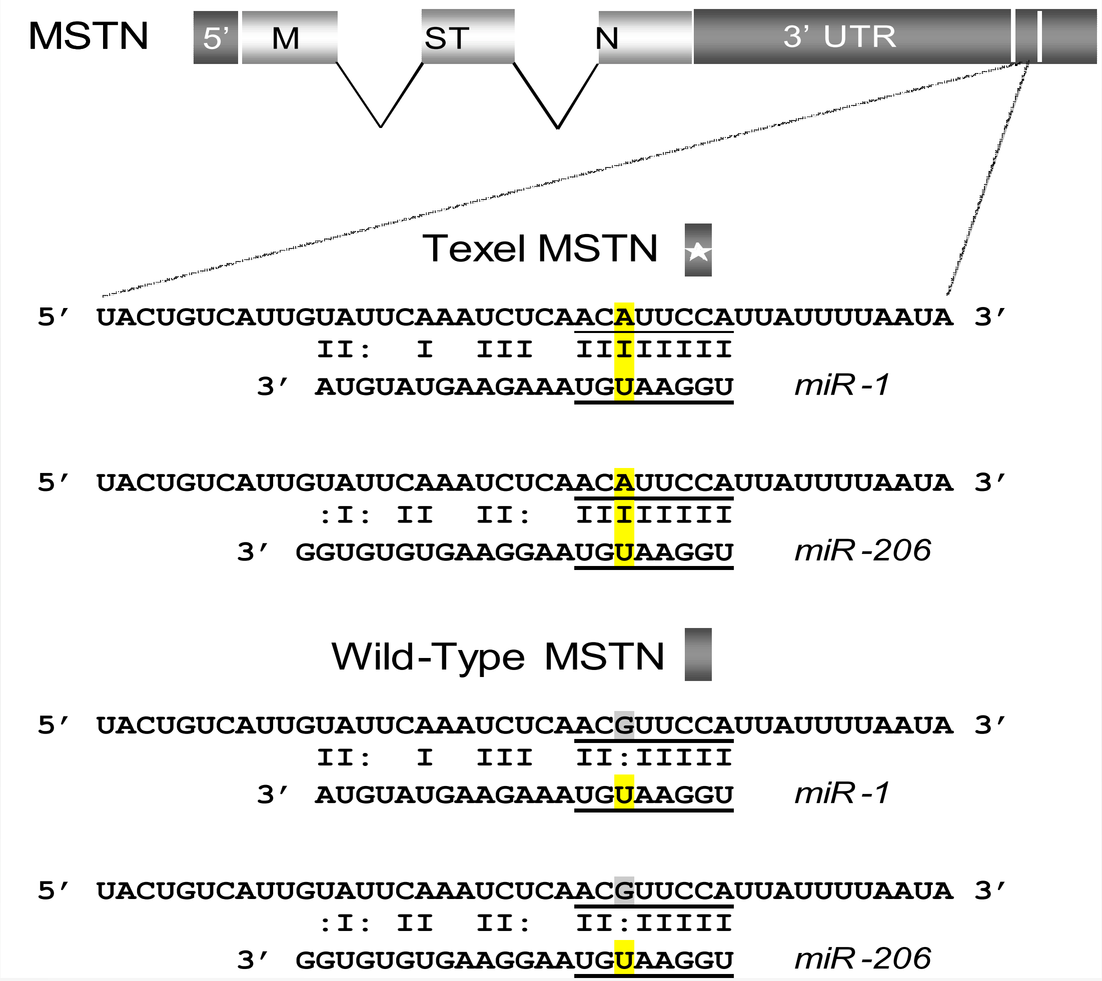
Patrocles: a database of polymorphic miRNA-mediated gene regulation

Satellite Eadgene Course "A primer in miRNA biology"
Liège, 3 march 2008

S. Hiard, D. Baurain, W. Coppieters, X. Tordoir, C. Charlier,
and M. Georges; University of Liège

Background



Texel MSTN

```
5′ UACUGUCAUUGUAUUCAAAUCUCAACAUUCCAUUAUUUUAAUA 3′
              II:  I  III  IIIIIIII
           3′ AUGUAUGAAGAAAUGUAAGGU      miR-1

5′ UACUGUCAUUGUAUUCAAAUCUCAACAUUCCAUUAUUUUAAUA 3′
              :I: II  II:  IIIIIIII
          3′ GGUGUGUGAAGGAAUGUAAGGU      miR-206
```

Wild-Type MSTN

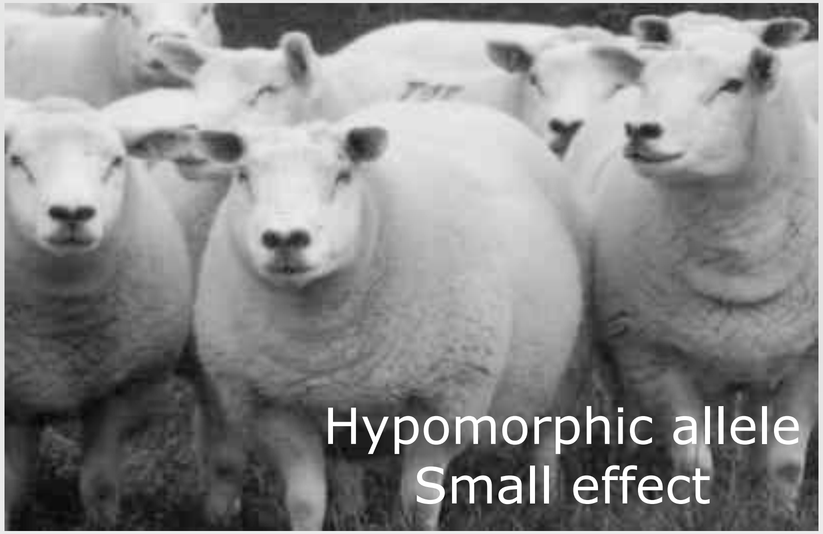
```
5′ UACUGUCAUUGUAUUCAAAUCUCAACGUUCCAUUAUUUUAAUA 3′
              II:  I  III  II:IIIII
           3′ AUGUAUGAAGAAAUGUAAGGU      miR-1

5′ UACUGUCAUUGUAUUCAAAUCUCAACGUUCCAUUAUUUUAAUA 3′
              :I: II  II:  II:IIIII
          3′ GGUGUGUGAAGGAAUGUAAGGU      miR-206
```

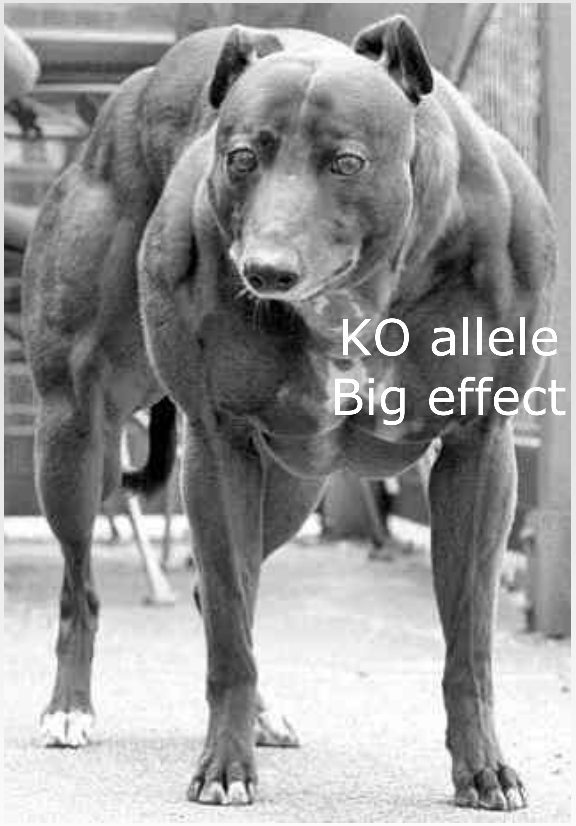
Texel sheep myostatin allele

Clop, A., Marcq, F., Takeda, H., Pirottin, D., Tordoir, X., Bibé, B., Bouix, J., Caiment, F., Elsen, J.M., Eychenne, F., Larzul, C., Laville, E., Meish, F. , Milenkovic, D. , Tobin, J., Charlier, C., Georges, M. (2006)
A mutation creating a potential illegitimate miRNA target site in the myostatin gene affects muscularity in sheep
Nature Genetics, **38**: 813-818

Motivation

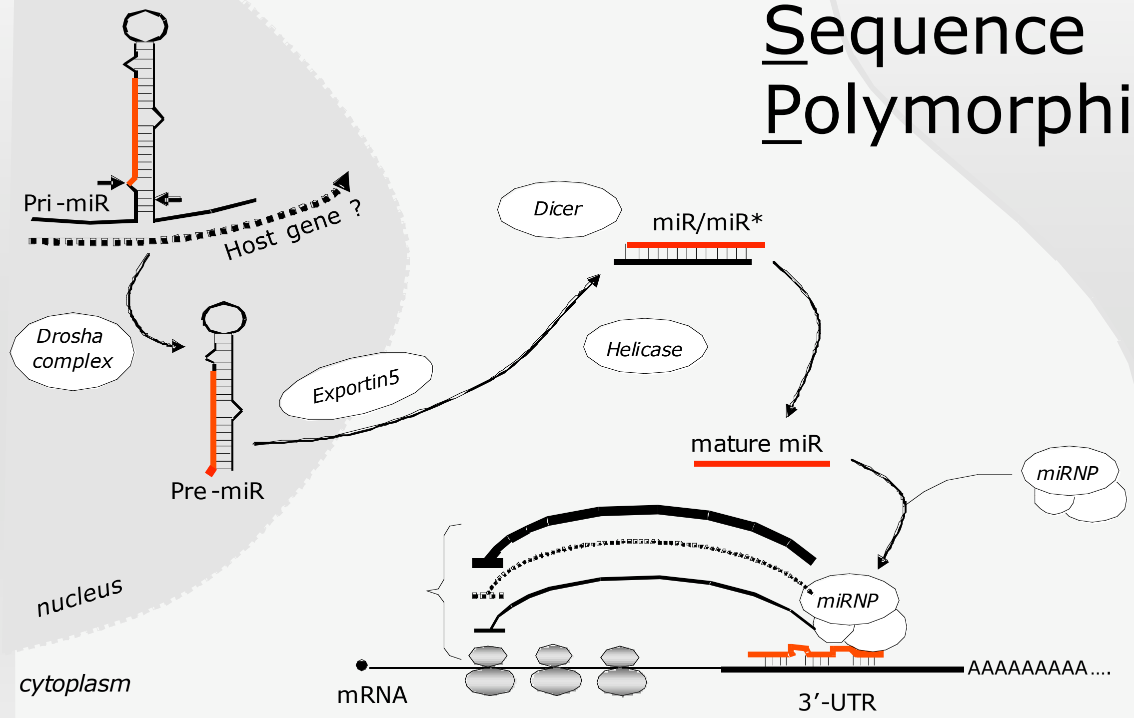
- ☐ Translational inhibition of the Texel *MSTN* allele


Hypomorphic allele
Small effect


KO allele
Big effect

- ☐ Are Texel-like mutations common?
  - Patrocles database

Scope of Patrocles database



DNA
Sequence
Polymorphisms: DSPs

- 1 – Targets (1)

  V

- 2 – miRNAs (100s)

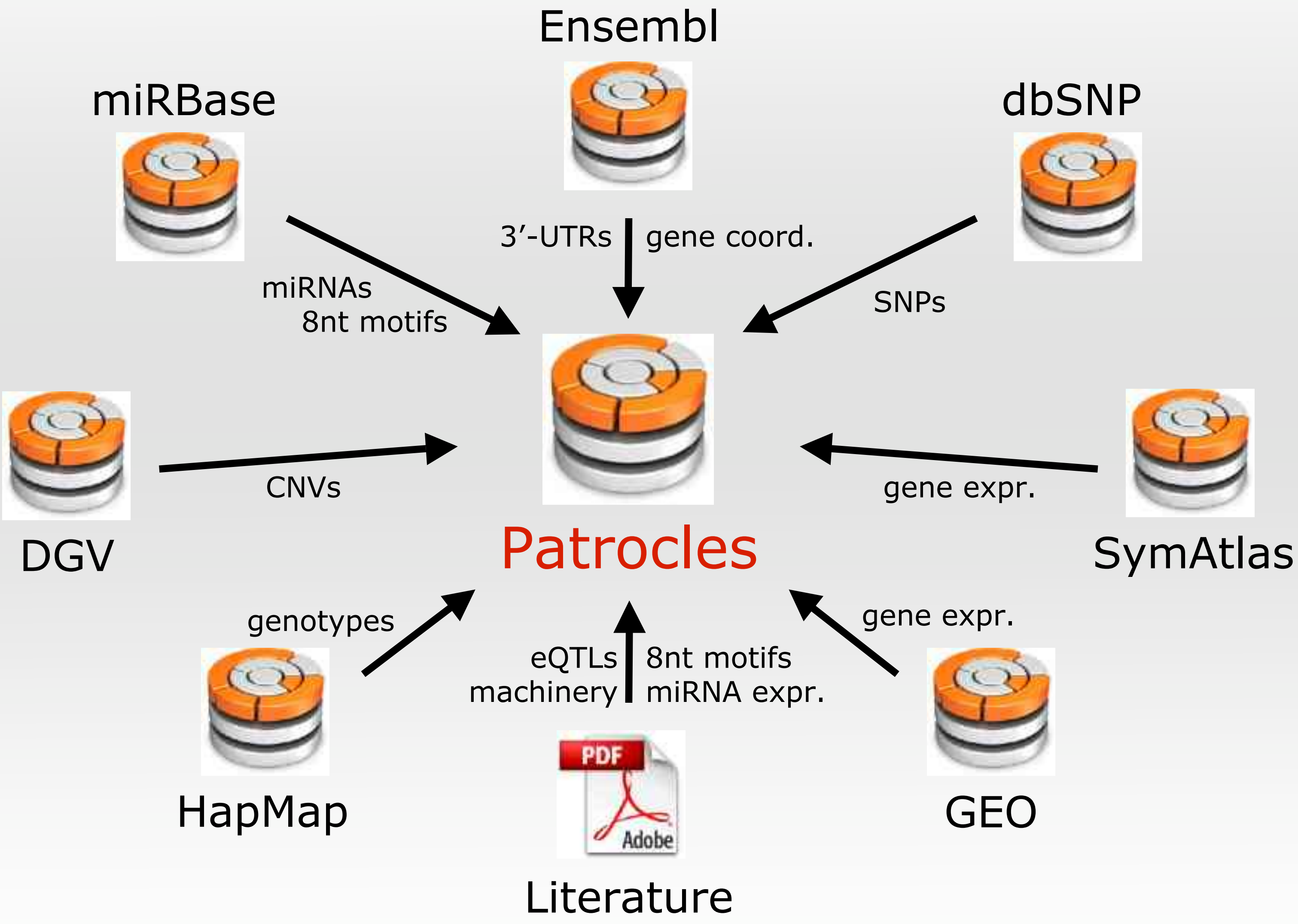
  V

- 3 – Silencing machinery (overall)

- Subtle variation in expression (hypo/hyper)
- Possibly no phenotypic expression
- Risk factor in complex diseases?

Georges M., Coppieters W., Charlier C. (2007)
Polymorphic miRNA-mediated gene regulation: contribution to phenotypic variation and disease
Current Opinion in Genetics & Development **17**:166-176

Integrating data from multiple sources



Ensembl

miRBase

dbSNP

3’-UTRs

gene coord.

miRNAs

8nt motifs

SNPs

CNVs

gene expr.

DGV

Patrocles

SymAtlas

genotypes

gene expr.

eQTLs

machinery

8nt motifs

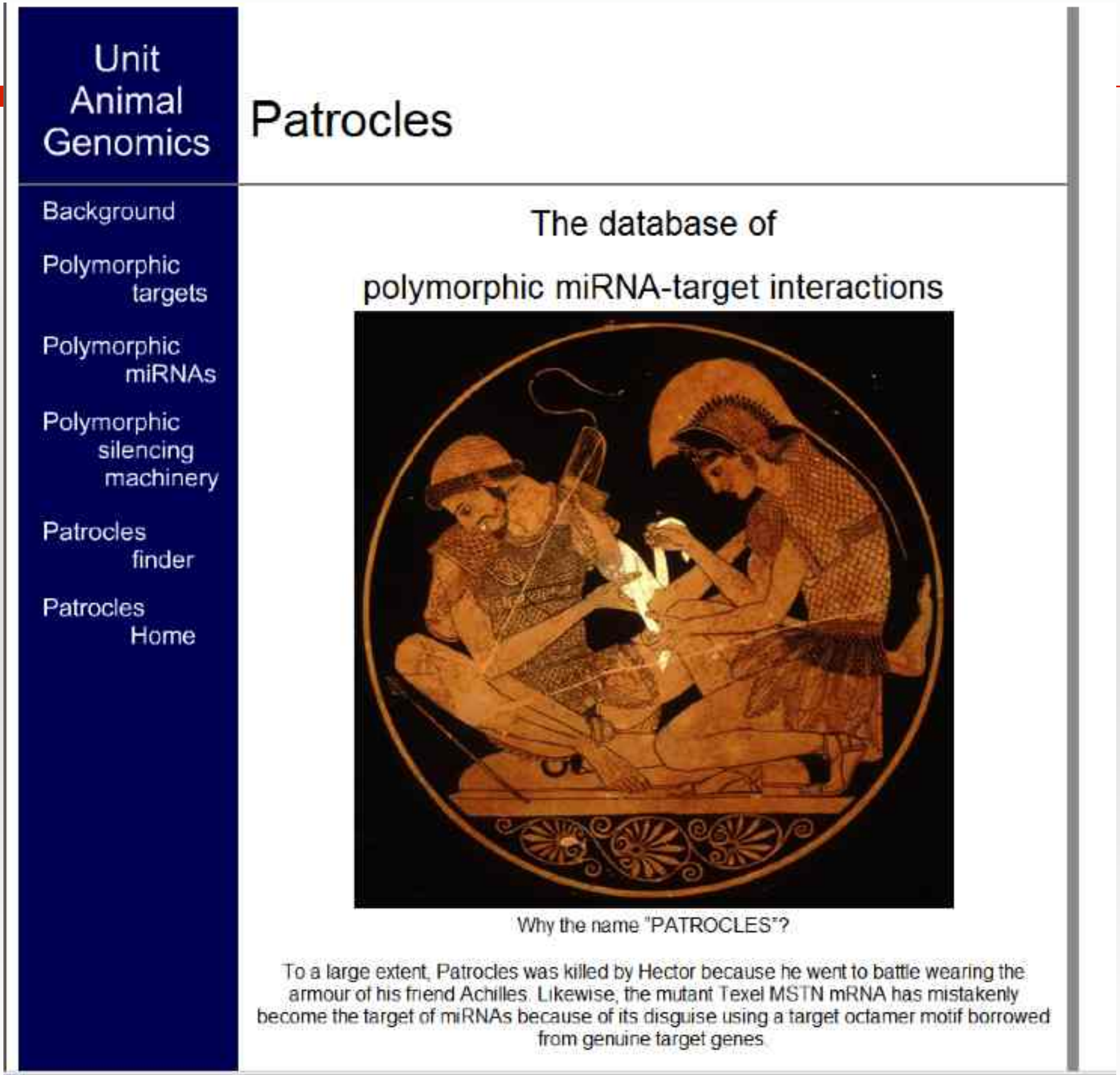
miRNA expr.

HapMap

GEO

Literature

Synchronisation…



Unit Animal Genomics

Patrocles

- Background
- 1. Polymorphic targets
- 2. Polymorphic miRNAs
- 3. Polymorphic silencing machinery
- 4. Patrocles finder
- Patrocles Home

The database of

polymorphic miRNA-target interactions

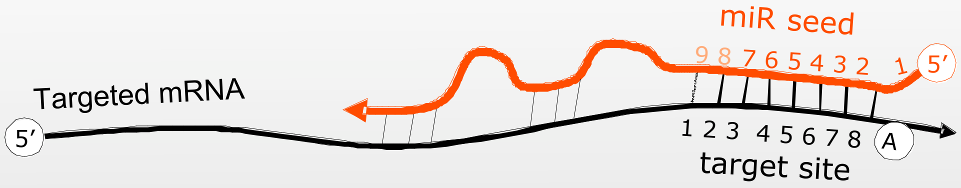
Why the name "PATROCLES"?

To a large extent, Patrocles was killed by Hector because he went to battle wearing the armour of his friend Achilles. Likewise, the mutant Texel MSTN mRNA has mistakenly become the target of miRNAs because of its disguise using a target octamer motif borrowed from genuine target genes.

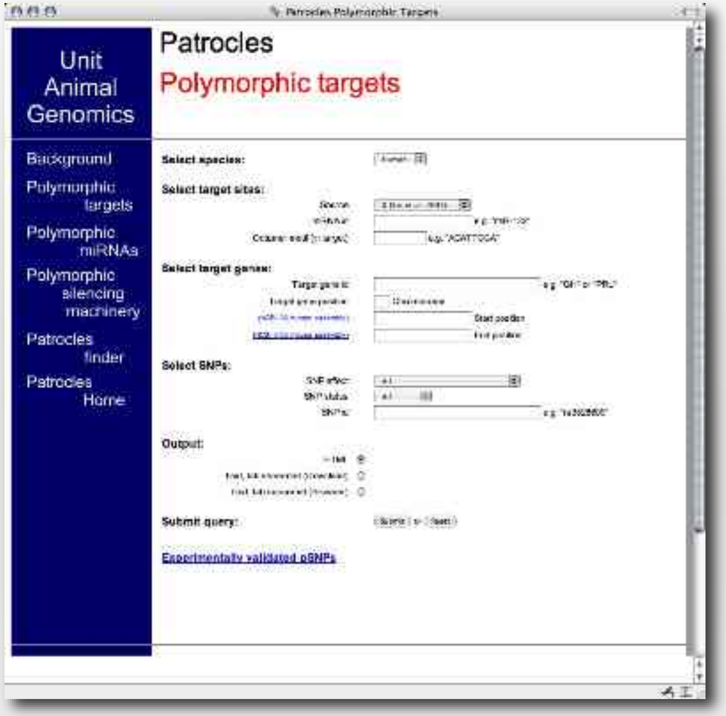
☐ Currently,
  - **human**
  - **mouse**
  - bovine
  - dog
  - chicken

http://www.patrocles.org/

Part 1 – Polymorphic targets



- ☐ 8nt motifs
  - Xie et al. (2005): 540 8nt motif (mammals) conserved, putative miR target sites
  - Lewis et al. (2005): 569 8nt motifs (human) mature miRNA in miRBase 2-8 + A



Unit Animal Genomics

Background
Polymorphic targets
Polymorphic miRNAs
Polymorphic silencing machinery
Patrocles finder
Patrocles Home

Patrocles
Polymorphic targets

Select species:

Select target sites:

Select target genes:

Select SNPs:

Output:

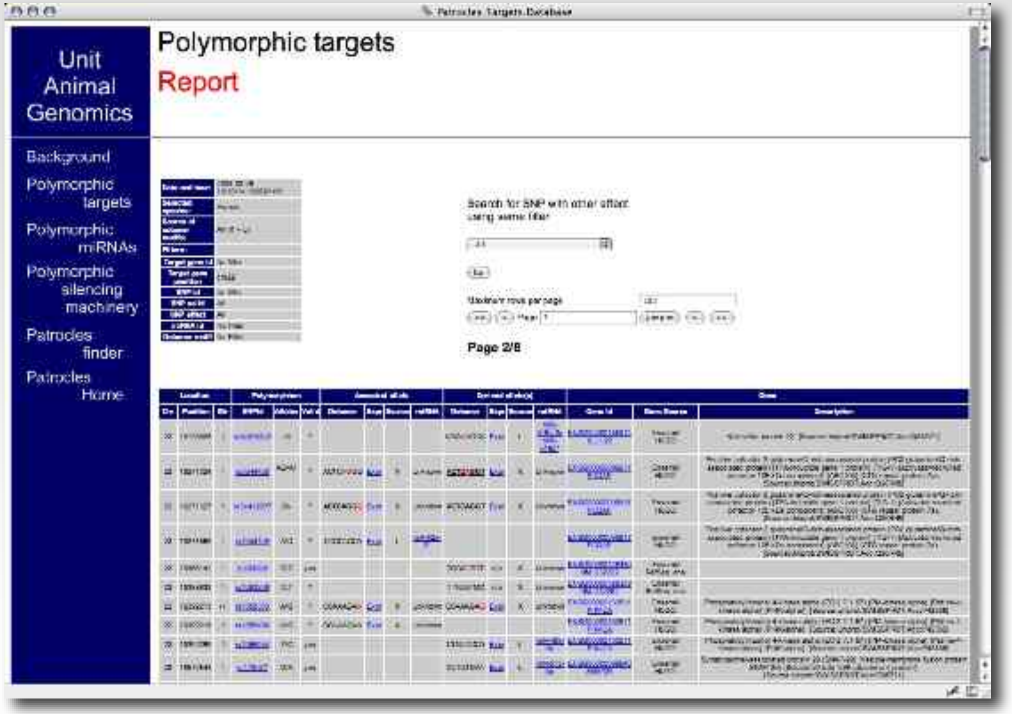
Submit query:

Experimentally validated pSNPs

input form

- ☐ various filters
  - SNP
  - target gene
  - miRNA

output report



Unit Animal Genomics

Background
Polymorphic targets
Polymorphic miRNAs
Polymorphic silencing machinery
Patrocles finder
Patrocles Home

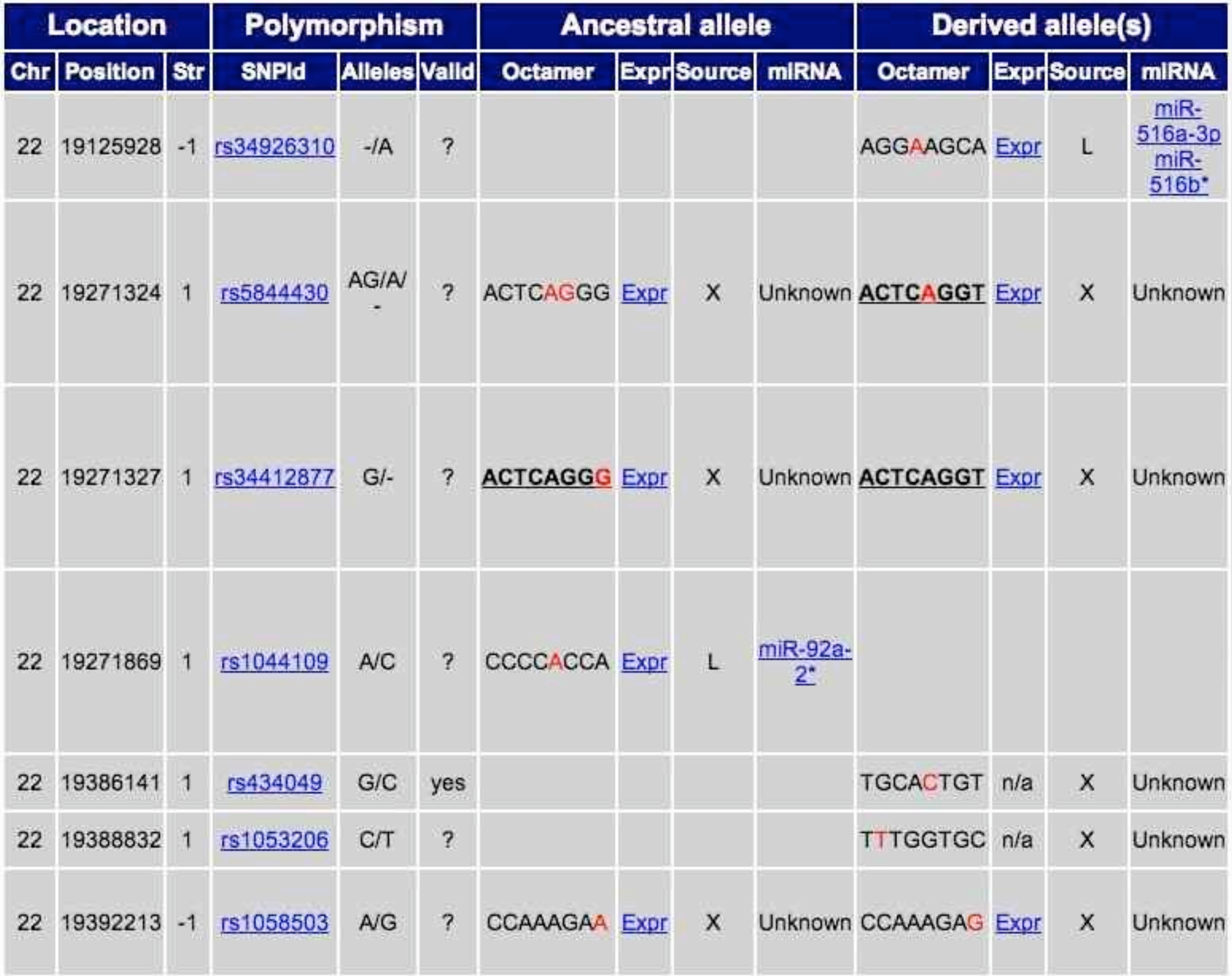
Polymorphic targets
Report

Search for SNP with other effect using same filter

Maximum rows per page

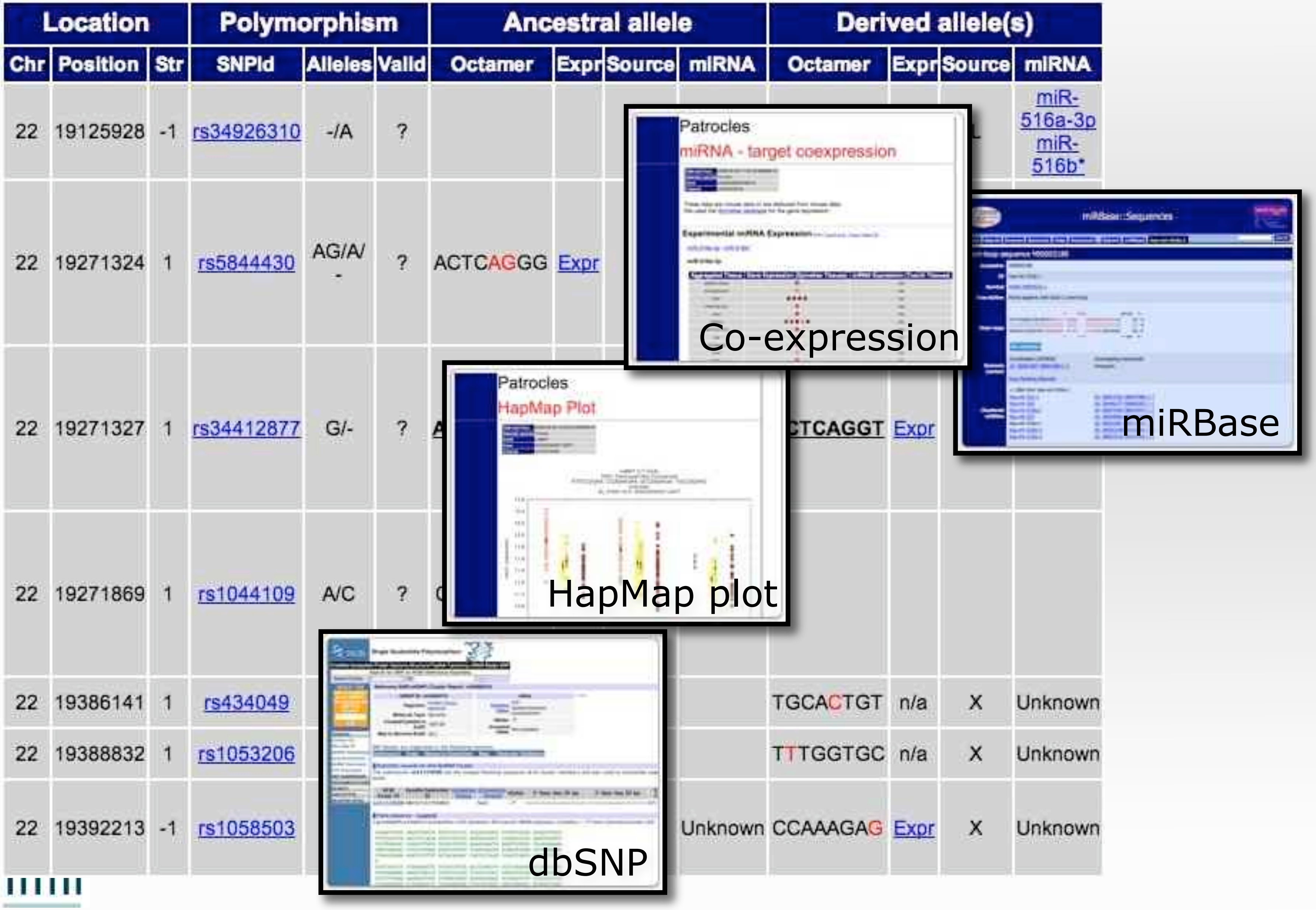
Page 2/8

Targets (zoomed output)



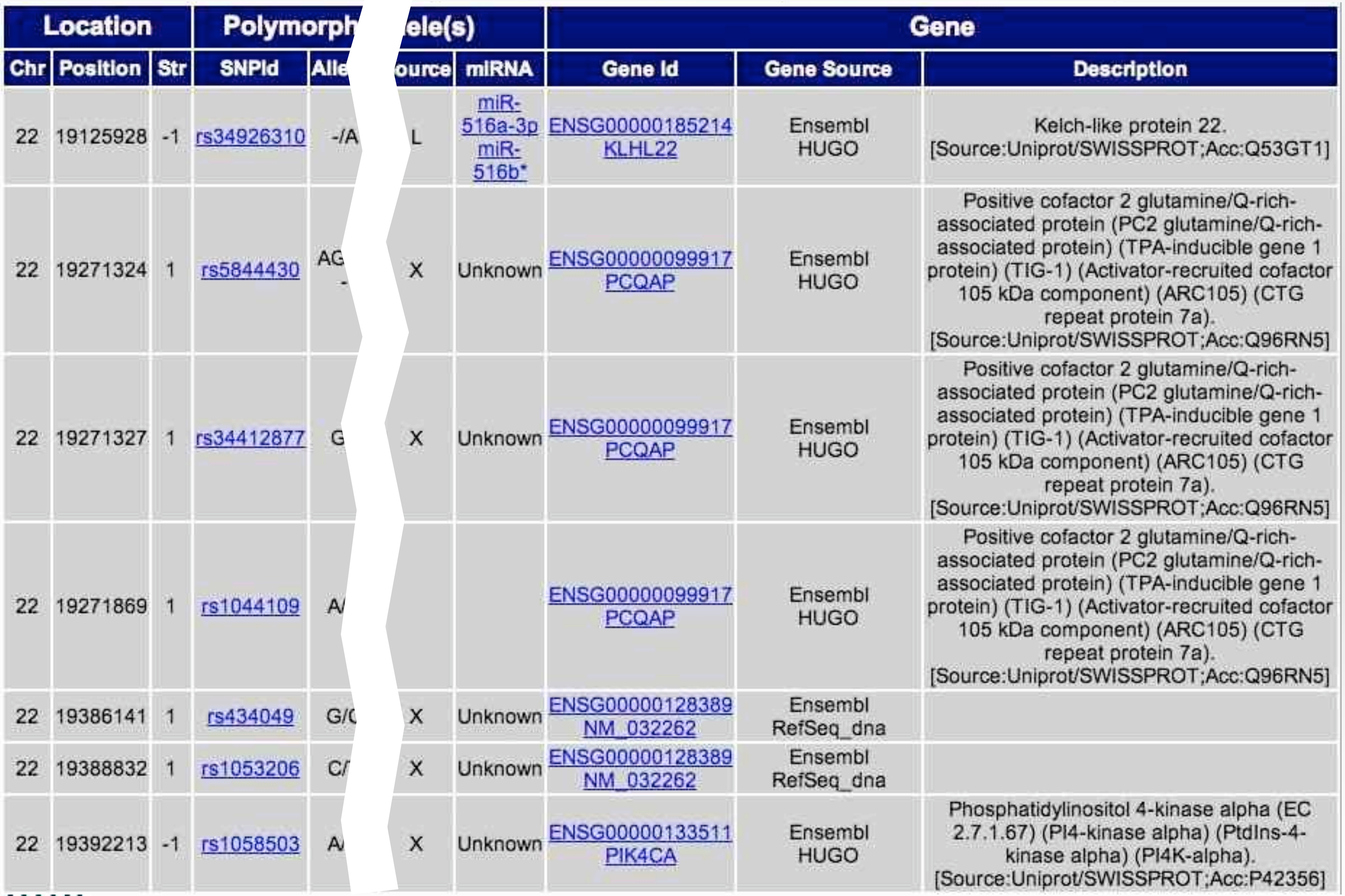
Location			Polymorphism			Ancestral allele				Derived allele(s)			
Chr	Position	Str	SNPId	Alleles	Valid	Octamer	Expr	Source	miRNA	Octamer	Expr	Source	miRNA
22	19125928	-1	rs34926310	-/A	?					AGGAAGCA	Expr	L	miR-516a-3p miR-516b*
22	19271324	1	rs5844430	AG/A/-	?	ACTCAGGG	Expr	X	Unknown	**ACTCAGGT**	Expr	X	Unknown
22	19271327	1	rs34412877	G/-	?	**ACTCAGGG**	Expr	X	Unknown	**ACTCAGGT**	Expr	X	Unknown
22	19271869	1	rs1044109	A/C	?	CCCCACCA	Expr	L	miR-92a-2*				
22	19386141	1	rs434049	G/C	yes					TGCACTGT	n/a	X	Unknown
22	19388832	1	rs1053206	C/T	?					TTTGGTGC	n/a	X	Unknown
22	19392213	-1	rs1058503	A/G	?	CCAAAGAA	Expr	X	Unknown	CCAAAGAG	Expr	X	Unknown

Targets (zoomed output)



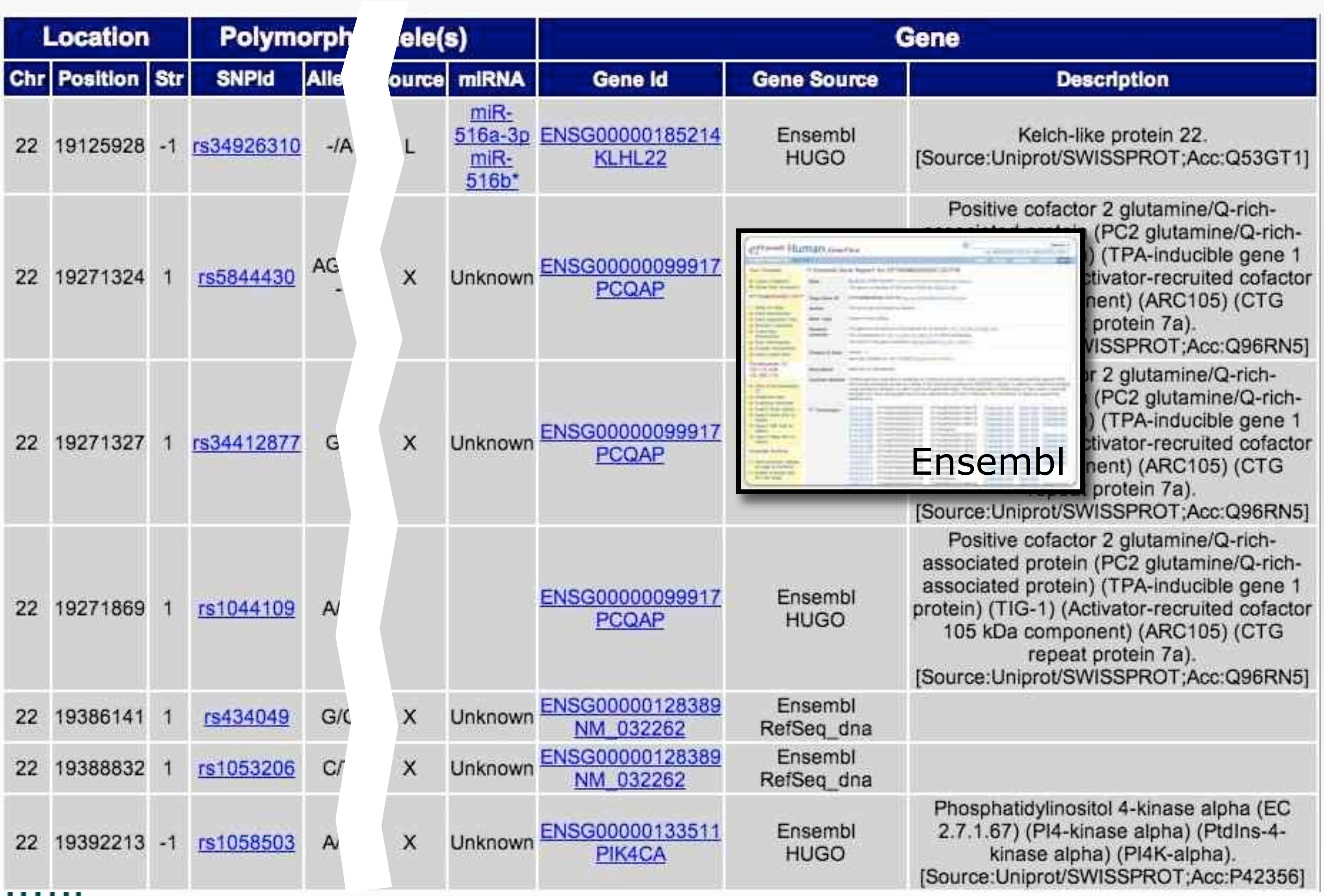
Location			Polymorphism			Ancestral allele				Derived allele(s)			
Chr	Position	Str	SNPId	Alleles	Valid	Octamer	Expr	Source	miRNA	Octamer	Expr	Source	miRNA
22	19125928	-1	rs34926310	-/A	?								miR-516a-3p miR-516b*
22	19271324	1	rs5844430	AG/A/-	?	ACTCAGGG	Expr						
22	19271327	1	rs34412877	G/-	?	A				CTCAGGT	Expr		
22	19271869	1	rs1044109	A/C	?								
22	19386141	1	rs434049							TGCACTGT	n/a	X	Unknown
22	19388832	1	rs1053206							TTTGGTGC	n/a	X	Unknown
22	19392213	-1	rs1058503						Unknown	CCAAAGAG	Expr	X	Unknown

Targets (zoomed output)



Location			Polymorph	ele(s)			Gene		
Chr	Position	Str	SNPId	Alle	ource	miRNA	Gene Id	Gene Source	Description
22	19125928	-1	rs34926310	-/A	L	miR-516a-3p miR-516b*	ENSG00000185214 KLHL22	Ensembl HUGO	Kelch-like protein 22. [Source:Uniprot/SWISSPROT;Acc:Q53GT1]
22	19271324	1	rs5844430	AG -	X	Unknown	ENSG00000099917 PCQAP	Ensembl HUGO	Positive cofactor 2 glutamine/Q-rich-associated protein (PC2 glutamine/Q-rich-associated protein) (TPA-inducible gene 1 protein) (TIG-1) (Activator-recruited cofactor 105 kDa component) (ARC105) (CTG repeat protein 7a). [Source:Uniprot/SWISSPROT;Acc:Q96RN5]
22	19271327	1	rs34412877	G	X	Unknown	ENSG00000099917 PCQAP	Ensembl HUGO	Positive cofactor 2 glutamine/Q-rich-associated protein (PC2 glutamine/Q-rich-associated protein) (TPA-inducible gene 1 protein) (TIG-1) (Activator-recruited cofactor 105 kDa component) (ARC105) (CTG repeat protein 7a). [Source:Uniprot/SWISSPROT;Acc:Q96RN5]
22	19271869	1	rs1044109	A/			ENSG00000099917 PCQAP	Ensembl HUGO	Positive cofactor 2 glutamine/Q-rich-associated protein (PC2 glutamine/Q-rich-associated protein) (TPA-inducible gene 1 protein) (TIG-1) (Activator-recruited cofactor 105 kDa component) (ARC105) (CTG repeat protein 7a). [Source:Uniprot/SWISSPROT;Acc:Q96RN5]
22	19386141	1	rs434049	G/C	X	Unknown	ENSG00000128389 NM_032262	Ensembl RefSeq_dna	
22	19388832	1	rs1053206	C/	X	Unknown	ENSG00000128389 NM_032262	Ensembl RefSeq_dna	
22	19392213	-1	rs1058503	A/	X	Unknown	ENSG00000133511 PIK4CA	Ensembl HUGO	Phosphatidylinositol 4-kinase alpha (EC 2.7.1.67) (PI4-kinase alpha) (PtdIns-4-kinase alpha) (PI4K-alpha). [Source:Uniprot/SWISSPROT;Acc:P42356]

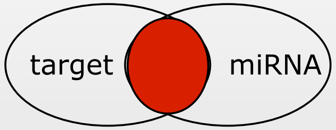
Targets (zoomed output)



Location			Polymorph		ele(s)		Gene		
Chr	Position	Str	SNPId	Alle	ource	miRNA	Gene Id	Gene Source	Description
22	19125928	-1	rs34926310	-/A	L	miR-516a-3p miR-516b*	ENSG00000185214 KLHL22	Ensembl HUGO	Kelch-like protein 22. [Source:Uniprot/SWISSPROT;Acc:Q53GT1]
22	19271324	1	rs5844430	AG -	X	Unknown	ENSG00000099917 PCQAP		Positive cofactor 2 glutamine/Q-rich- (PC2 glutamine/Q-rich-) (TPA-inducible gene 1 ctivator-recruited cofactor nent) (ARC105) (CTG protein 7a). ISSPROT;Acc:Q96RN5]
22	19271327	1	rs34412877	G	X	Unknown	ENSG00000099917 PCQAP		r 2 glutamine/Q-rich- (PC2 glutamine/Q-rich-) (TPA-inducible gene 1 ctivator-recruited cofactor nent) (ARC105) (CTG protein 7a). [Source:Uniprot/SWISSPROT;Acc:Q96RN5]
22	19271869	1	rs1044109	A/			ENSG00000099917 PCQAP	Ensembl HUGO	Positive cofactor 2 glutamine/Q-rich-associated protein (PC2 glutamine/Q-rich-associated protein) (TPA-inducible gene 1 protein) (TIG-1) (Activator-recruited cofactor 105 kDa component) (ARC105) (CTG repeat protein 7a). [Source:Uniprot/SWISSPROT;Acc:Q96RN5]
22	19386141	1	rs434049	G/(	X	Unknown	ENSG00000128389 NM_032262	Ensembl RefSeq_dna	
22	19388832	1	rs1053206	C/	X	Unknown	ENSG00000128389 NM_032262	Ensembl RefSeq_dna	
22	19392213	-1	rs1058503	A/	X	Unknown	ENSG00000133511 PIK4CA	Ensembl HUGO	Phosphatidylinositol 4-kinase alpha (EC 2.7.1.67) (PI4-kinase alpha) (PtdIns-4-kinase alpha) (PI4K-alpha). [Source:Uniprot/SWISSPROT;Acc:P42356]

Co-expression: cloning data

Aggregated Tissue	Gene Expression (Symatlas Tissues)	miRNA Expression (Tuschl Tissues)
adipose tissue	●	● ●
B-lymphocyte	●	○ ○ ○ ○ ○ ○ ○ ○ ◎ ○ ○ ○
brain	◎ ◎ ◎ ◎	◉ ● ◉
collecting duct	◉	◎
colon	●	○
embryo	◉ ◉ ◉ ◉ ◉	◎ ●
epithelial cell	◉	● ◎
eye	●	◎
glomerulus	◉	●
heart	●	●
kidney	◉	◉ ●
liver	◉	●
lung	◉	●
ovary	●	●
pancreatic alpha cell	◎	●
pancreatic beta cell	◎	●
placenta	◉	●
skin	◉	●
small intestine	●	○
spinal cord	● ●	●
spleen	◎	◎
testis	◎	● ●
thymus	●	◍
T-lymphocyte CD8	◎	○ ◍ ○



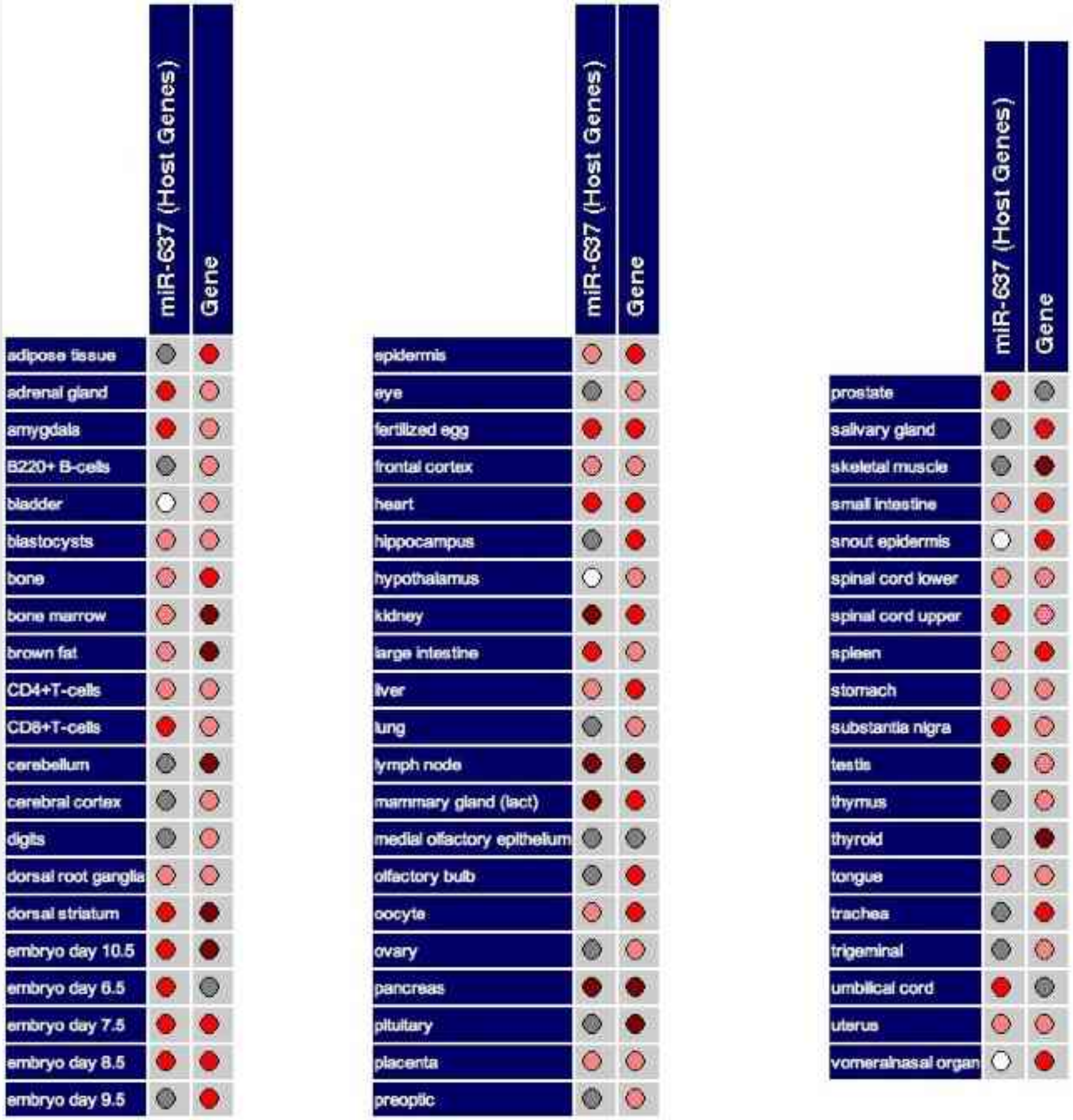
score?
tissue mapping?

- target gene
  - SymAtlas
- miRNA
  - cloning
  - host gene
  - computational

○ [0.0,0.2[
◍ [0.2,0.4[
◎ [0.4,0.6[
● [0.6,0.8[
◉ [0.8,1.0]

Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, Pfeffer S, Rice A, Kamphorst AO, Landthaler M, Lin C, Socci ND, Hermida L, Fulci V, Chiaretti S, Foà R, Schliwka J, Fuchs U, Novosel A, Müller RU, Schermer B, Bissels U, Inman J, Phan Q, Chien M, Weir DB, Choksi R, De Vita G, Frezzetti D, Trompeter HI, Hornung V, Teng G, Hartmann G, Palkovits M, Di Lauro R, Wernet P, Macino G, Rogler CE, Nagle JW, Ju J, Papavasiliou FN, Benzing T, Lichter P, Tam W, Brownstein MJ, Bosio A, Borkhardt A, Russo JJ, Sander C, Zavolan M, Tuschl T (2007) A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**: 1401-1414

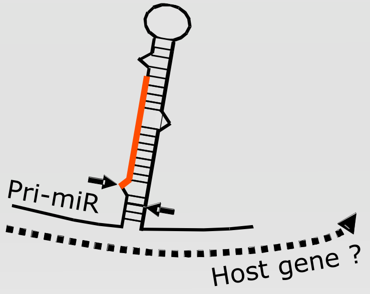
Co-expression: ‘proxy’ data



	miR-637 (Host Genes)	Gene
adipose tissue	●	●
adrenal gland	●	●
amygdala	●	●
B220+ B-cells	●	●
bladder	○	●
blastocysts	●	●
bone	●	●
bone marrow	●	●
brown fat	●	●
CD4+T-cells	●	●
CD8+T-cells	●	●
cerebellum	●	●
cerebral cortex	●	●
digits	●	●
dorsal root ganglia	●	●
dorsal striatum	●	●
embryo day 10.5	●	●
embryo day 6.5	●	●
embryo day 7.5	●	●
embryo day 8.5	●	●
embryo day 9.5	●	●

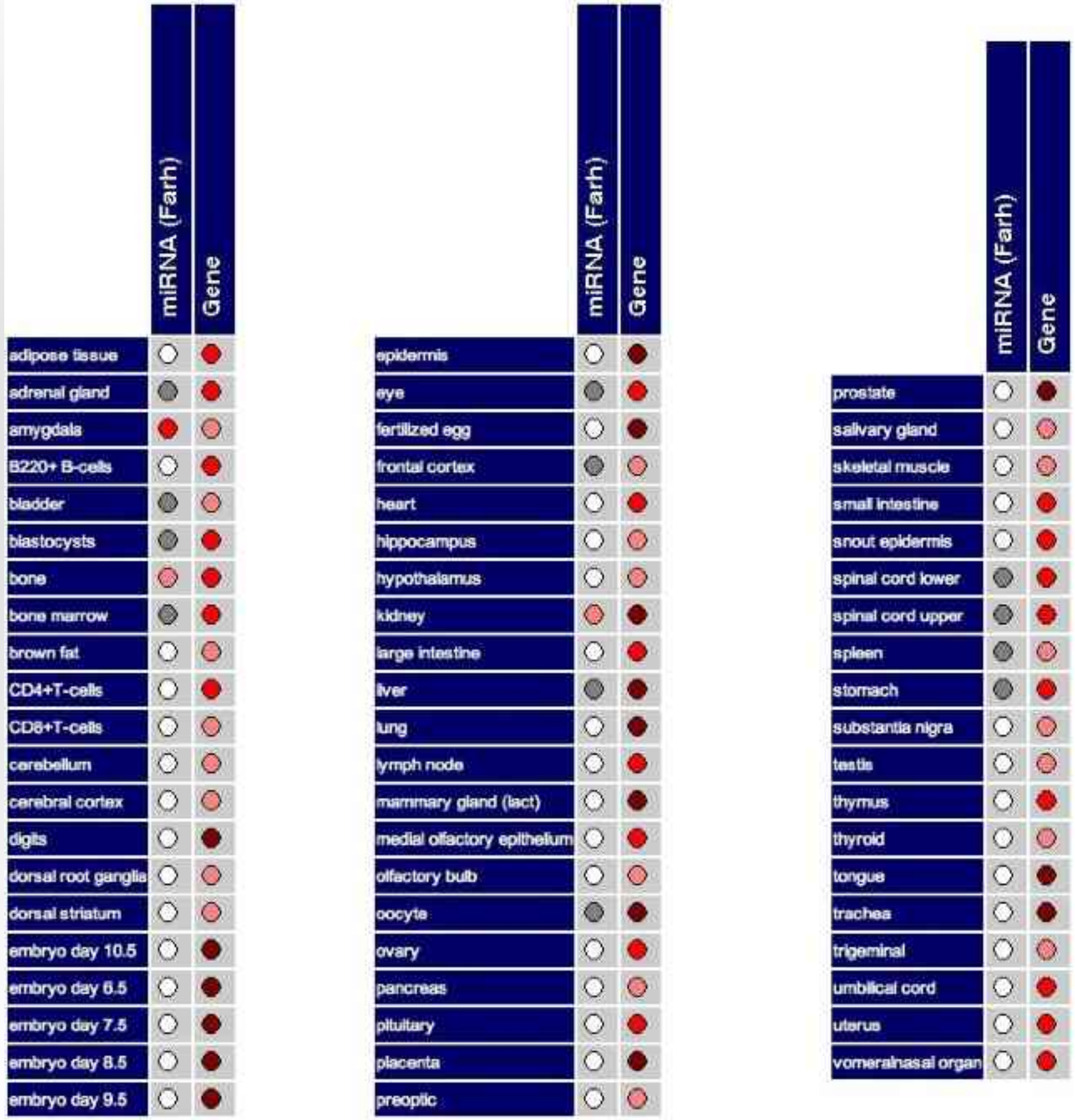
	miR-637 (Host Genes)	Gene
epidermis	●	●
eye	●	●
fertilized egg	●	●
frontal cortex	●	●
heart	●	●
hippocampus	●	●
hypothalamus	○	●
kidney	●	●
large intestine	●	●
liver	●	●
lung	●	●
lymph node	●	●
mammary gland (lact)	●	●
medial olfactory epithelium	●	●
olfactory bulb	●	●
oocyte	●	●
ovary	●	●
pancreas	●	●
pituitary	●	●
placenta	●	●
preoptic	●	●

	miR-637 (Host Genes)	Gene
prostate	●	●
salivary gland	●	●
skeletal muscle	●	●
small intestine	●	●
snout epidermis	○	●
spinal cord lower	●	●
spinal cord upper	●	●
spleen	●	●
stomach	●	●
substantia nigra	●	●
testis	●	●
thymus	●	●
thyroid	●	●
tongue	●	●
trachea	●	●
trigeminal	●	●
umbilical cord	●	●
uterus	●	●
vomeralnasal organ	○	●



SymAtlas

Co-expression: computational data



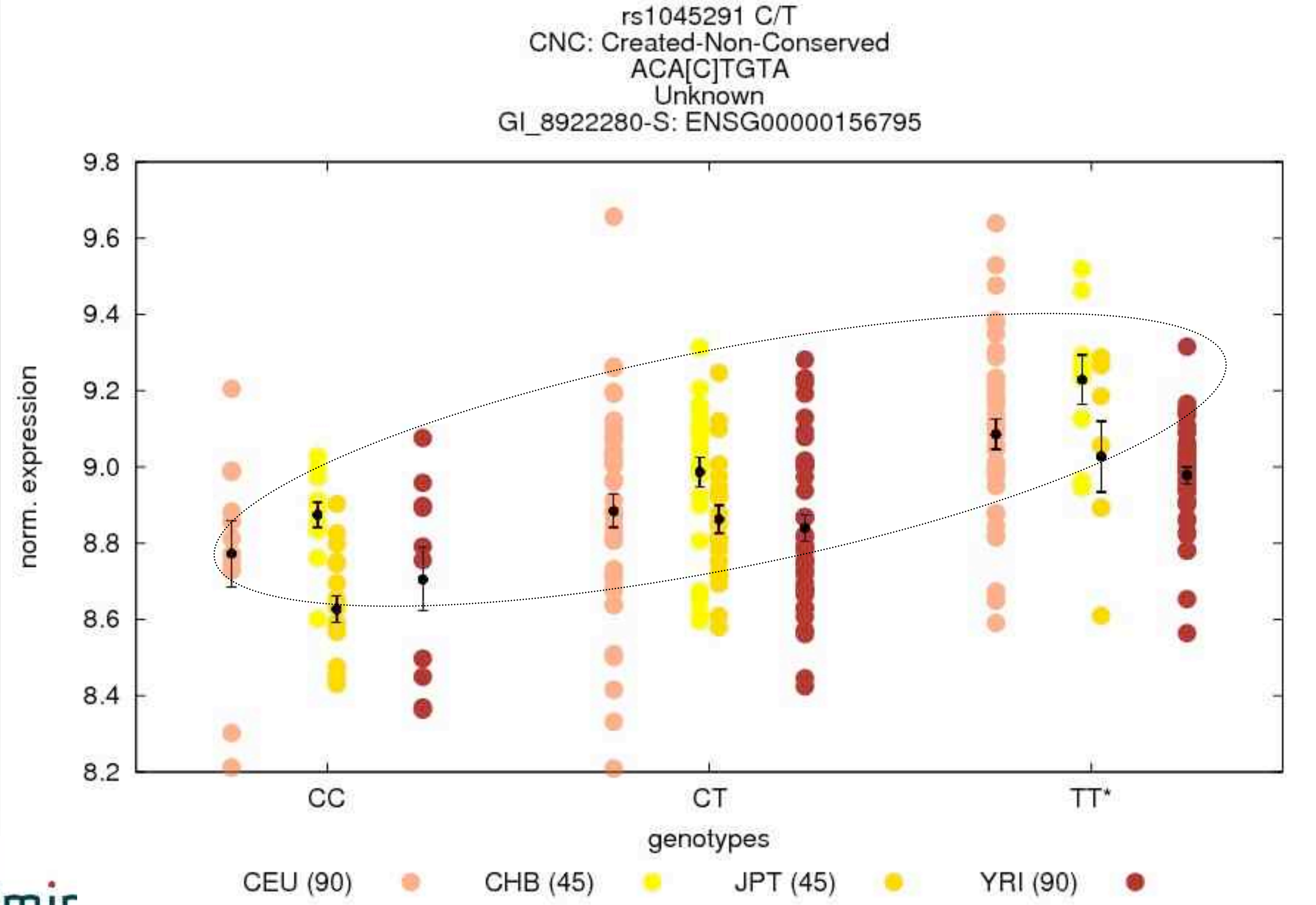
	miRNA (Farh)	Gene
adipose tissue	○	🔴
adrenal gland	●	🔴
amygdala	🔴	◎
B220+ B-cells	○	🔴
bladder	●	◎
blastocysts	●	🔴
bone	◎	🔴
bone marrow	●	🔴
brown fat	○	◎
CD4+T-cells	○	🔴
CD8+T-cells	○	◎
cerebellum	○	◎
cerebral cortex	○	◎
digits	○	🟤
dorsal root ganglia	○	◎
dorsal striatum	○	◎
embryo day 10.5	○	🟤
embryo day 6.5	○	🟤
embryo day 7.5	○	🟤
embryo day 8.5	○	🟤
embryo day 9.5	○	🟤

	miRNA (Farh)	Gene
epidermis	○	🟤
eye	●	🔴
fertilized egg	○	🟤
frontal cortex	●	◎
heart	○	🔴
hippocampus	○	◎
hypothalamus	○	◎
kidney	◎	🟤
large intestine	○	🔴
liver	●	🟤
lung	○	🟤
lymph node	○	🔴
mammary gland (lact)	○	🟤
medial olfactory epithelium	○	🔴
olfactory bulb	○	◎
oocyte	●	🟤
ovary	○	🔴
pancreas	○	◎
pituitary	○	🔴
placenta	○	🟤
preoptic	○	◎

	miRNA (Farh)	Gene
prostate	○	🟤
salivary gland	○	◎
skeletal muscle	○	◎
small intestine	○	🔴
snout epidermis	○	🔴
spinal cord lower	●	🔴
spinal cord upper	●	🔴
spleen	●	◎
stomach	●	🔴
substantia nigra	○	◎
testis	○	◎
thymus	○	🔴
thyroid	○	◎
tongue	○	🟤
trachea	○	🟤
trigeminal	○	◎
umbilical cord	○	🔴
uterus	○	🔴
vomeronasal organ	○	🔴

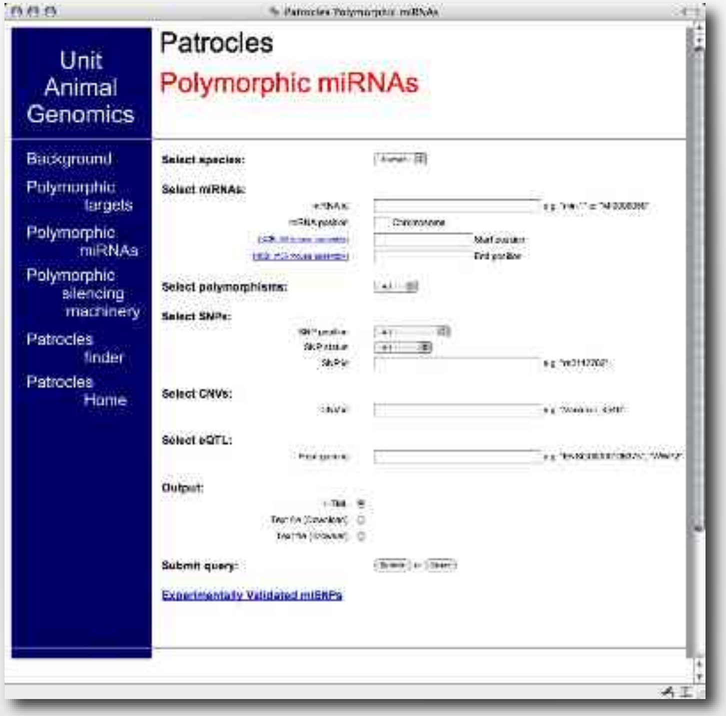
Selective avoidance as in Farh KK, Grimson A, Jan C, Lewis BP, Johnston WK, Lim LP, Burge CB, Bartel DP (2005) The widespread impact of mammalian MicroRNAs on mRNA repression and evolution. *Science* **310**: 1817-1821

HapMap plots



Part 2 – Polymorphic miRNAs

Université de Liège



Patrocles

Polymorphic miRNAs

Select species:

Select miRNAs:

Chromosome

Start position

End position

Select polymorphisms:

Select SNPs:

Select CNVs:

Select eQTL:

Output:

Submit query:

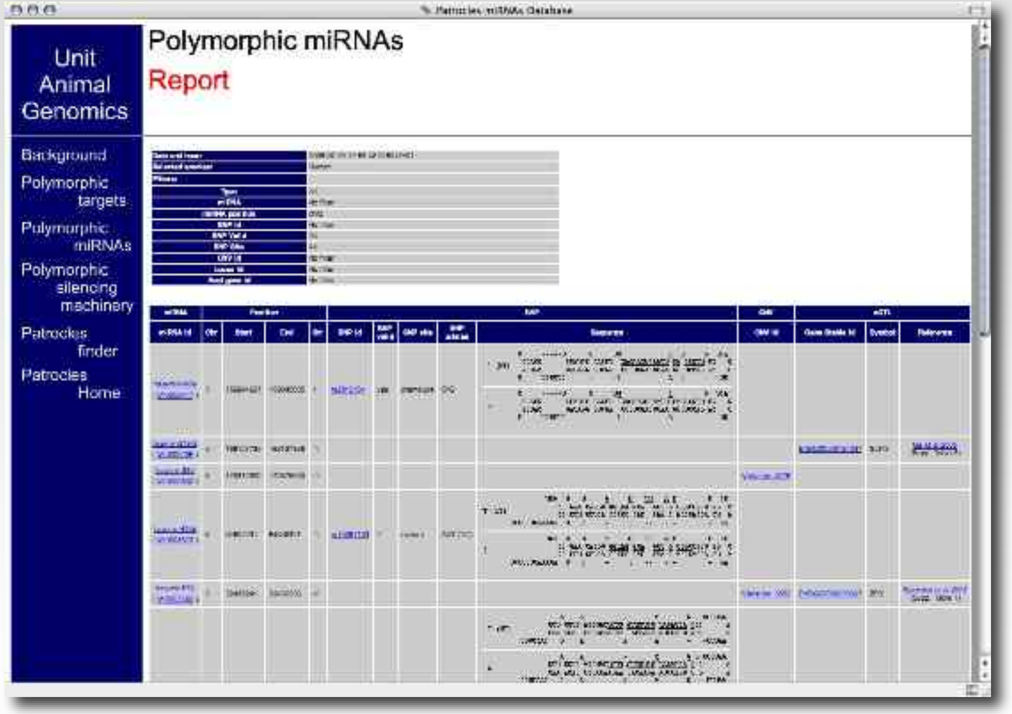
Experimentally Validated miSNPs

input form

- ☐ various filters
  - SNP
  - host gene
  - miRNA

- ☐ DSPs altering miRNA sequence
  - (de-)stabilizing interaction
- ☐ DSPs altering miRNA concentration
  - SNPs altering processing efficiency
  - CNVs encompassing miRNA genes
  - eQTLs corresponding to host genes

output report

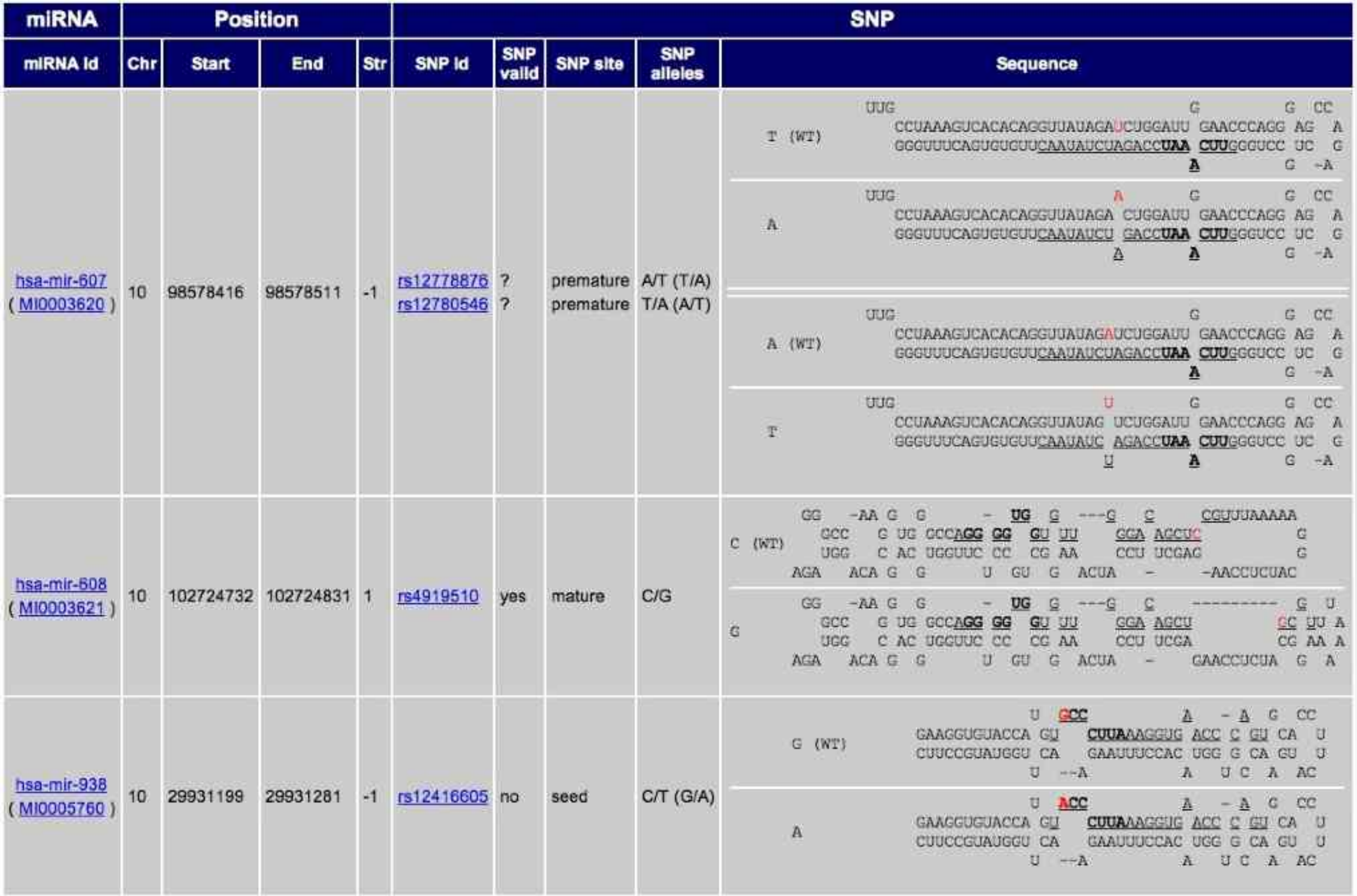


Polymorphic miRNAs

Report

Callimir

miRNAs (zoomed output)



miRNA	Position				SNP			
miRNA Id	Chr	Start	End	Str	SNP Id	SNP valid	SNP site	SNP alleles
hsa-mir-607 (MI0003620)	10	98578416	98578511	-1	rs12778876 rs12780546	? ?	premature premature	A/T (T/A) T/A (A/T)
hsa-mir-608 (MI0003621)	10	102724732	102724831	1	rs4919510	yes	mature	C/G
hsa-mir-938 (MI0005760)	10	29931199	29931281	-1	rs12416605	no	seed	C/T (G/A)

Sequence

hsa-mir-607

```
          UUG                               G         G  CC
T (WT)       CCUAAAGUCACACAGGUUAUAGAUCUGGAUU GAACCCAGG AG  A
             GGGUUUCAGUGUGUUCAAUAUCUAGACCUAA CUUGGGUCC UC  G
                                            A         G  -A
```

```
          UUG                       A       G         G  CC
A            CCUAAAGUCACACAGGUUAUAGA CUGGAUU GAACCCAGG AG  A
             GGGUUUCAGUGUGUUCAAUAUCU GACCUAA CUUGGGUCC UC  G
                                    A       A         G  -A
```

```
          UUG                               G         G  CC
A (WT)       CCUAAAGUCACACAGGUUAUAGAUCUGGAUU GAACCCAGG AG  A
             GGGUUUCAGUGUGUUCAAUAUCUAGACCUAA CUUGGGUCC UC  G
                                            A         G  -A
```

```
          UUG                      U        G         G  CC
T            CCUAAAGUCACACAGGUUAUAG UCUGGAUU GAACCCAGG AG  A
             GGGUUUCAGUGUGUUCAAUAUC AGACCUAA CUUGGGUCC UC  G
                                   U        A         G  -A
```

hsa-mir-608

```
        GG   -AA G  G      -  UG  G  ---G   C     CGUUUAAAAA
C  (WT)   GCC   G UG GCCAGG GG  GU UU    GGA AGCUC          G
          UGG   C AC UGGUUC CC  CG AA    CCU UCGAG          G
       AGA   ACA G  G      U  GU  G  ACUA   -     -AACCUCUAC
```

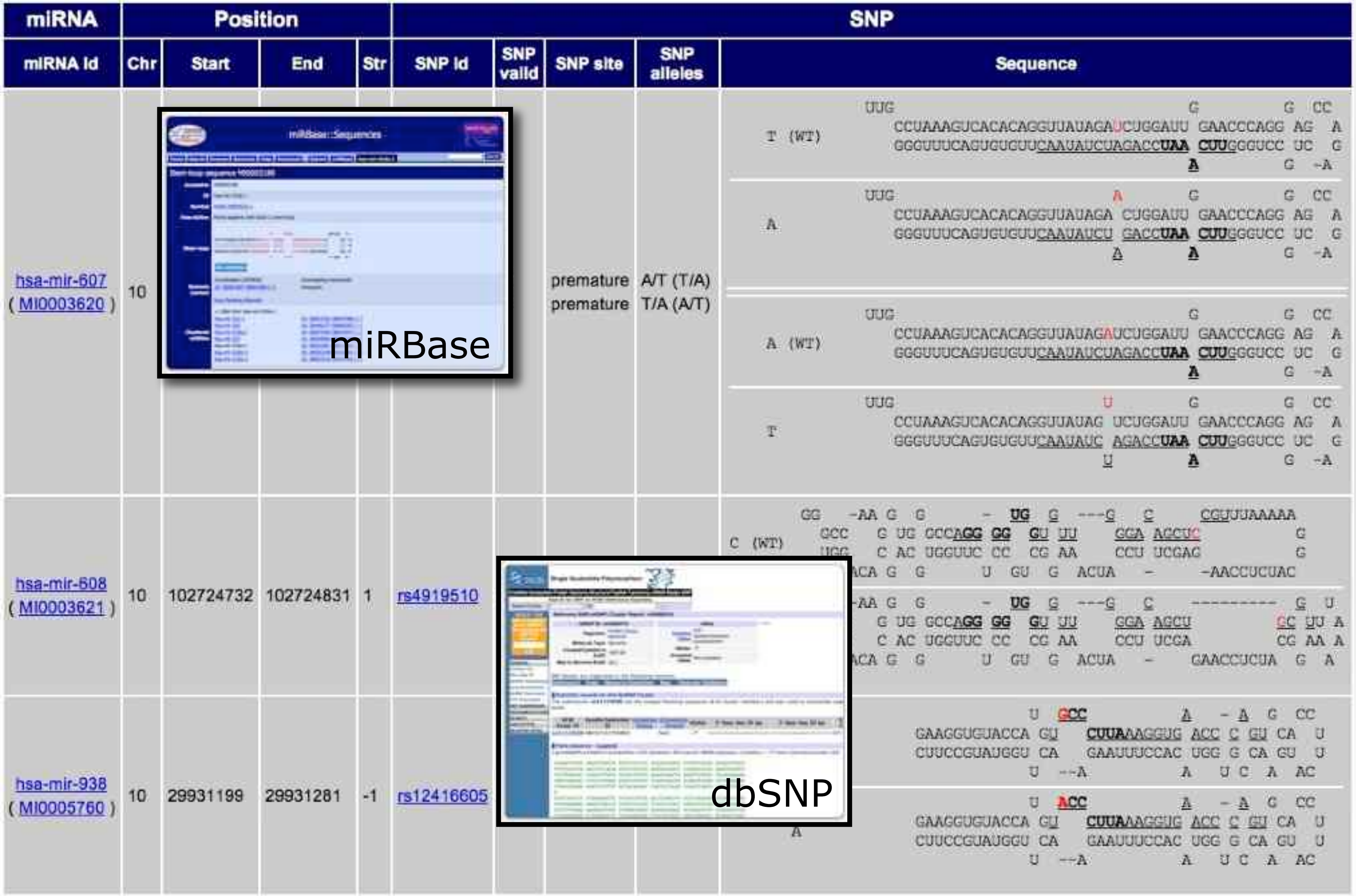
```
        GG   -AA G  G      -  UG  G  ---G   C    ---------  G  U
G         GCC   G UG GCCAGG GG  GU UU    GGA AGCU         GC UU A
          UGG   C AC UGGUUC CC  CG AA    CCU UCGA         CG AA A
       AGA   ACA G  G      U  GU  G  ACUA   -    GAACCUCUA  G  A
```

hsa-mir-938

```
                         U  GCC          A   - A  G  CC
G  (WT)      GAAGGUGUACCA GU   CUUAAAGGUG ACC C GU CA  U
             CUUCCGUAUGGU CA   GAAUUUCCAC UGG G CA GU  U
                         U  --A          A   U C  A  AC
```

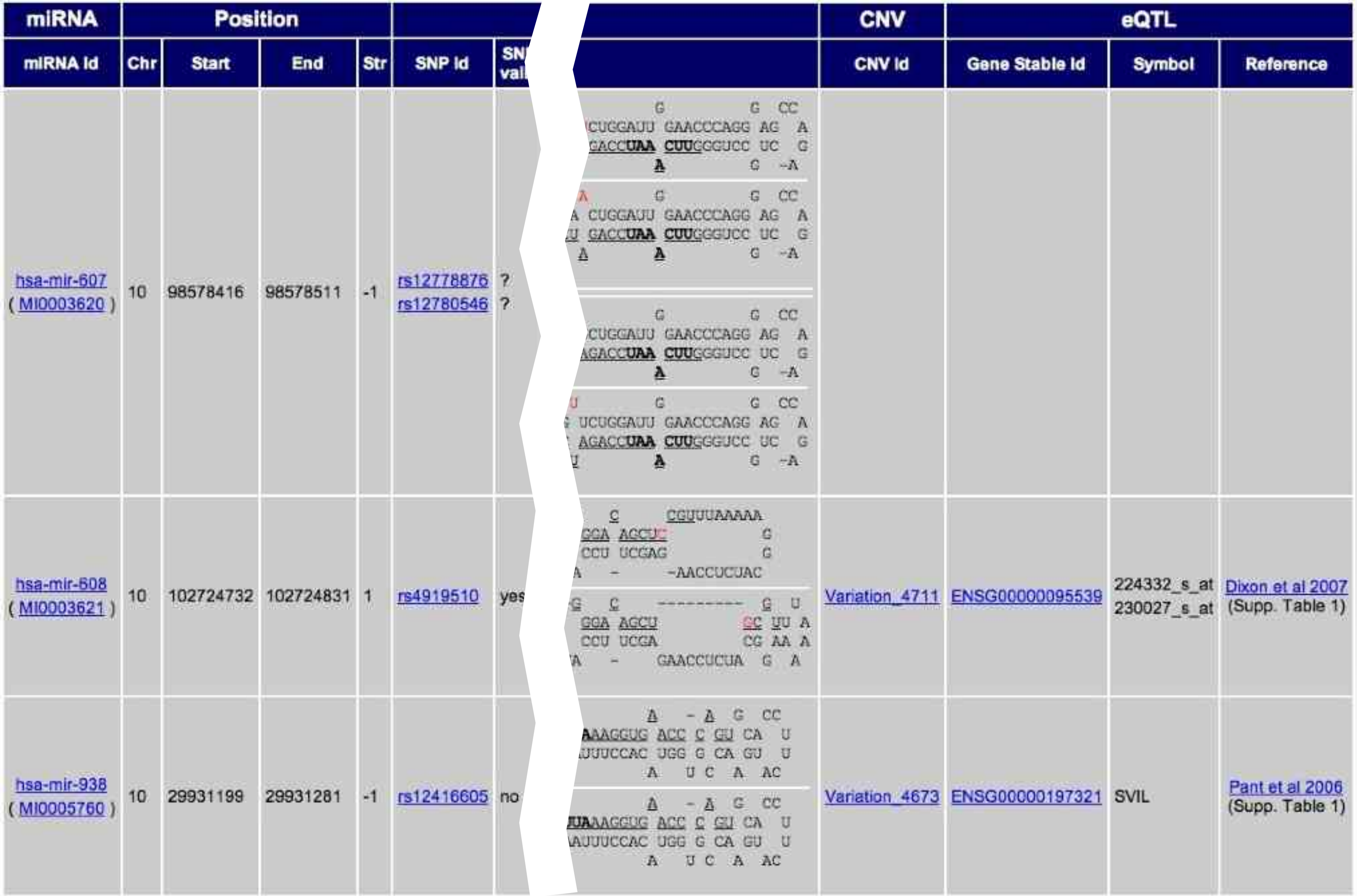
```
                         U  ACC          A   - A  G  CC
A            GAAGGUGUACCA GU   CUUAAAGGUG ACC C GU CA  U
             CUUCCGUAUGGU CA   GAAUUUCCAC UGG G CA GU  U
                         U  --A          A   U C  A  AC
```

miRNAs (zoomed output)



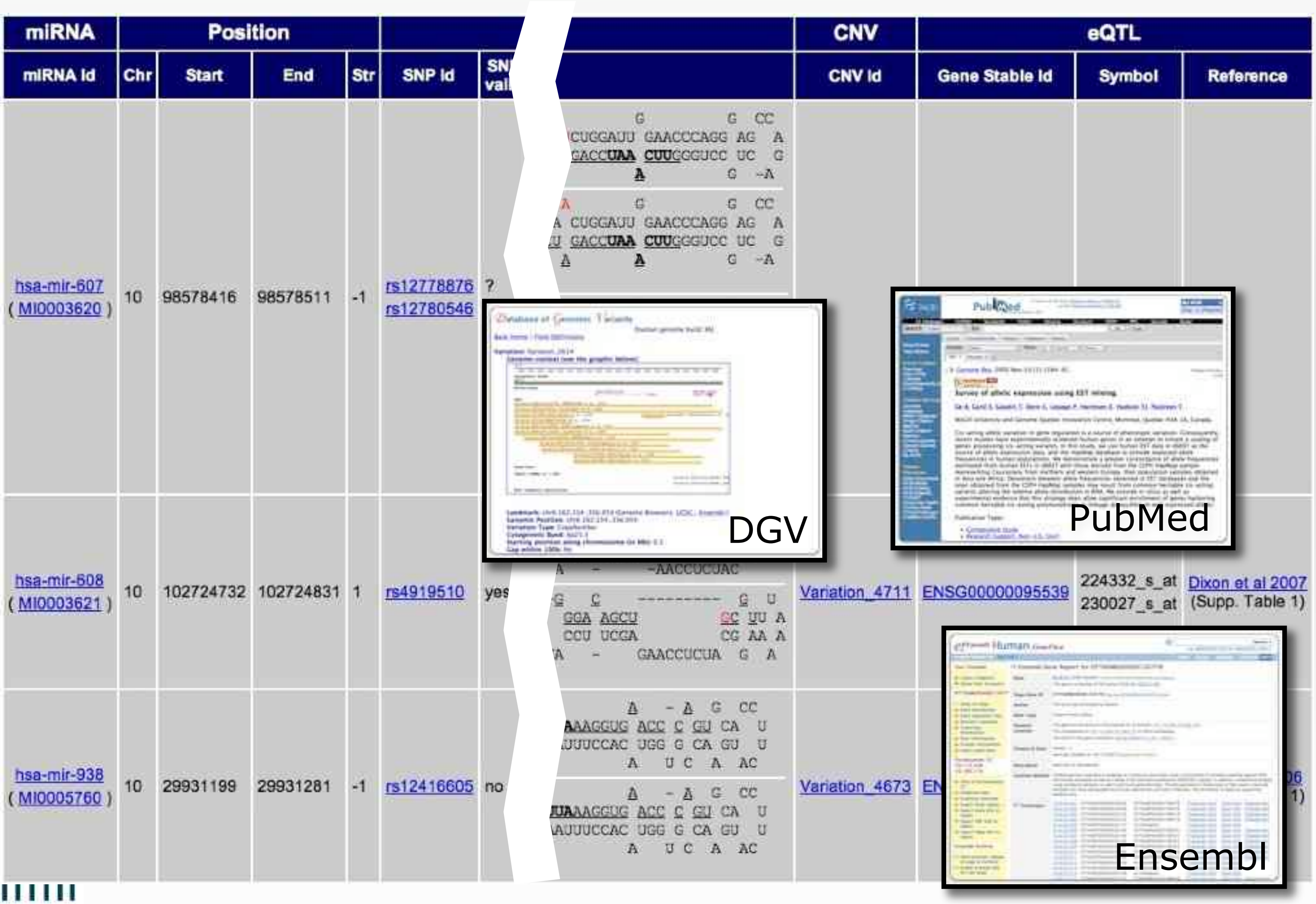
miRNA	Position				SNP				
miRNA Id	Chr	Start	End	Str	SNP Id	SNP valid	SNP site	SNP alleles	Sequence
hsa-mir-607 (MI0003620)	10						premature premature	A/T (T/A) T/A (A/T)	T (WT) UUG G G CC CCUAAAGUCACACAGGUUAUAGAUCUGGAUU GAACCCAGG AG A GGGUUUCAGUGUGUUCAAUAUCUAGACC**UAA** **CUU**GGGUCC UC G **A** G -A A UUG A G G CC CCUAAAGUCACACAGGUUAUAGA CUGGAUU GAACCCAGG AG A GGGUUUCAGUGUGUUCAAUAUCU GACC**UAA** **CUU**GGGUCC UC G A **A** G -A A (WT) UUG G G CC CCUAAAGUCACACAGGUUAUAGAUCUGGAUU GAACCCAGG AG A GGGUUUCAGUGUGUUCAAUAUCUAGACC**UAA** **CUU**GGGUCC UC G **A** G -A T UUG U G G CC CCUAAAGUCACACAGGUUAUAG UCUGGAUU GAACCCAGG AG A GGGUUUCAGUGUGUUCAAUAUC AGACC**UAA** **CUU**GGGUCC UC G U **A** G -A
hsa-mir-608 (MI0003621)	10	102724732	102724831	1	rs4919510				C (WT) GG -AA G G - **UG** G ---G C CGUUUAAAAA GCC G UG GCCA**GG** **GG** **G**U UU GGA AGCUC G UGG C AC UGGUUC CC CG AA CCU UCGAG G ACA G G U GU G ACUA - -AACCUCUAC -AA G G - **UG** G ---G C ---------- G U G UG GCCA**GG** **GG** **G**U UU GGA AGCU CC UU A C AC UGGUUC CC CG AA CCU UCGA CG AA A ACA G G U GU G ACUA - GAACCUCUA G A
hsa-mir-938 (MI0005760)	10	29931199	29931281	-1	rs12416605				U GCC A - A G CC GAAGGUGUACCA GU **CUUA**AAGGUG ACC C GU CA U CUUCCGUAUGGU CA GAAUUUCCAC UGG G CA GU U U --A A U C A AC A U ACC A - A G CC GAAGGUGUACCA GU **CUUA**AAGGUG ACC C GU CA U CUUCCGUAUGGU CA GAAUUUCCAC UGG G CA GU U U --A A U C A AC

miRNAs (zoomed output)



miRNA	Position							CNV	eQTL		
miRNA Id	Chr	Start	End	Str	SNP Id	SN val		CNV Id	Gene Stable Id	Symbol	Reference
hsa-mir-607 (MI0003620)	10	98578416	98578511	-1	rs12778876 rs12780546	? ?	G G CC CUGGAUU GAACCCAGG AG A GACC**UAA** **CUU**GGGUCC UC G A G -A A G G CC A CUGGAUU GAACCCAGG AG A U GACC**UAA** **CUU**GGGUCC UC G A A G -A G G CC CUGGAUU GAACCCAGG AG A AGACC**UAA** **CUU**GGGUCC UC G A G -A U G G CC UCUGGAUU GAACCCAGG AG A AGACC**UAA** **CUU**GGGUCC UC G U A G -A				
hsa-mir-608 (MI0003621)	10	102724732	102724831	1	rs4919510	yes	C CGUUUAAAAA GGA AGCUC G CCU UCGAG G A - -AACCUCUAC G C ---------- G U GGA AGCU GC UU A CCU UCGA CG AA A A - GAACCUCUA G A	Variation_4711	ENSG00000095539	224332_s_at 230027_s_at	Dixon et al 2007 (Supp. Table 1)
hsa-mir-938 (MI0005760)	10	29931199	29931281	-1	rs12416605	no	A - A G CC AAAGGUG ACC C GU CA U UUUCCAC UGG G CA GU U A U C A AC A - A G CC UAAAGGUG ACC C GU CA U AUUUCCAC UGG G CA GU U A U C A AC	Variation_4673	ENSG00000197321	SVIL	Pant et al 2006 (Supp. Table 1)

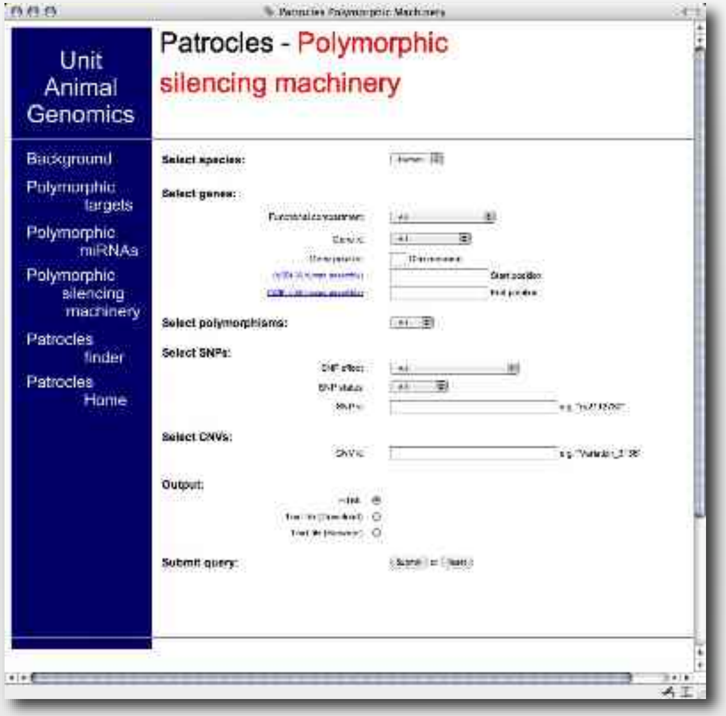
miRNAs (zoomed output)



miRNA	Position							CNV	eQTL		
miRNA Id	Chr	Start	End	Str	SNP Id	SN val		CNV Id	Gene Stable Id	Symbol	Reference
hsa-mir-607 (MI0003620)	10	98578416	98578511	-1	rs12778876 rs12780546	?	G G CC CUGGAUU GAACCCAGG AG A GACCUAA CUUGGGUCC UC G A G -A A G G CC A CUGGAUU GAACCCAGG AG A U GACCUAA CUUGGGUCC UC G A A G -A				
hsa-mir-608 (MI0003621)	10	102724732	102724831	1	rs4919510	yes	A – –AACCUCUAC G C ---------- G U GGA AGCU GC UU A CCU UCGA CG AA A A – GAACCUCUA G A	Variation_4711	ENSG00000095539	224332_s_at 230027_s_at	Dixon et al 2007 (Supp. Table 1)
hsa-mir-938 (MI0005760)	10	29931199	29931281	-1	rs12416605	no	A - A G CC AAAGGUG ACC C GU CA U UUUCCAC UGG G CA GU U A U C A AC A - A G CC UAAAAGGUG ACC C GU CA U AUUUCCAC UGG G CA GU U A U C A AC	Variation_4673	EN		06 1)

Part 3 – Polymorphic machinery

- ☐ 3 main compartments (48)
  - miRNA biogenesis (4)
  - RISC/mRNP (14)
  - P-bodies (30)


Patrocles - Polymorphic silencing machinery

Unit Animal Genomics

Background
Polymorphic targets
Polymorphic miRNAs
Polymorphic silencing machinery
Patrocles finder
Patrocles Home

Select species:

Select genes:

Select polymorphisms:

Select SNPs:

Select CNVs:

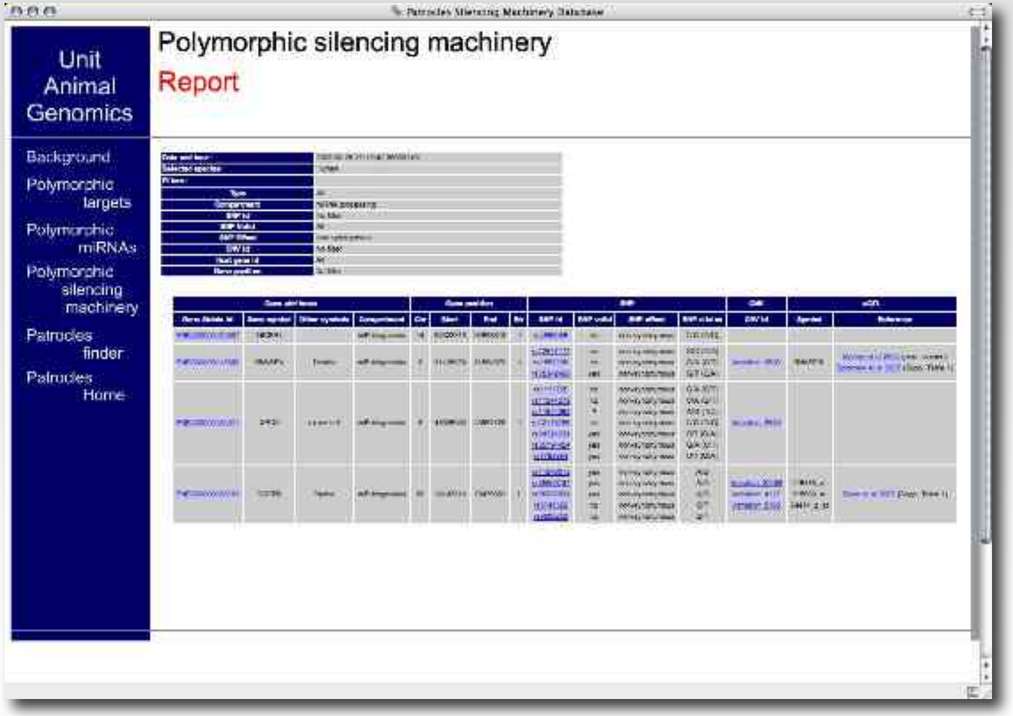
Output:

Submit query:

input form

- ☐ various filters
  - SNP
  - machinery gene

output report

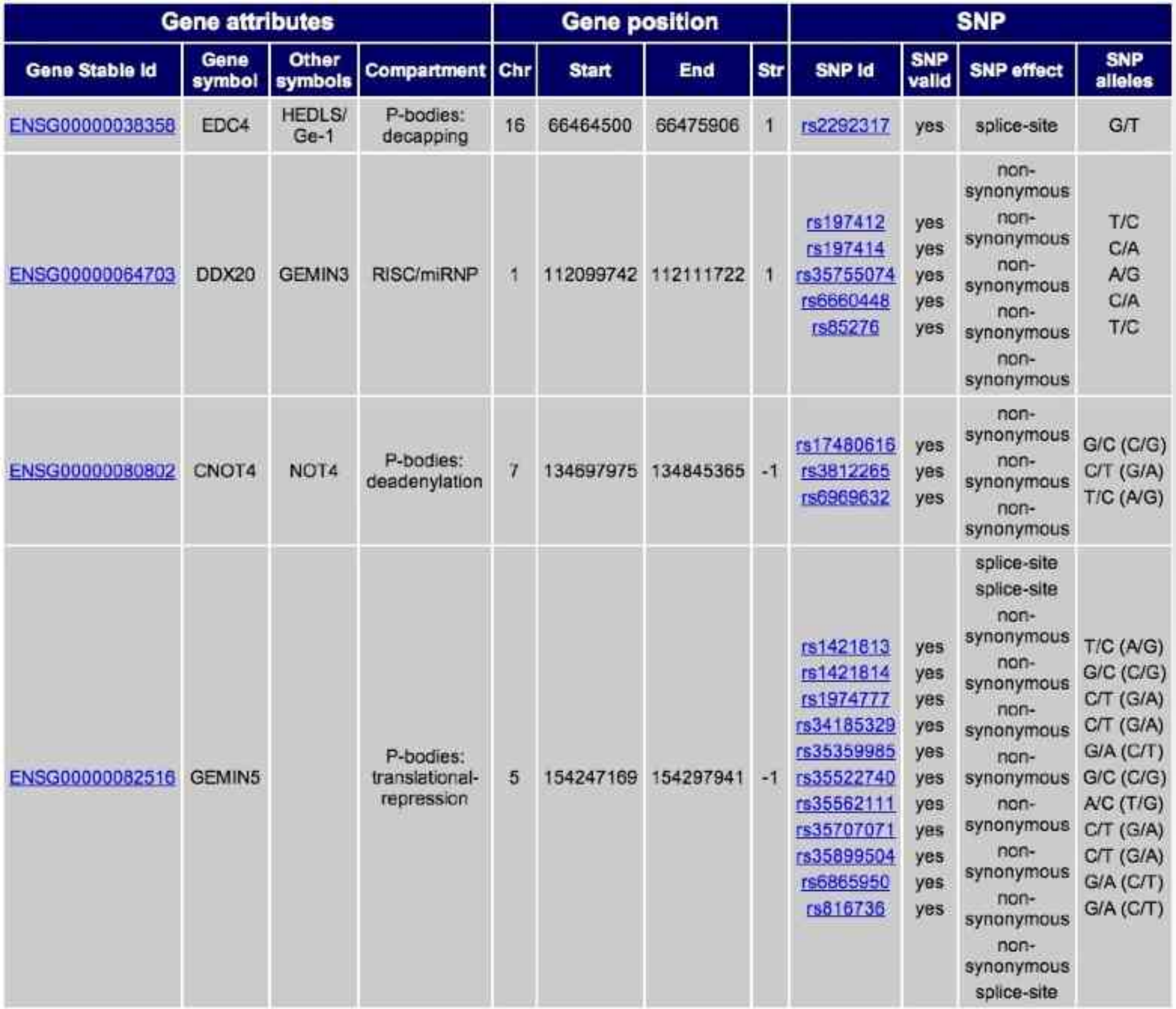

Polymorphic silencing machinery

Report

Unit Animal Genomics

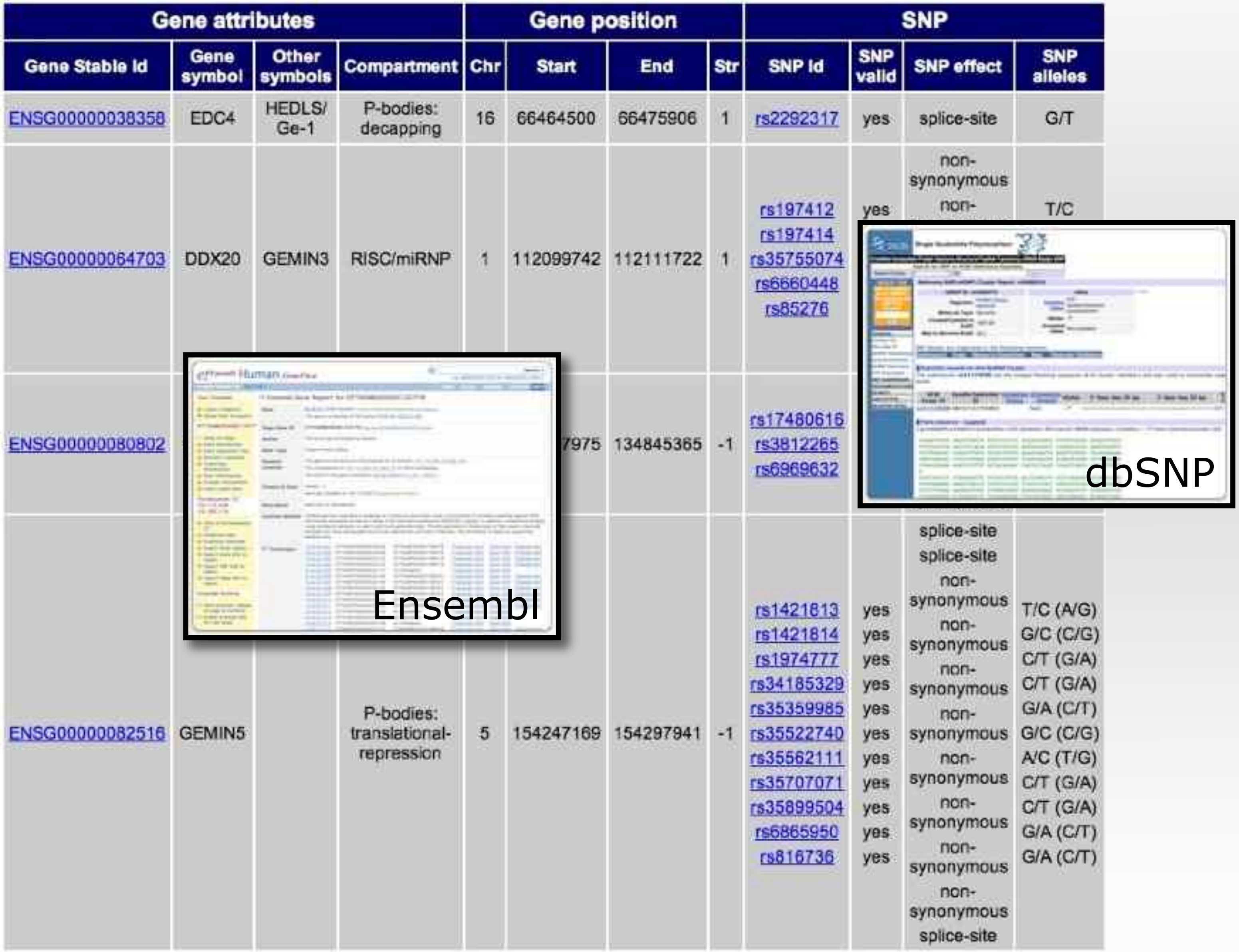
Background
Polymorphic targets
Polymorphic miRNAs
Polymorphic silencing machinery
Patrocles finder
Patrocles Home

Machinery (zoomed output)



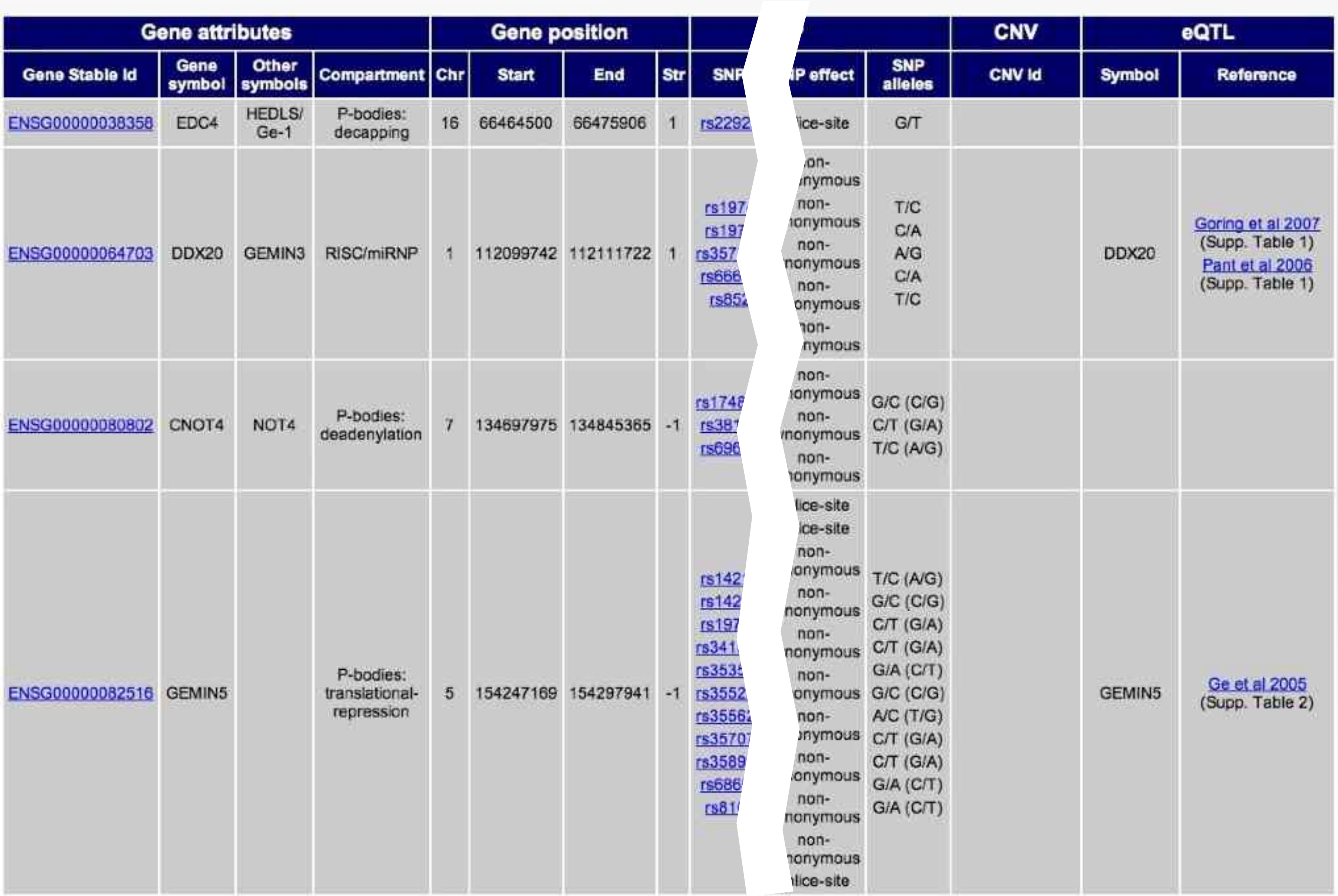
Gene attributes				Gene position				SNP			
Gene Stable Id	Gene symbol	Other symbols	Compartment	Chr	Start	End	Str	SNP Id	SNP valid	SNP effect	SNP alleles
ENSG00000038358	EDC4	HEDLS/ Ge-1	P-bodies: decapping	16	66464500	66475906	1	rs2292317	yes	splice-site	G/T
ENSG00000064703	DDX20	GEMIN3	RISC/miRNP	1	112099742	112111722	1	rs197412 rs197414 rs35755074 rs6660448 rs85276	yes yes yes yes yes	non-synonymous non-synonymous non-synonymous non-synonymous non-synonymous	T/C C/A A/G C/A T/C
ENSG00000080802	CNOT4	NOT4	P-bodies: deadenylation	7	134697975	134845365	-1	rs17480616 rs3812265 rs6969632	yes yes yes	non-synonymous non-synonymous non-synonymous	G/C (C/G) C/T (G/A) T/C (A/G)
ENSG00000082516	GEMIN5		P-bodies: translational-repression	5	154247169	154297941	-1	rs1421813 rs1421814 rs1974777 rs34185329 rs35359985 rs35522740 rs35562111 rs35707071 rs35899504 rs6865950 rs816736	yes yes yes yes yes yes yes yes yes yes yes	splice-site splice-site non-synonymous non-synonymous non-synonymous non-synonymous non-synonymous non-synonymous non-synonymous non-synonymous splice-site	T/C (A/G) G/C (C/G) C/T (G/A) C/T (G/A) G/A (C/T) G/C (C/G) A/C (T/G) C/T (G/A) C/T (G/A) G/A (C/T) G/A (C/T)

Machinery (zoomed output)



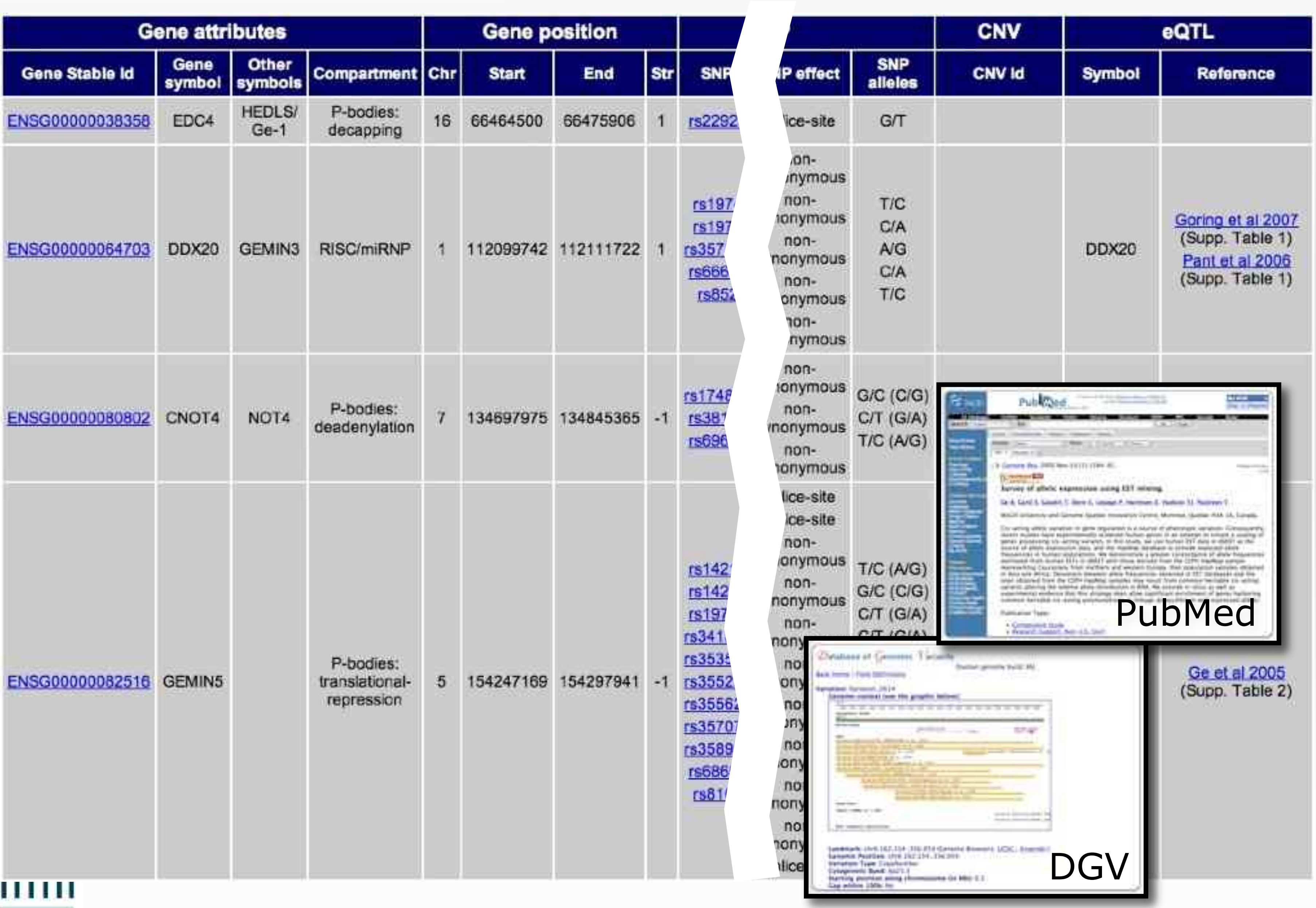
Gene attributes				Gene position				SNP			
Gene Stable Id	Gene symbol	Other symbols	Compartment	Chr	Start	End	Str	SNP Id	SNP valid	SNP effect	SNP alleles
ENSG00000038358	EDC4	HEDLS/ Ge-1	P-bodies: decapping	16	66464500	66475906	1	rs2292317	yes	splice-site	G/T
ENSG00000064703	DDX20	GEMIN3	RISC/miRNP	1	112099742	112111722	1	rs197412 rs197414 rs35755074 rs6660448 rs85276	yes	non-synonymous non-	T/C
ENSG00000080802					7975	134845365	-1	rs17480616 rs3812265 rs6969632			
ENSG00000082516	GEMIN5		P-bodies: translational-repression	5	154247169	154297941	-1	rs1421813 rs1421814 rs1974777 rs34185329 rs35359985 rs35522740 rs35562111 rs35707071 rs35899504 rs6865950 rs816736	yes yes yes yes yes yes yes yes yes yes yes	splice-site splice-site non-synonymous non-synonymous non-synonymous non-synonymous non-synonymous non-synonymous non-synonymous non-synonymous splice-site	T/C (A/G) G/C (C/G) C/T (G/A) C/T (G/A) G/A (C/T) G/C (C/G) A/C (T/G) C/T (G/A) C/T (G/A) G/A (C/T) G/A (C/T)

Machinery (zoomed output)



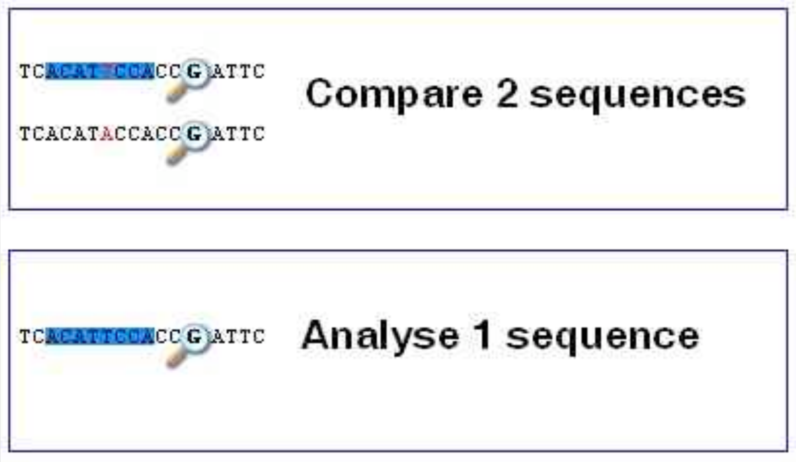
Gene attributes				Gene position							CNV	eQTL	
Gene Stable Id	Gene symbol	Other symbols	Compartment	Chr	Start	End	Str	SNP	P effect	SNP alleles	CNV Id	Symbol	Reference
ENSG00000038358	EDC4	HEDLS/ Ge-1	P-bodies: decapping	16	66464500	66475906	1	rs2292	lice-site	G/T			
ENSG00000064703	DDX20	GEMIN3	RISC/miRNP	1	112099742	112111722	1	rs197 rs197 rs357 rs666 rs852	on- nymous non- nonymous non- nonymous non- onymous on- nymous	T/C C/A A/G C/A T/C		DDX20	Goring et al 2007 (Supp. Table 1) Pant et al 2006 (Supp. Table 1)
ENSG00000080802	CNOT4	NOT4	P-bodies: deadenylation	7	134697975	134845365	-1	rs1748 rs381 rs696	non- nonymous non- ynonymous non- nonymous	G/C (C/G) C/T (G/A) T/C (A/G)			
ENSG00000082516	GEMIN5		P-bodies: translational-repression	5	154247169	154297941	-1	rs142 rs142 rs197 rs341 rs3535 rs3552 rs3556 rs3570 rs3589 rs686 rs81	lice-site ice-site non- onymous non- nonymous non- nonymous non- onymous non- onymous non- onymous non- nonymous non- nonymous lice-site	T/C (A/G) G/C (C/G) C/T (G/A) C/T (G/A) G/A (C/T) G/C (C/G) A/C (T/G) C/T (G/A) C/T (G/A) G/A (C/T) G/A (C/T)		GEMIN5	Ge et al 2005 (Supp. Table 2)

Machinery (zoomed output)

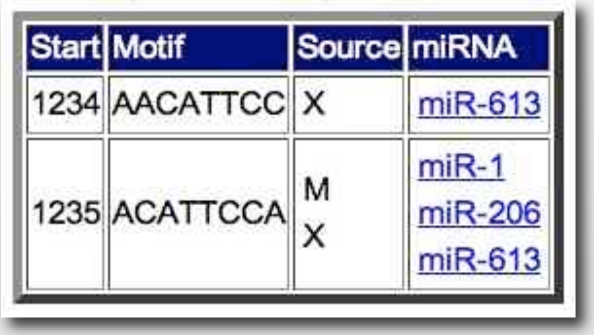


Gene attributes				Gene position							CNV	eQTL	
Gene Stable Id	Gene symbol	Other symbols	Compartment	Chr	Start	End	Str	SNP	P effect	SNP alleles	CNV Id	Symbol	Reference
ENSG00000038358	EDC4	HEDLS/ Ge-1	P-bodies: decapping	16	66464500	66475906	1	rs2292	ice-site	G/T			
ENSG00000064703	DDX20	GEMIN3	RISC/miRNP	1	112099742	112111722	1	rs197 rs197 rs357 rs666 rs852	on- nymous non- ionymous non- nonymous non- onymous ion- nymous	T/C C/A A/G C/A T/C		DDX20	Goring et al 2007 (Supp. Table 1) Pant et al 2006 (Supp. Table 1)
ENSG00000080802	CNOT4	NOT4	P-bodies: deadenylation	7	134697975	134845365	-1	rs1748 rs381 rs696	non- ionymous non- ynonymous non- ionymous	G/C (C/G) C/T (G/A) T/C (A/G)			
ENSG00000082516	GEMIN5		P-bodies: translational-repression	5	154247169	154297941	-1	rs142 rs142 rs197 rs341 rs3535 rs3552 rs3556 rs3570 rs3589 rs686 rs81	lice-site ice-site non- ionymous non- nonymous non- nony no ony no ony no ony no nony no nony lice	T/C (A/G) G/C (C/G) C/T (G/A) C/T (C/A)			Ge et al 2005 (Supp. Table 2)

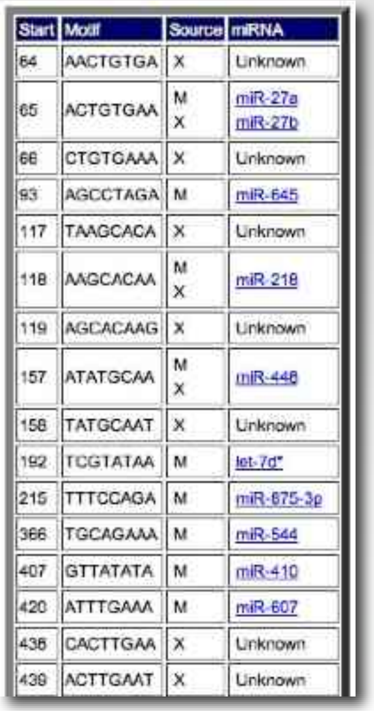
Part 4 – Custom 3′-UTR sequences



3′-UTR of MSTN allele (Texel *vs.* WT)



Start	Motif	Source	miRNA
1234	AACATTCC	X	miR-613
1235	ACATTCCA	M X	miR-1 miR-206 miR-613

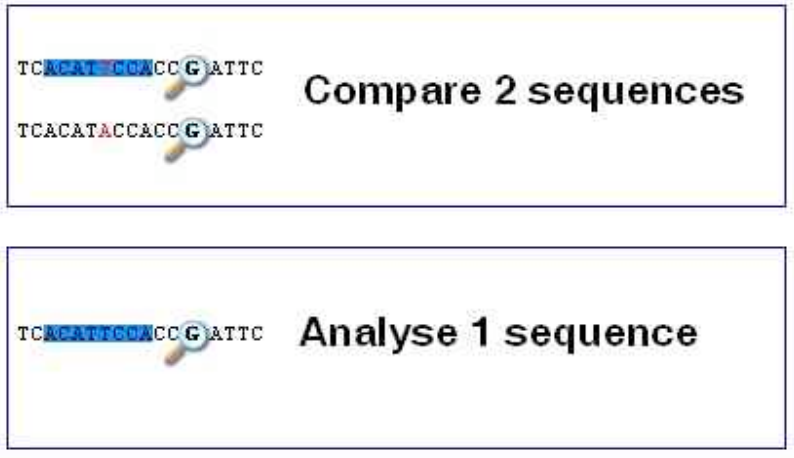


Start	Motif	Source	miRNA
64	AACTGTGA	X	Unknown
65	ACTGTGAA	M X	miR-27a miR-27b
66	CTGTGAAA	X	Unknown
93	AGCCTAGA	M	miR-645
117	TAAGCACA	X	Unknown
118	AAGCACAA	M X	miR-218
119	AGCACAAG	X	Unknown
157	ATATGCAA	M X	miR-448
158	TATGCAAT	X	Unknown
192	TCGTATAA	M	let-7d*
215	TTTCCAGA	M	miR-875-3p
366	TGCAGAAA	M	miR-544
407	GTTATATA	M	miR-410
420	ATTTGAAA	M	miR-607
436	CACTTGAA	X	Unknown
439	ACTTGAAT	X	Unknown

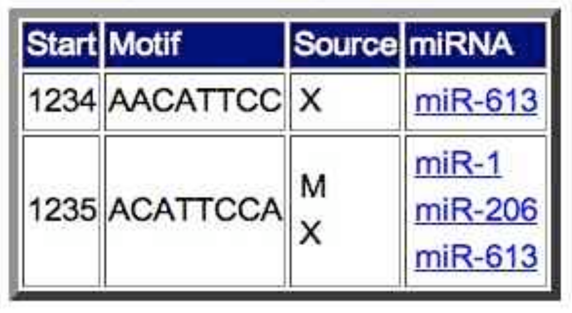
useful for species
not yet fully annotated...

or even not yet fully sequenced!

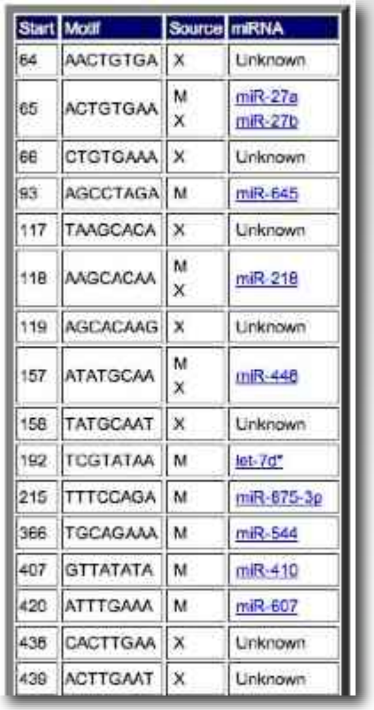
Part 4 – Custom 3’-UTR sequences



3’-UTR of MSTN allele (Texel *vs.* WT)



Start	Motif	Source	miRNA
1234	AACATTCC	X	miR-613
1235	ACATTCCA	M X	miR-1 miR-206 miR-613



Start	Motif	Source	miRNA
64	AACTGTGA	X	Unknown
65	ACTGTGAA	M X	miR-27a miR-27b
66	CTGTGAAA	X	Unknown
93	AGCCTAGA	M	miR-645
117	TAAGCACA	X	Unknown
118	AAGCACAA	M X	miR-218
119	AGCACAAG	X	Unknown
157	ATATGCAA	M X	miR-448
158	TATGCAAT	X	Unknown
192	TCGTATAA	M	let-7d*
215	TTTCCAGA	M	miR-875-3p
366	TGCAGAAA	M	miR-544
407	GTTATATA	M	miR-410
420	ATTTGAAA	M	miR-607
436	CACTTGAA	X	Unknown
439	ACTTGAAT	X	Unknown

Conclusions

- ☐ Tool to help prioritizing wet-lab experiments
- ☐ Reports of polymorphic target sites / miRNAs
  - SLITRK1 (Abelson et al., Science, 2005)
  - AGTR1 (Sethupathy et al., AJHG, 2007)
  - AGTR1 (Martin et al., JBC, 2007)
  - DHFR (Mishra et al., PNAS, 2007)
  - HLA-G (Tan et al., AJHG, 2007)
  - REEP1, FGF20 (ASHG meeting 2007) ...
  - miR125a (Duan et al., HMG, 2007) ...
- ☐ We hope it will be useful...

http://www.patrocles.org/