



# L'analyse de marqueurs minisatellites variables suggèrent de fréquents échanges génétiques entre sous-espèces de la souris commune

Eric Rivals

## ► To cite this version:

Eric Rivals. L'analyse de marqueurs minisatellites variables suggèrent de fréquents échanges génétiques entre sous-espèces de la souris commune. Bioinformatique, Modélisation des Systèmes Biologiques, Oct 2007, Paris, France. lirmm-00193841

HAL Id: lirmm-00193841

<https://hal-lirmm.ccsd.cnrs.fr/lirmm-00193841>

Submitted on 4 Dec 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

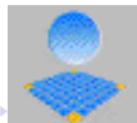
# Minisatellite Markers Suggests Frequent Genetic Exchanges among House Mouse Subspecies

Eric Rivals

LIRMM, CNRS, Université Montpellier 2

rivals\_AT\_lirmm.fr

<http://www.lirmm.fr/~rivals>



# Outline

1. Minisatellite data

2. Methods

2.1 Protocol

2.2 Molecular Divergence Estimation - Alignment

2.3 Robustness and Confidence

3. Results

3.1 Coalescence

3.2 Intruders

4. Conclusion

# Minisatellites (MS)

- **Tandem repeat** loci present in genome of all kingdom

## Minisatellites (MS)

- ▶ **Tandem repeat** loci present in genome of all kingdom
- ▶ > 5% in bacteria *E. ruminatum*

## Minisatellites (MS)

- ▶ **Tandem repeat** loci present in genome of all kingdom
- ▶ > 5% in bacteria *E. ruminatum*
- ▶ pattern between 7 and 100 bps, up to 20 Kb long

## Minisatellites (MS)

- ▶ **Tandem repeat** loci present in genome of all kingdom
- ▶ > 5% in bacteria *E. ruminatum*
- ▶ pattern between 7 and 100 bps, up to 20 Kb long
- ▶ **Polymorphism:** Hypervariable in *H.s.*, (> 0.5% per gamete)

# Minisatellites (MS)

- ▶ **Tandem repeat** loci present in genome of all kingdom
- ▶ > 5% in bacteria *E. ruminatum*
- ▶ pattern between 7 and 100 bps, up to 20 Kb long
- ▶ **Polymorphism:** Hypervariable in *H.s.*, ( $> 0.5\%$  per gamete)  
Simply variable in *M.m.*(add  $> 3$  units:  $< 5.10^{-6}$  per gamete)

# Minisatellites (MS)

- ▶ **Tandem repeat** loci present in genome of all kingdom
- ▶ > 5% in bacteria *E. ruminatum*
- ▶ pattern between 7 and 100 bps, up to 20 Kb long
- ▶ **Polymorphism:** Hypervariable in *H.s.*, ( $> 0.5\%$  per gamete)  
Simply variable in *M.m.*(add  $> 3$  units:  $< 5 \cdot 10^{-6}$  per gamete)
- ▶ Mouse: simple intra-allelic tandem duplication and contraction  
[Bois et al., 2002]

# Minisatellites (MS)

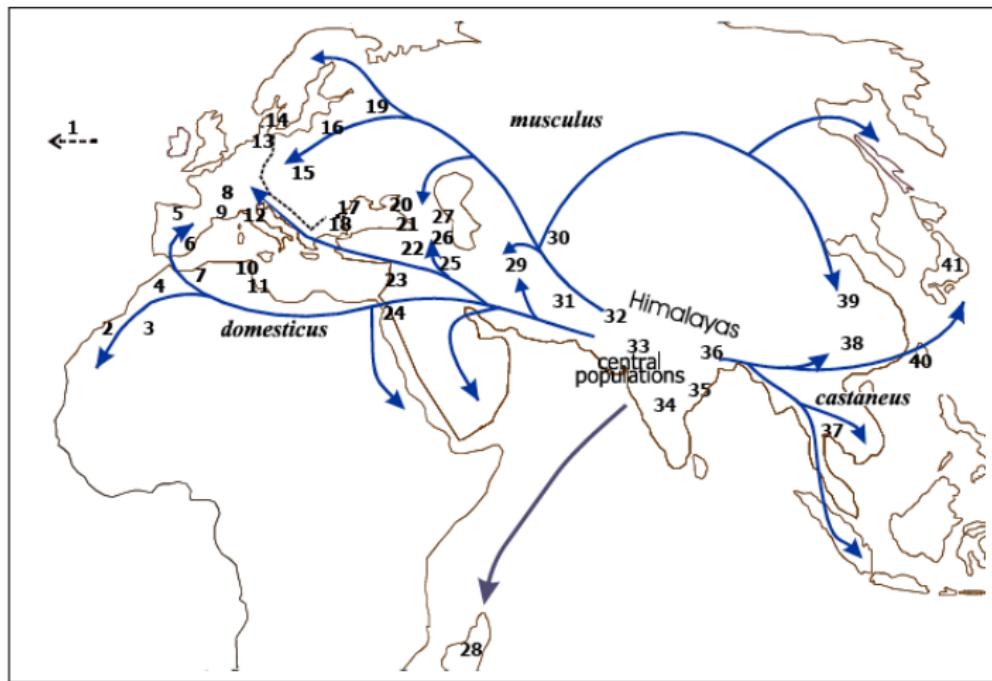
- ▶ **Tandem repeat** loci present in genome of all kingdom
- ▶ > 5% in bacteria *E. ruminatum*
- ▶ pattern between 7 and 100 bps, up to 20 Kb long
- ▶ **Polymorphism:** Hypervariable in *H.s.*, ( $> 0.5\%$  per gamete)  
Simply variable in *M.m.*(add  $> 3$  units:  $< 5 \cdot 10^{-6}$  per gamete)
- ▶ Mouse: simple intra-allelic tandem duplication and contraction  
[Bois et al., 2002]
- ▶ MS informative for intra-species evolution



## Minisatellite data

- ▶ Four minisatellite loci: MMS 24, 26, 80, and 30 respectively on chromosomes 7 (22 cM), 9 (68 cM and 79 cM), and X (43 cM)
- ▶ Panel of 116 individuals of various geographical origins
- ▶ Maps obtained by MVR-PCR as in [Bois et al., 2002]
- ▶ High diversity in length and array structure  
haplotypic diversity ( $H_e$ ) in [0.90, 0.99]

# Geographical origin of sampled wild mice



World-wide sample at four distinct MS loci.

No other equivalent data set (SNPs or microsatellites)

# Mouse minisatellite maps of MMS30 (X chr)

**Repeat unit:** 39 bps

Variant code    sequence

$K = \text{"aggagattcaGttcaCaCtatacagaagatggtgtcagc"}$

$L = \text{"aggagattcaCttcaGaGtatacagaagatggtgtcagc"}$

# Mouse minisatellite maps of MMS30 (X chr)

**Repeat unit:** 39 bps

Variant code    sequence

$K$  = "aggagattcaGttcaCaCtatacagaagatggtgtcagc"

$L$  = "aggagattcaCttcaGaGtatacagaagatggtgtcagc"

Example of **maps** in wild mice [Bois et al., 2002]:

espèce origine carte

---

**SPR** Fran KGKGKLKHLKKLLKYKG

**CEN** Inde GKKKKWGKKYKWKGWGHoGoKWKKKoLYY

**CEN** Inde GKKKKWGKKYKWKGWGHoGoKWKKKoKYY

**DOM** Geor GYKKKWGKLKYoWKGWGKoGGYWYKKoKYYYKG

**DOM** Ocea GYKKKWGKLKYoWKGWGKoGGYWYKKoKKKYYYKG

**MUS** Geor GYYKGYKYKGYKKKGKoKYoWKYYKG

---

# Methods

# Protocol of the analysis

Input: set of sequences (maps)

# Protocol of the analysis

Input: set of sequences (maps)

1. Comparison all against all  $\Rightarrow$  pairwise distance matrix

# Protocol of the analysis

Input: set of sequences (maps)

1. Comparison all against all  $\Rightarrow$  pairwise distance matrix  
*MS\_Align* [Bérard, Rivals, 2003]

# Protocol of the analysis

Input: set of sequences (maps)

1. Comparison all against all  $\Rightarrow$  pairwise distance matrix  
*MS\_Align* [Bérard, Rivals, 2003]
2. Inference of evolutionary tree from the distances

# Protocol of the analysis

Input: set of sequences (maps)

1. Comparison all against all  $\Rightarrow$  pairwise distance matrix  
*MS\_Align* [Bérard, Rivals, 2003]
2. Inference of evolutionary tree from the distances  
*FastME* [Desper, Gascuel, 2002]

# Protocol of the analysis

Input: set of sequences (maps)

1. Comparison all against all  $\Rightarrow$  pairwise distance matrix  
*MS\_Align* [Bérard, Rivals, 2003]
2. Inference of evolutionary tree from the distances  
*FastME* [Desper, Gascuel, 2002]
3. Test robustness of the trees w.r.t. alignment parameters

# Protocol of the analysis

Input: set of sequences (maps)

1. Comparison all against all  $\Rightarrow$  pairwise distance matrix  
*MS\_Align* [Bérard, Rivals, 2003]
2. Inference of evolutionary tree from the distances  
*FastME* [Desper, Gascuel, 2002]
3. Test robustness of the trees w.r.t. alignment parameters  
Criterion: Percentage of explained variance (VAF)

# Protocol of the analysis

Input: set of sequences (maps)

1. Comparison all against all  $\Rightarrow$  pairwise distance matrix  
*MS\_Align* [Bérard, Rivals, 2003]
2. Inference of evolutionary tree from the distances  
*FastME* [Desper, Gascuel, 2002]
3. Test robustness of the trees w.r.t. alignment parameters  
Criterion: Percentage of explained variance (VAF)
4. Assess confidence of the tree internal nodes

# Protocol of the analysis

Input: set of sequences (maps)

1. Comparison all against all  $\Rightarrow$  pairwise distance matrix  
*MS\_Align* [Bérard, Rivals, 2003]
2. Inference of evolutionary tree from the distances  
*FastME* [Desper, Gascuel, 2002]
3. Test robustness of the trees w.r.t. alignment parameters  
Criterion: Percentage of explained variance (VAF)
4. Assess confidence of the tree internal nodes  
Criterion: Rate of elementary well designed quartets (Re)  
*Qualitree* [Garreta, Guénoche, 2000]

## Evolutionary model

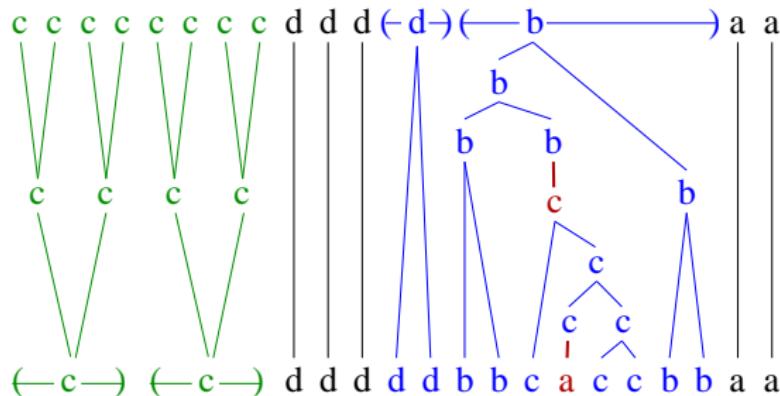
- ▶ Substitution : WGY → WKY
- ▶ Deletion: WGY → WY
- ▶ Insertion (dual): WY → WGY
- ▶ Tandem duplication: WKY → WKKY
- ▶ Tandem contraction (dual): WKKY → WKY

## Evolutionary model

- ▶ Substitution : WGY → WKY
- ▶ Deletion: WGY → WY
- ▶ Insertion (dual): WY → WGY
- ▶ Tandem duplication: WKY → WKKY
- ▶ Tandem contraction (dual): WKKY → WKY

⇒ variation in their number of units

## Example of an alignment of 2 maps



- ▶ contexte dependency - non commutativity
- ▶ Version 1 : correspondance with a graph algorithm  
stable max in a *circle graph* [Bérard Rivals *J Comp Bio* 2003]
- ▶ Version 2 : dynamic programming  
[Bérard et al., *Evol. Bioinfo.* 2006]

## Percentage of explained variance (VAF)

$$\text{VAF} = 1 - \frac{\sum_{(i,j):i < j} (D(i,j) - T(i,j))^2}{\sum_{(i,j):i < j} (D(i,j) - D_m)^2}$$

where

$D(i,j)$  : alignment distance between  $i$  and  $j$

$T(i,j)$  : tree distance between  $i$  and  $j$

$D_m$  : average alignment distance over all pairs  $(i,j)$

Value in  $[0, 1]$

## Rate of elementary well designed quartets (Re)

For an internal edge  $e$ , for all quartets  $(i, j, k, l)$   
s.t.  $e$  splits  $(i, j)$  and  $(k, l)$ :

$R(e) = \text{percentage of these quartets satisfying}$

$$(D(i, j) + D(k, l)) < \min(D(i, l) + D(j, k), (D(i, k) + D(j, l)))$$

where

$D(i, j)$  : alignment distance between  $i$  and  $j$

Value in  $[0, 1]$

# Results

## Identical alleles

Locus	2	3	$\geq 4$	Total	$\neq$ origin	$\neq$ subspecies
MMS 24	7	4	1	27	8 (18)	0 (1)
MMS 26	6	2	3	36	10 (26)	5 (1)
MMS 30	9	2	1	38	8 (18)	2 (0)
MMS 80	10	8	0	44	10 (22)	0 (0)

## Identical alleles

Locus	2	3	$\geq 4$	Total	$\neq$ origin	$\neq$ subspecies
MMS 24	7	4	1	27	8 (18)	0 (1)
MMS 26	6	2	3	36	10 (26)	5 (1)
MMS 30	9	2	1	38	8 (18)	2 (0)
MMS 80	10	8	0	44	10 (22)	0 (0)

### Example

At MMS 30: DOM\_BULG\_Vlas\_DBV,  
DOM\_TUNI\_Monastir\_22MO, and SPR\_MARO\_Azzemour\_9852  
GYKKKGWGKoGGYWYKKoKKKYYYYKG

## Identical alleles

Locus	2	3	$\geq 4$	Total	$\neq$ origin	$\neq$ subspecies
MMS 24	7	4	1	27	8 (18)	0 (1)
MMS 26	6	2	3	36	10 (26)	5 (1)
MMS 30	9	2	1	38	8 (18)	2 (0)
MMS 80	10	8	0	44	10 (22)	0 (0)

### Example

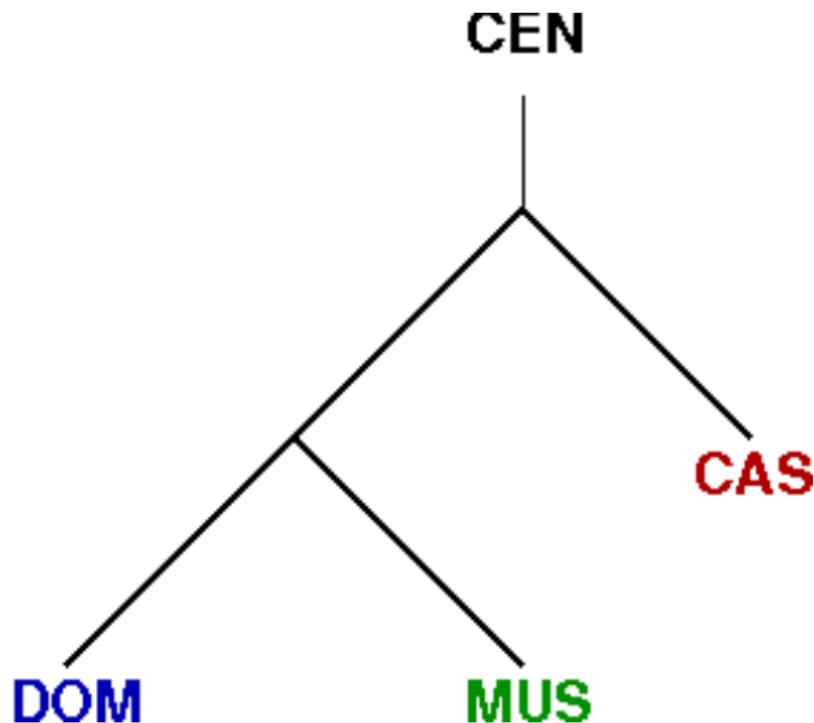
At MMS 30: DOM\_BULG\_Vlas\_DBV,  
DOM\_TUNI\_Monastir\_22MO, and SPR\_MARO\_Azzemour\_9852  
GYKKKGWGKoGGYWYKKoKKKYYYYKG

At MMS 26: CEN\_INDE\_Dehli\_DHA; DOM\_OCEA\_Tahiti\_DOT  
YGGGGGGGGAGGGAGAAGGYAAGGGAAAAGAGAAGAAGGGG

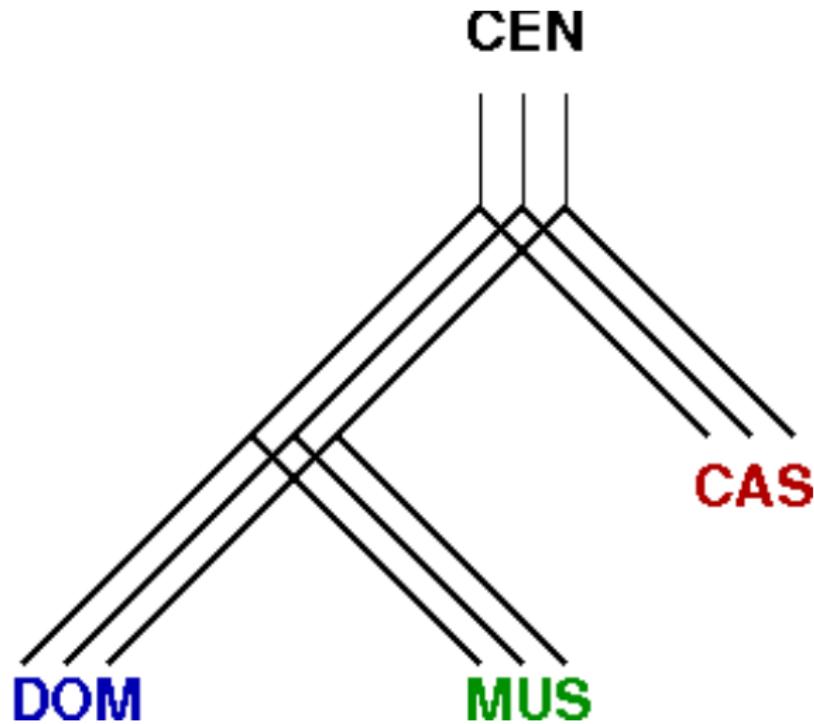
## General features

- ▶ Average distance between subspecies  
((CAS, CEN) (DOM, MUS))  
identical at all loci
- ▶ Questions:
  1. isolated subspecies: clear separation in the tree
  2. subspecies in contact: mix of haplotypes, no separation.

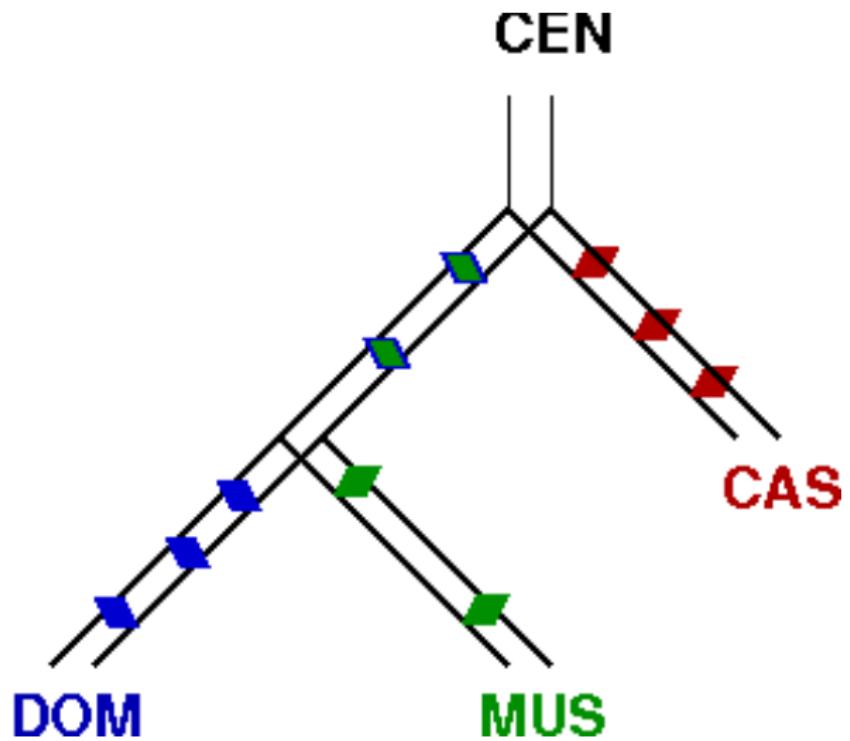
What do we expect?



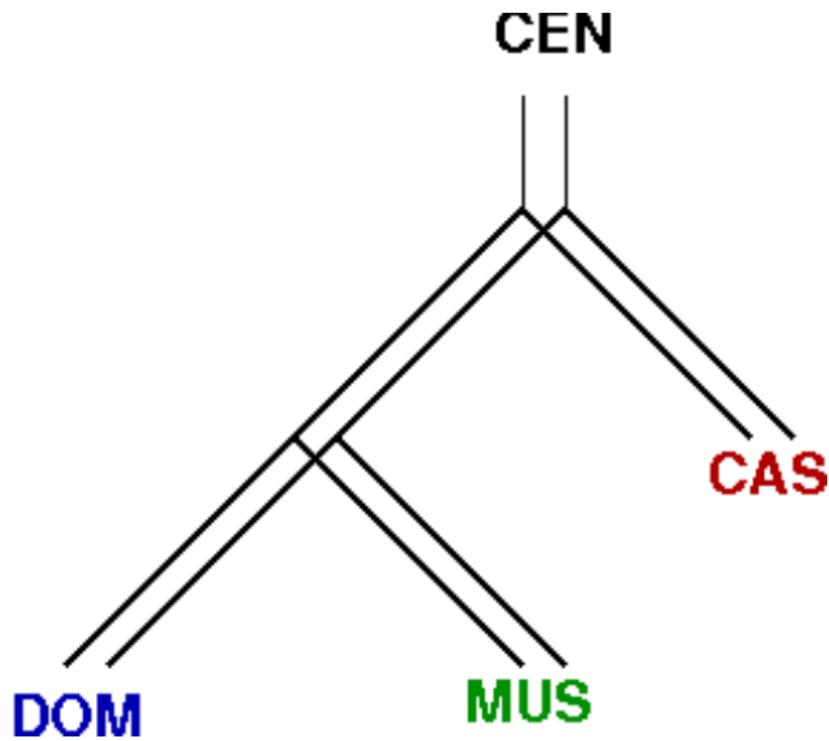
What do we expect?



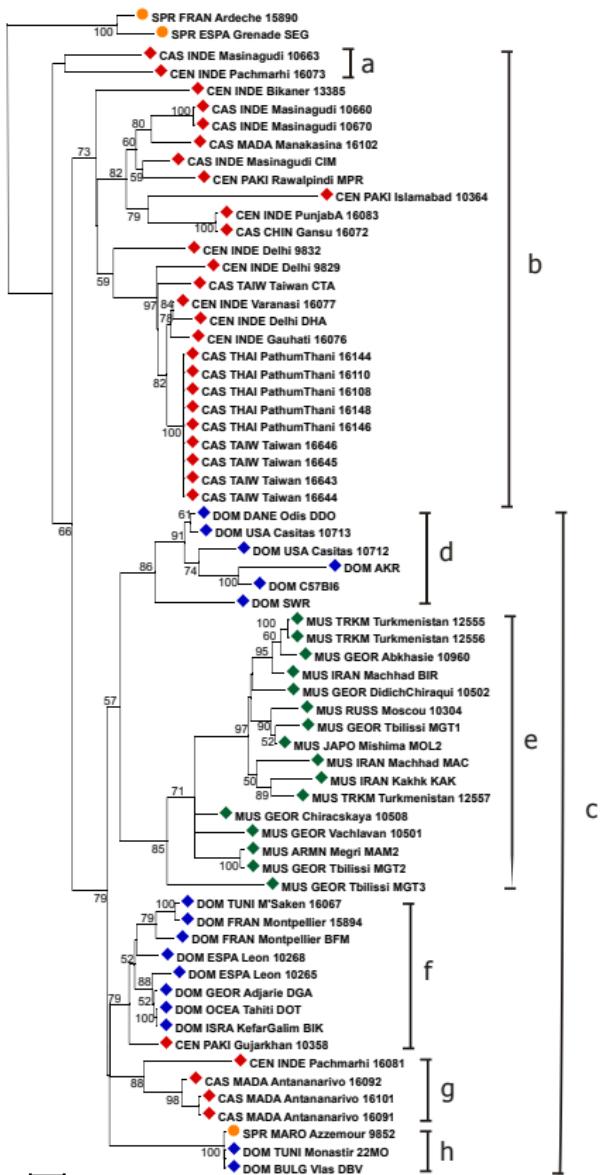
What do we expect?



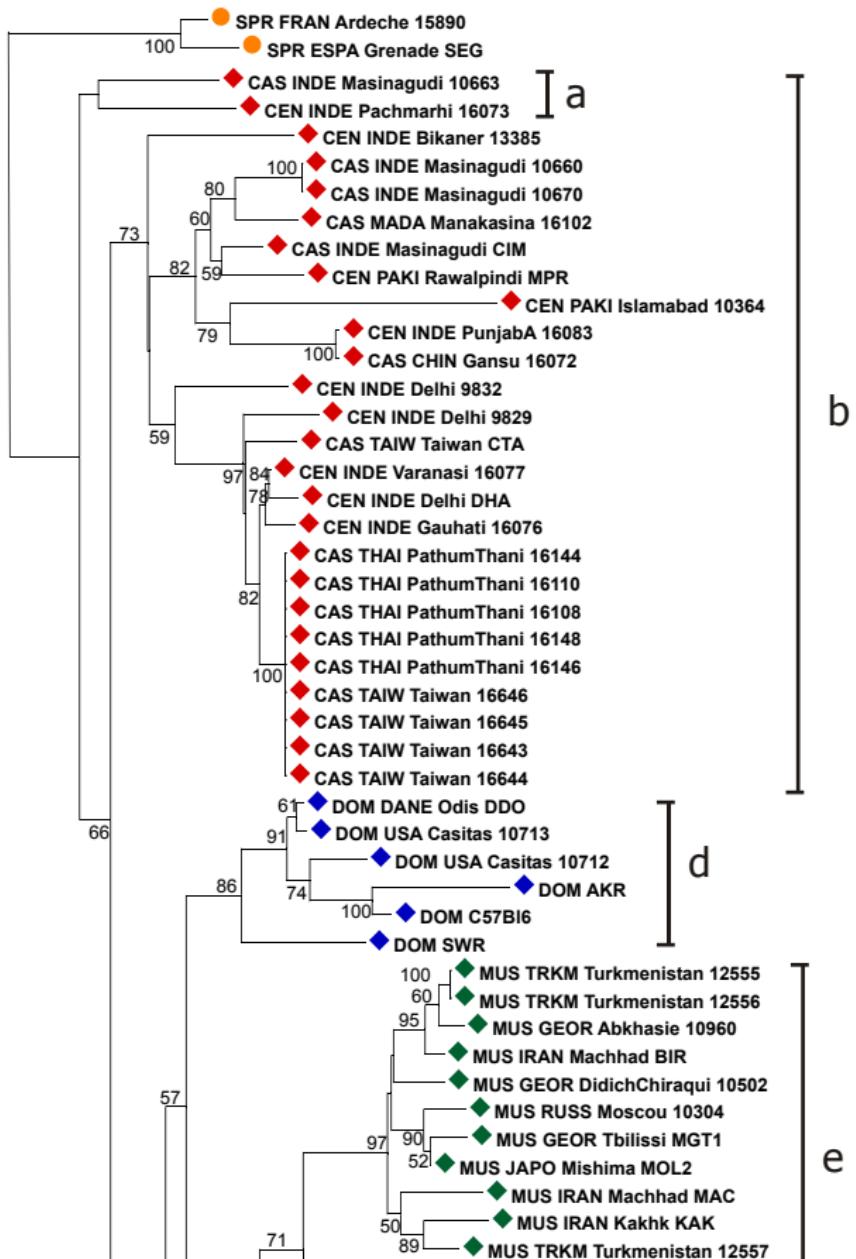
What do we expect?



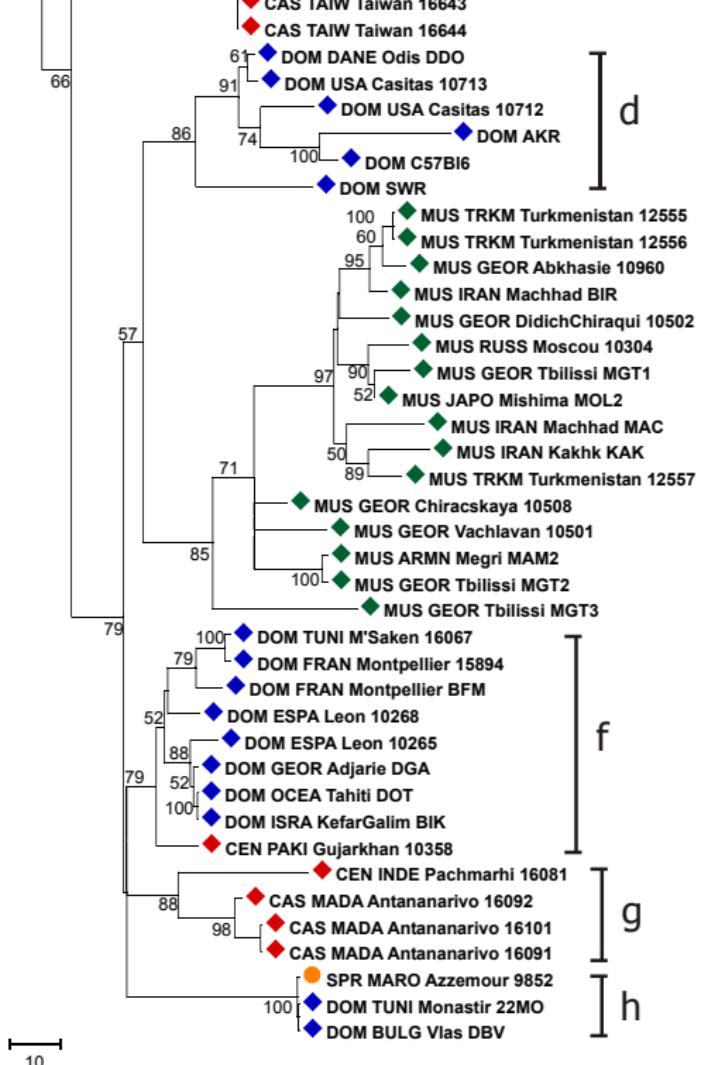
# MMS 30 coalescence



# MMS 30 coalescence



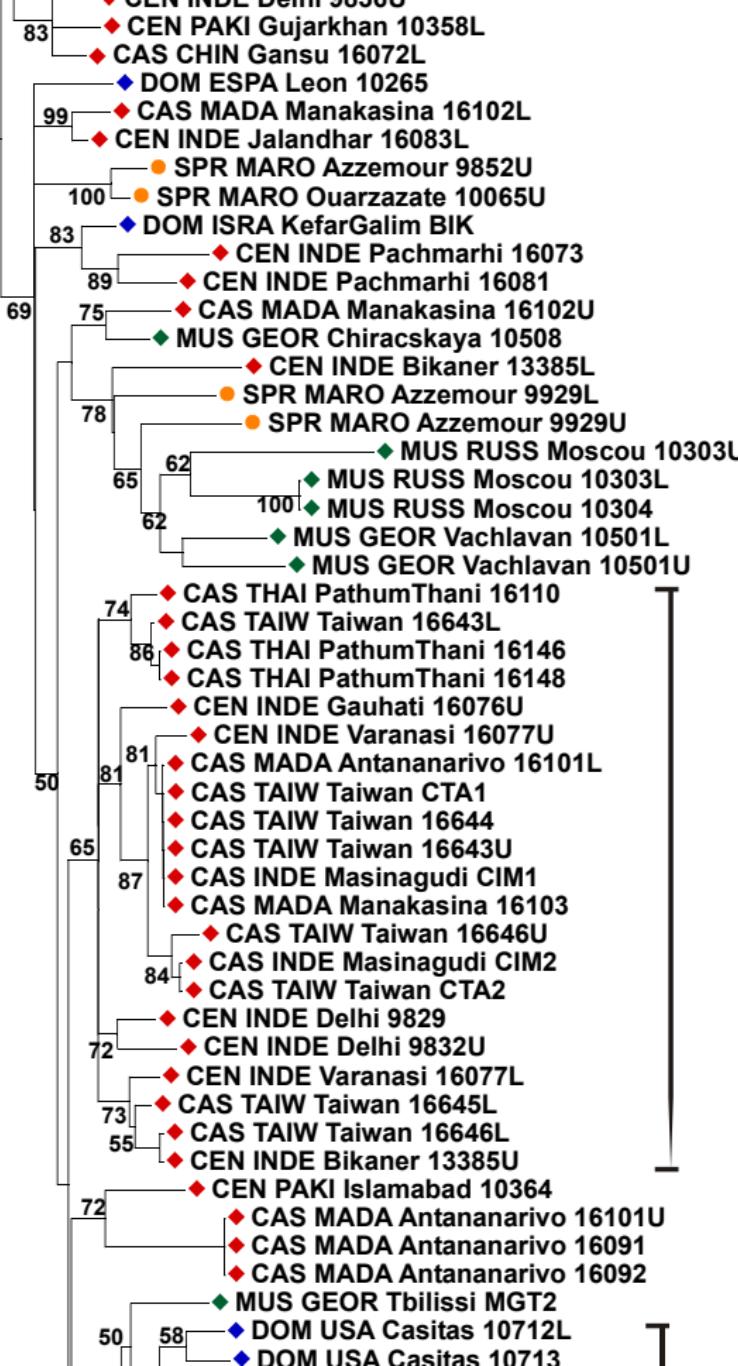
# MMS 30 coalescence



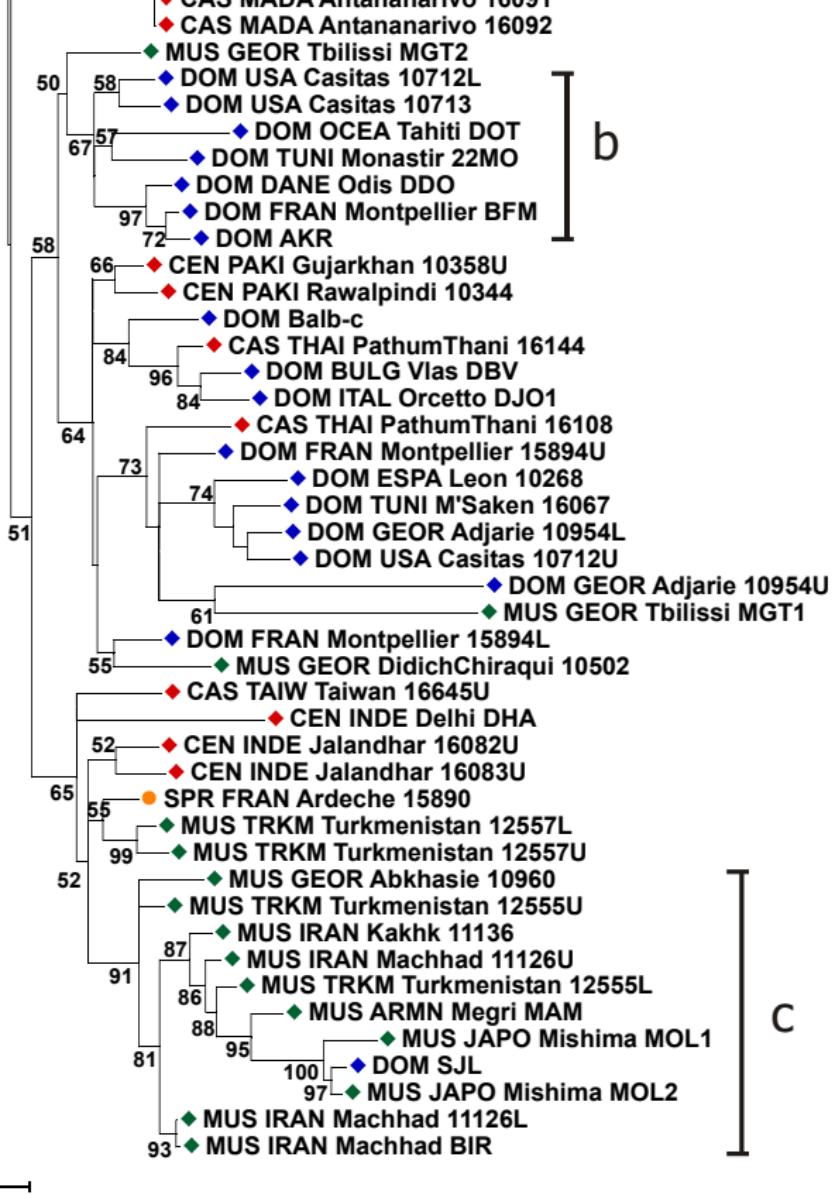
# MMS 24 coalescence



# MMS 24 coalescence

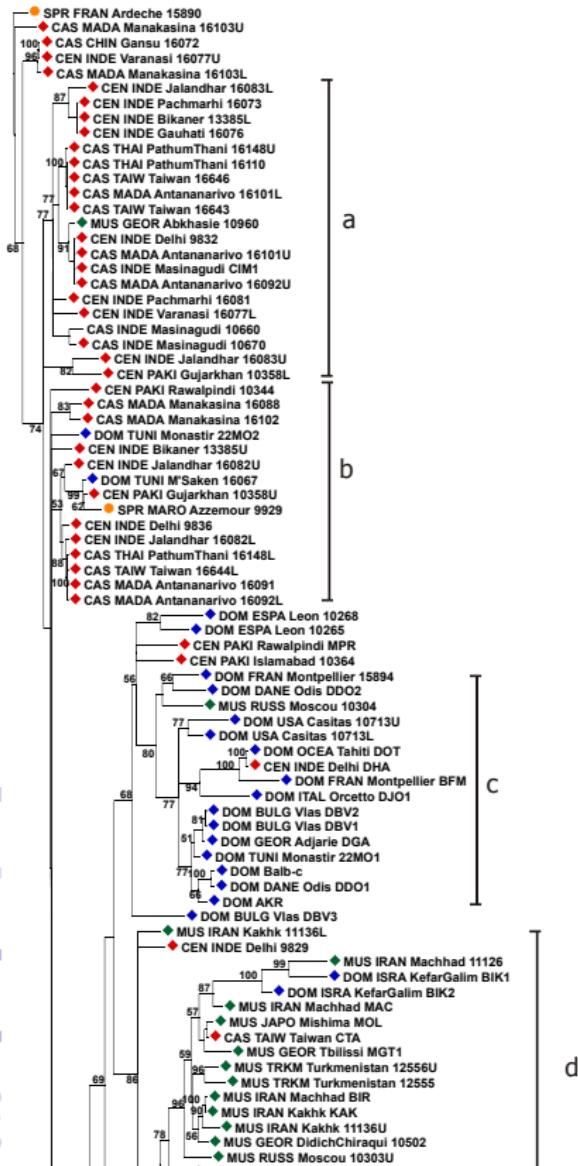


# MMS 24 coalescence

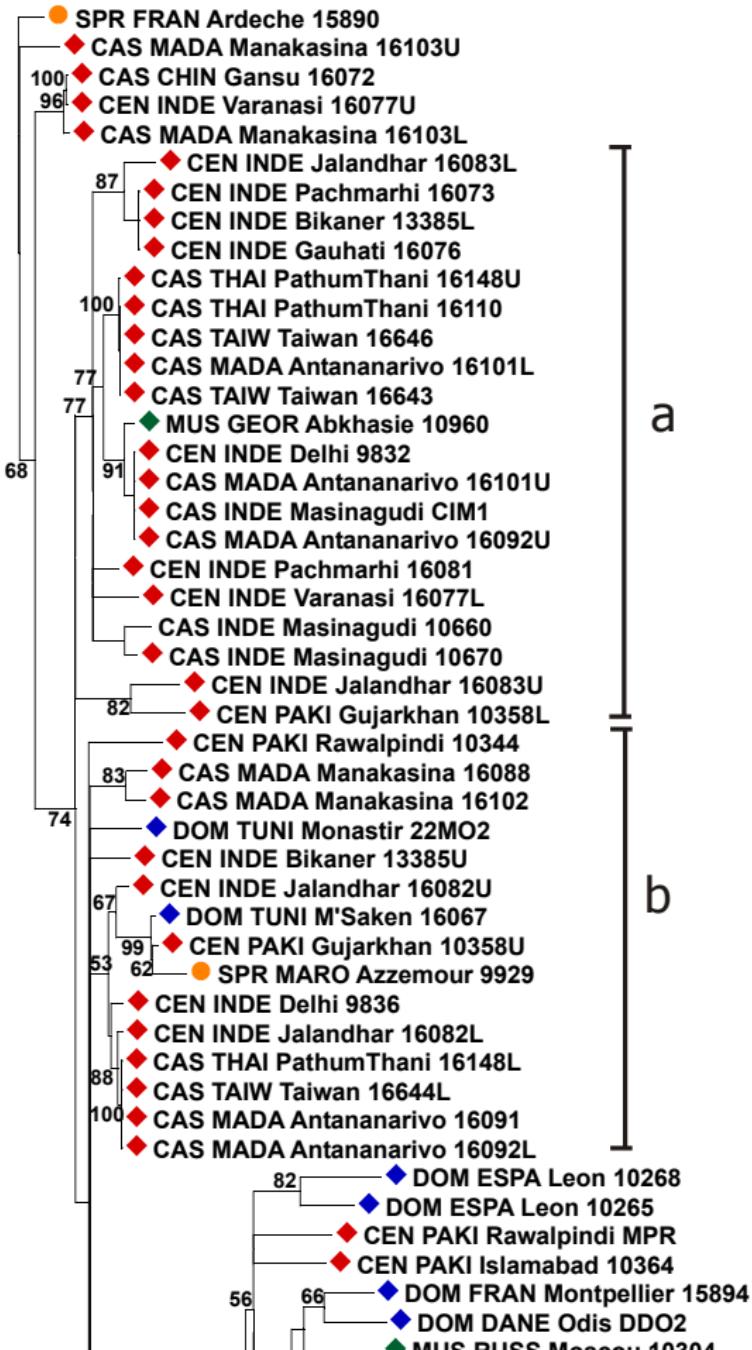


10

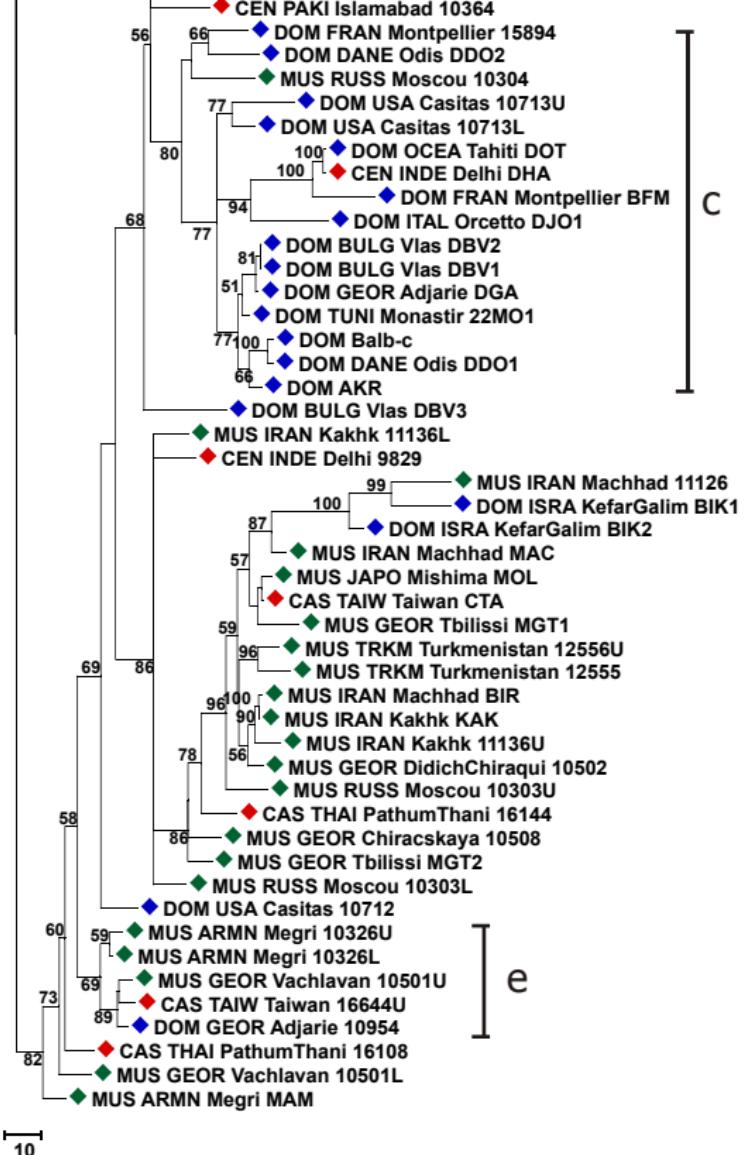
# MMS 26 coalescence



# MMS 26 coalescence

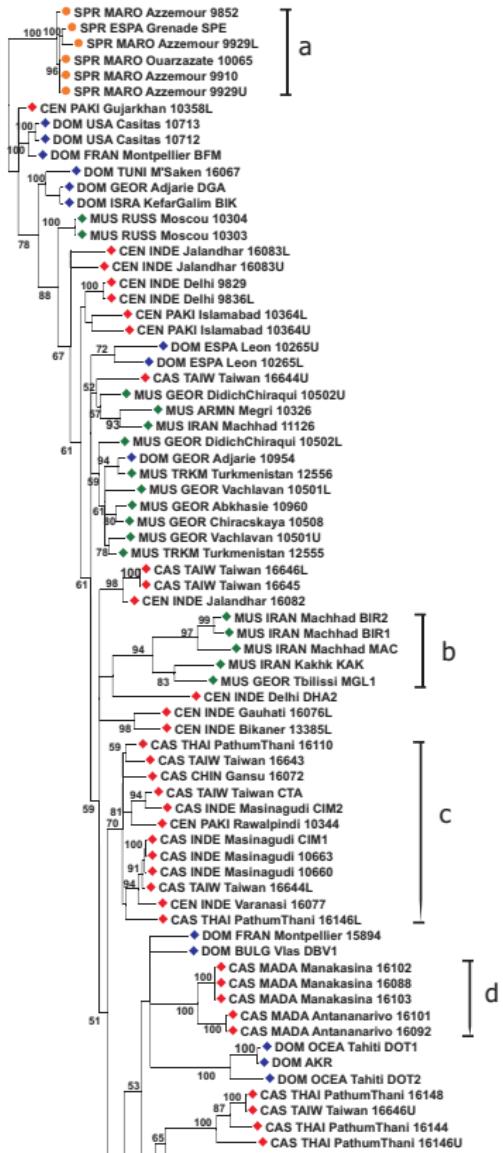


# MMS 26 coalescence

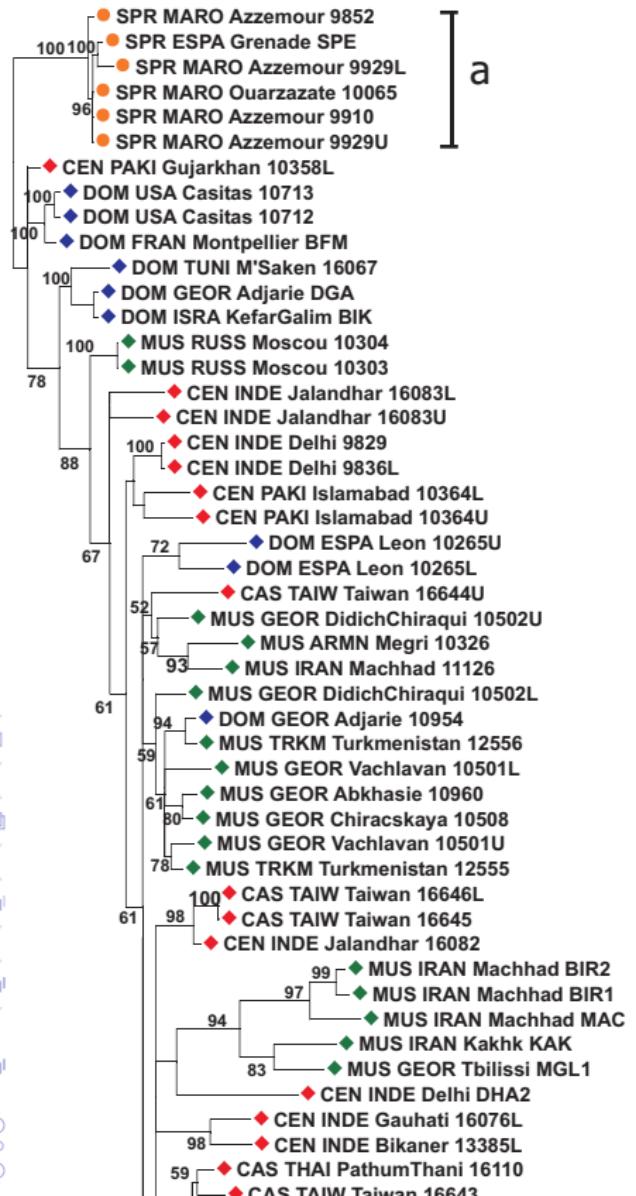


10

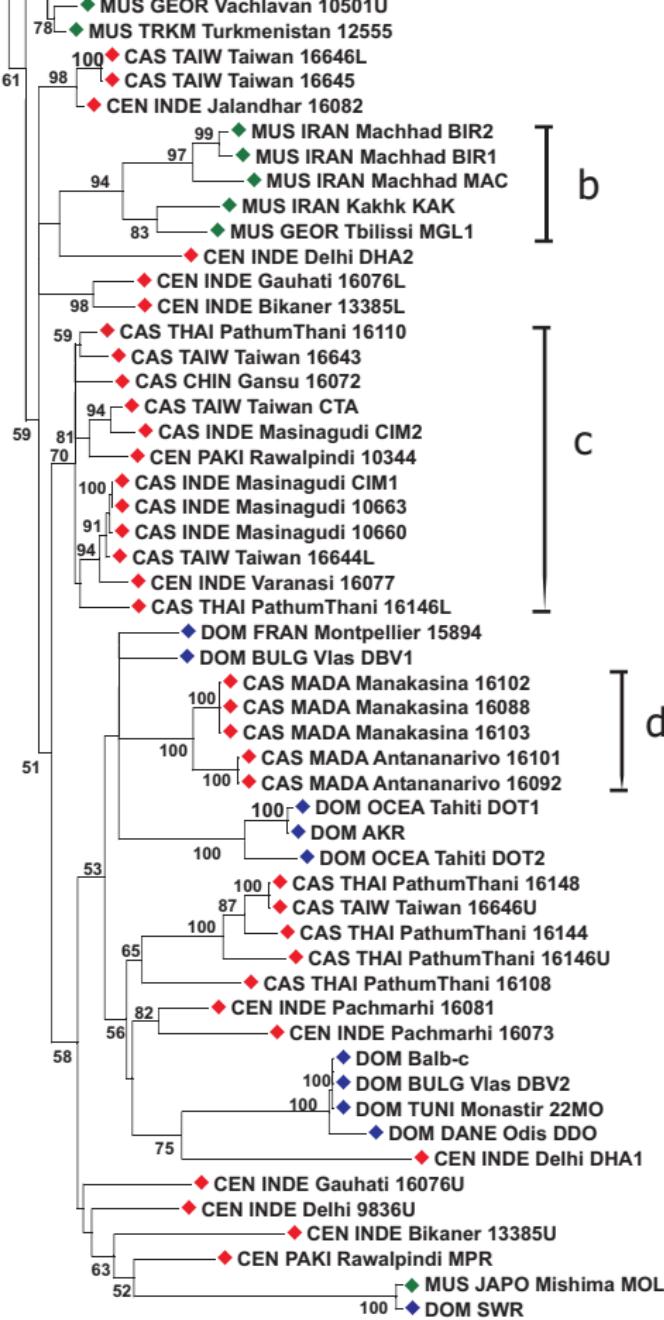
# MMS 80 coalescence



# MMS 80 coalescence



# MMS 80 coalescence



## Intruders

Locus ID			clade	Average distance to			
				CAS	CEN	DOM	MUS
24	CAS THAI	Pathumtani 16108	DOM	52	53	40	61
24	CAS THAI	Pathumtani 16144	DOM	45	47	37	60
24	DOM SJL		MUS	71	74	74	53
26	MUS GEOR	Abkhasie 10960	CAS	23	34	84	72
26	CEN INDE	Dehli DHA	DOM	99	96	58	91
26	MUS RUSS	Moscou 10304	DOM	69	62	60	92
26	CAS TAIW	Taiwan CTA	MUS	77	80	73	42
80	DOM GEOR	Adjarie 10954	MUS	37	39	56	28
80	DOM BULG	Vlas DBV1	CAS	40	50	50	51
80	DOM FRAN	Montpellier 15894	CAS	42	49	51	50

## MMS24 intruders alignments

DOM_ESPA_Leon	CTCCC~CCT~~oTCTTCoTC~oTCToT~~oTTCCCC~~
DOM_USA_Casit	CTCCo~CCo~oo~CTTCoTC~oTCTo~~ooTTCCCC~~
DOM_GEOR_Adja	CTCCT~CCTT~o~CTTCoTCCoTCTo~~ooTTCCCC~~
DOM_TUNI_M'Sa	CTCCT~CCT~~oTCTTCoTC~oTCTo~~ooTTC~~~~~
DOM_FRAN_Mont	CTCCT~CCT~~oTCTTC~~~~~oTCTT~~~oTTCCCC~~
CAS_THAI_Path	CTCCT~CCT~~oTCTTC~~~~~oTCoT~~~CT~~~~~
CAS_THAI_Path	CTCCT~CCTT~~TCTTCTTC~TTCTTT~~TTC~~~~~
DOM_BULG_Vlas	CTCCTTC~TT~~TCTTCTTC~TTCTTT~~TTCCCCCTC
DOM_ITAL_Orce	CTCCT~CCTT~~TCTTCTTC~TTCTTT~~oTTCCCCCT~
	**** *.. **** **. *..

## MMS30 intruders alignments

DOM ESPA	GYKKKGKLKYLWKGWGKoGGYWYKKoKKK~YYY~KG
DOM TURQ	GYKKKGKWKYoWKGWGKoGGYLYKKoKKK~YYY~KG
DOM TURQ	GYKKKGKWKYoWKGWGKoGGYLYKKoK~~~YYY~KG
DOM GEOR	GYKKKGKLKYoWKGWGKoGGYWYKKoK~~~YYY~KG
DOM ISRA	GYKKKGKLKYoWKGWGKoGGYWYKKoKKK~YYY~KG
DOM ESPA	GYKKKGKLKWoWKGWGKoGGYWYKKoKKK~YYY~KG
DOM FRAN	GYKKKGKLKYLWKGWGKoGGWGYWYKKoKKK~YY~~~G
DOM FRAN	GYKKKGKLKYoWKGWGKoGGYGGYYYKKoKKK~YY~~~G
DOM ITAL	GYKKKGKLKYoWKGWGKoGGYWYKKoKKK~YY~~~G
DOM TURQ	GYKKKGKLKYoWKGWGLWGGYWYKKoKKK~YYY~KG
CEN PAKI	GYKKKGKoKYKWKWGKoGGYWYKKoKKK~YYY~KG
CEN INDE	GYKKWGLWKYoKoWGLWGGYWYKKoKKKKYY~~KG
CAS MADA	GYKKK-GLLKYLoKGWGLoGGKWKKKoKKKKYY~~KG
CAS MADA	GYKKKGLLKYLoKGWGLoGGKWKKKoKKKKYY~~KG
	*****... .*.** .**.      **      **    .*

## MMS30 intruders alignments

DOM ESPA	GYKKK <b>WG</b> KLKYLWKGWGKoGGYWYKKoK <b>KK~YYY~KG</b>
DOM TURQ	GYKKK <b>WG</b> KWKYoWKGWGKoGGYLYKKoKK~YYY~KG
DOM TURQ	GYKKK <b>WG</b> KWKYoWKGWGKoGGYLYKKoK~~~YYY~KG
DOM GEOR	GYKKK <b>WG</b> KLKYoWKGWGKoGGYWYKKoK~~~YYY~KG
DOM ISRA	GYKKK <b>WG</b> KLKYoWKGWGKoGGYWYKKoKK~YYY~KG
DOM ESPA	GYKKK <b>WG</b> KLKWoWKGWGKoGGYWYKKoKK~YYY~KG
DOM FRAN	GYKKK <b>WG</b> KLKYLWKGWGK <b>WG</b> GYWYKKoKK~YY~~~G
DOM FRAN	GYKKK <b>WG</b> KLKYoWKGWGK <b>WG</b> GYYYYYKKoKK~YY~~~G
DOM ITAL	GYKKK <b>WG</b> KLKYoWKGWGKoGGYWYKKoKK~YY~~~G
DOM TURQ	GYKKK <b>WG</b> KLKYoWKGWGLWGGYWYKKoKK~YYY~KG
CEN PAKI	<b>GYKKK<b>WG</b></b> KoKYKWKWGKoGGYWYKKoK <b>KK~YYY~KG</b>
CEN INDE	GYKKK <b>WG</b> LWKYooKoWGLWGGYWYKKoKKKKYY~~KG
CAS MADA	GYKKK-GLLKYLoKGWGLoGGKWKKKoKKKKYY~~KG
CAS MADA	GYKKK <b>WG</b> LLKYLoKGWGLoGGKWKKKoKKKKYY~~KG
	*****... .*.** .**.      **      **      .*

## MMS30 intruders alignments

DOM ESPA	GYKKK <b>WG</b> KLKYLWKGWGKoGGYWYKKoKK~YYY~KG
DOM TURQ	GYKKK <b>WG</b> KWKYoWKGWGKoGGYLYKKoKK~YYY~KG
DOM TURQ	GYKKK <b>WG</b> KWKYoWKGWGKoGGYLYKKoK~~~YYY~KG
DOM GEOR	GYKKK <b>WG</b> KLKYoWKGWGKoGGYWYKKoK~~~YYY~KG
DOM ISRA	GYKKK <b>WG</b> KLKYoWKGWGKoGGYWYKKoKK~YYY~KG
DOM ESPA	GYKKK <b>WG</b> KLKWoWKGWGKoGGYWYKKoKK~YYY~KG
DOM FRAN	GYKKK <b>WG</b> KLKYLWKGWGKoGGYWGKoKK~YY~~~G
DOM FRAN	GYKKK <b>WG</b> KLKYoWKGWGKoGGYGGYYYYKKoKK~YY~~~G
DOM ITAL	GYKKK <b>WG</b> KLKYoWKGWGKoGGYWYKKoKK~YY~~~G
DOM TURQ	GYKKK <b>WG</b> KLKYoWKGWGLWGGYWYKKoKK~YYY~KG
CEN PAKI	<b>GYKKK<b>WG</b></b> KoKYWKWGKGKoGGYWYKKoKK~YYY~KG
CEN INDE	GYKKK <b>WG</b> LWKYooKoWGLWGGYWYKKoKKKKYY~~KG
CAS MADA	GYKKK-GLLKYLoKGWGLoGGKWKKKoKKKKYY~~KG
CAS MADA	GYKKK <b>WG</b> LLKYLoKGWGLoGGKWKKKoKKKKYY~~KG
	*****... .*.** .**.      **      **    .*

# Discussion

# Homoplasy vs genetic exchanges

- ▶ Are intruders or identical haplotypes due
  1. to homoplasy by convergent evolution
  2. or to genetic exchanges between subspecies?
- ▶ Some cases with short MS maps.
- ▶ Trees overall quality
- ▶ Most cases with long, complex maps
- ▶ Absence of long branches within subspecies specific clades

# Incomplete lineage sorting

- ▶ Intra subspecific coalescence depth versus divergence time
- ▶ Estimate from mitochondrial data [Prager et al., Genetics, 1998]  
DOM vs MUS : ratio = 0.27 < 1.
- ▶ Estimate for nuclear loci

DOM vs MUS	1.08	$\approx 1$
DOM vs CAS	1.75	$> 1$
MUS vs CAS	2.35	$> 1$
- ▶ Agreement with the MMS 30 tree on X chromosome and with trees of autosomal loci MMS 24, 26, 80

# Conclusion

# Conclusion

- ▶ MS analysis revealed:
  1. species-wide genetic flow
  2. past and present exchanges within the species range

# Conclusion

- ▶ MS analysis revealed:
  1. species-wide genetic flow
  2. past and present exchanges within the species range
- ▶ when handled with appropriate comparison algorithms  
[Bérard, Rivals, 2003]

# Conclusion

- ▶ MS analysis revealed:
  1. species-wide genetic flow
  2. past and present exchanges within the species range
- ▶ when handled with appropriate comparison algorithms  
[Bérard, Rivals, 2003]
- ▶ MS versus SNPs

# Conclusion

- ▶ MS analysis revealed:
  1. species-wide genetic flow
  2. past and present exchanges within the species range
- ▶ when handled with appropriate comparison algorithms  
[Bérard, Rivals, 2003]
- ▶ MS versus SNPs and micro-satellites

# Conclusion

- ▶ MS analysis revealed:
  1. species-wide genetic flow
  2. past and present exchanges within the species range
- ▶ when handled with appropriate comparison algorithms  
[Bérard, Rivals, 2003]
- ▶ MS versus **SNPs** and **micro-satellites**
- ▶ Mice genomes: set of interrelated gene pools, still able to exchange genes, **even on chromosome X**
- ▶ Difficult to find genes important for speciation

## Conclusion: Immunological Implications

Comment from J. Howard (Institut fur Genetik, RFA), F1000 Biology

*This paper establishes once and for all that there is continuous introgression of genetic material between the three subspecies of *Mus musculus**

*... This is by far the most extensive and accurate analysis of the *Mus musculus* species complex. It shows that none of the three sub-species is yet fully isolated.*

*... When immunologists consider the genetic aspects of immune responsiveness in the mouse, they should now be aware that the genomes they are comparing belong at least in part to three really quite well differentiated subspecies ...*

Faculty of 1000 Biology: for Bonhomme et al Genome Biol 2007

<http://www.f1000biology.com/article/id/1084935/evaluation>

## Algorithms' publications

- ▶ S. Bérard, E. Rivals, *Comparison of Minisatellites*, *J. of Computational Biology*, p. 357-372, vol. 10(3-4), 2003.
- ▶ S. Bérard, F. Nicolas, J. Buard, O. Gascuel, E. Rivals, *A Fast and Specific Alignment Method for Minisatellite Maps*, *Evolutionary Bioinformatics*, 2:327–344, 2006.

*MS\_Align* <http://atgc.lirmm.fr/>

## Credits and publication

- ISEM, Montpellier, France: F. Bonhomme, A. Orth
- Dpt Genetics, Univ. Leicester, GB: G. Grant, A.J. Jeffreys
- Scripps Research Institute, Jupiter, Florida: P. Bois
- **Support:** BioSTIC LR, Génopole LR,

ACI IMPBIO REPEVOL

<http://www.lirmm.fr/~rivals/RESEARCH/REPEVOL/>

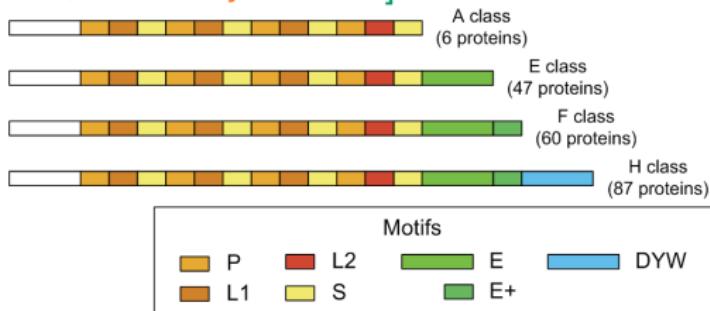
- F. Bonhomme, E. Rivals, A. Orth, G.R. Grant, A. J. Jeffreys, P.R.J. Bois

*Species wide distribution of highly polymorphic minisatellite markers suggests past and present genetic exchanges among house mouse subspecies*

*Genome Biology*, 8:R80, 2007.

# Autres applications *MS\_Align*

1. Micro-évolution du chromosome Y chez l'humain  
[Bérard et al., *Evol Bioinfo Online* 2006]
2. Formation de la famille protéines PPR chez les plantes  
[Rivals et al., *Plant Physiol* 2006]



3. Évolution d'un marqueur microsatellite composé chez les esturgeons