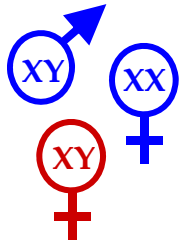
# Evolution of the *Sry* gene within the African pygmy mice *Nannomys*

- Subgenus of the genus *Mus*

- Widespread in Sub-Saharan Africa

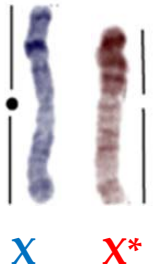
- ~ 20 species

- *Mus minutoides*



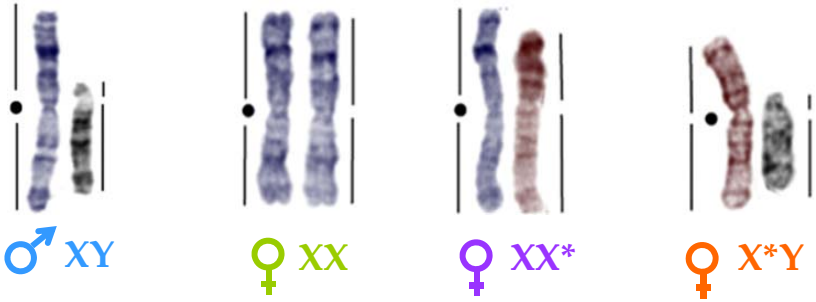

➡ Very high proportion (> 75%)
of fertile sex-reversed females

- **Sex reversal is due to a X-linked mutation**




➡ Two morphologically different X chromosomes,
one always associated to sex-reversed females (X*)

- **1 type of Males, 3 types of Females**



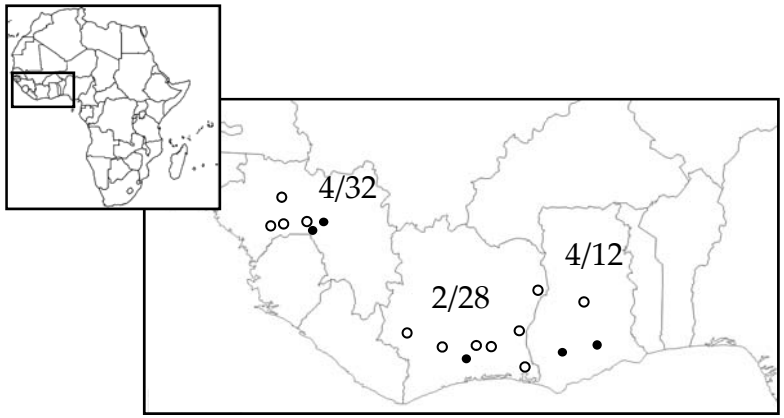

Veyrunes et al *Proc R Soc B* 2010

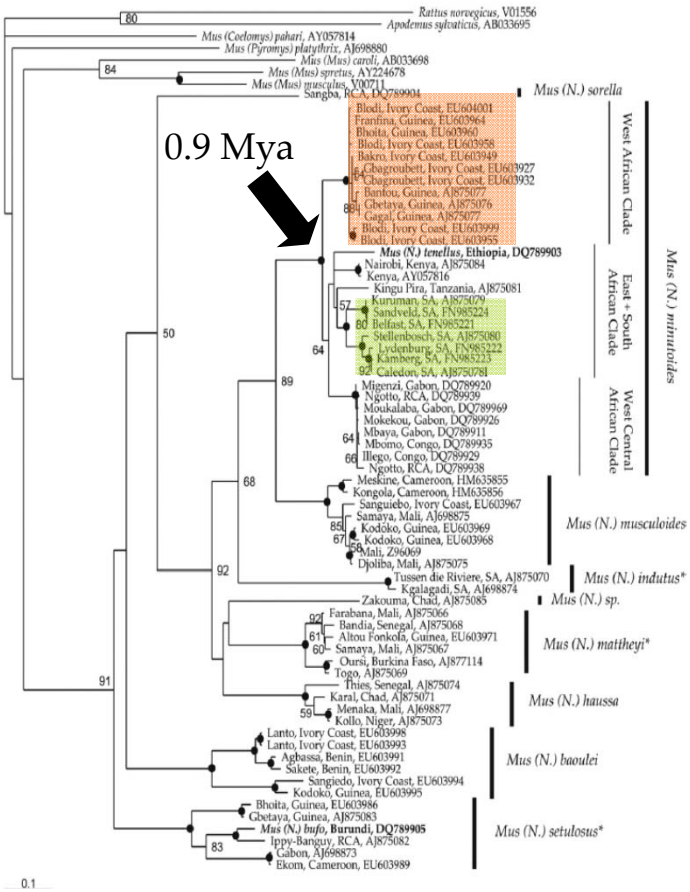
- X*Y females were only known from Southern Africa

→ Data spanning the distribution range were required to better assess the origin and the evolutionary history of the sex reversal mutation

- Karyotyping analyses require live animals, which is often a limitation.

→ Indirect approach (PCR of the *Sry* gene) performed on a large sample size (n = 72 females).



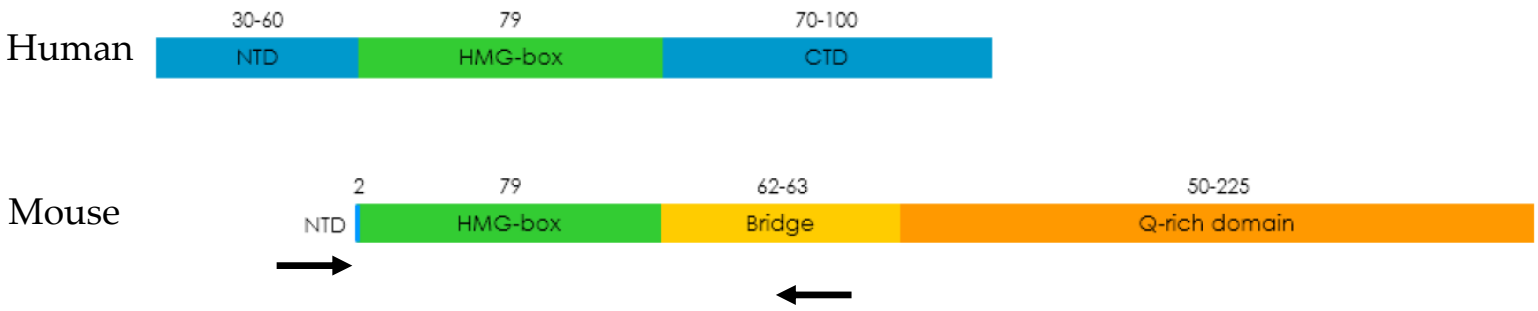

Veyrunes et al. 2013 *Sex Dev*

- **Sry** = Sex Determining Gene in Therian Mammals, Y-specific

→ single-exon gene in a single copy in most mammals, but multiple functional copies are present in several rodents

→ Only one extremely conserved region (**HMG-box**)
Most of the mutations associated with sex-reversal in humans and mice occur within this region

→ Structure not sequence of the CTD is important




Veyrunes et al. 2013 *Sex Dev*

## Cloning & Sequencing

Bantou: 5 haplotypes, but probably more copies

Mutanda M: 8 haplotypes

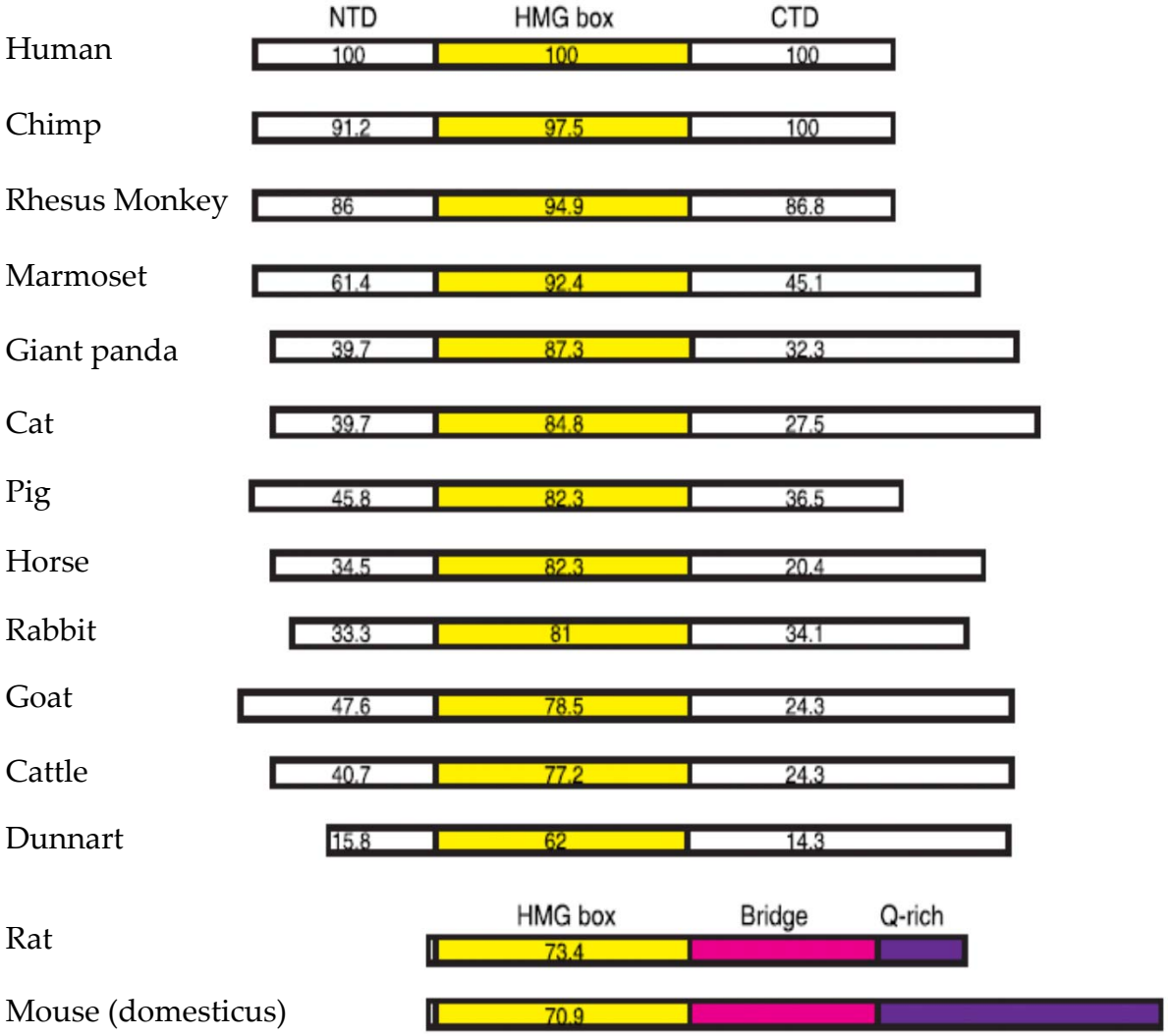
Mutanda F: 7 haplotypes

| | 302 233 2176 | | 2130 0881 0437 0027 002 | | 318 230 2131 1114 1088 0088 027 |
|---|---|---|---|---|---|
| **Bantou_F1** | ATRT | **Mutanda_M** | TRTCATS | **Mutanda_F1** | AMCAGTSG |
| clone_1 | G-A- | clone_1 | -G----G | clone_1 | -C--A-C- |
| clone_2 | --A- | clone_2 | -G-AGCG | clone_2 | GC---CG- |
| clone_3 | --G- | clone_3 | -A---CG | clone_3 | GC----C- |
| clone_4 | --G- | clone_4 | -A---CG | clone_4 | GA-G-CC- |
| clone_5 | -CA- | clone_5 | -G---CG | clone_5 | GCT---GA |
| clone_6 | --A- | clone_6 | -A----C | clone_6 | -C----C- |
| clone_7 | --A- | clone_7 | CG---CA | clone_7 | -C----C- |
| clone_8 | --G- | clone_8 | -A---CA | clone_8 | -C----C- |
| clone_9 | --A- | clone_9 | -GCAGCA | clone_9 | GA-G-CG- |
| clone_10 | --AC | clone_10 | -A----C | clone_10 | -C----C- |

→ Large number of *Sry* copies, and presence of polymorphism among these copies

→ Some of these mutations lead to AA substitutions, but almost all predicted proteins are presumed to be functional, except one.

Veyrunes et al. 2013 *Sex Dev*




## Within the genus

| | HMG box | Partial Bridge |
|---|---|---|
| *M. musculus* | 100 | 100 |
| *M. spretus* | 98.7 | 97.2 |
| *M. pahari* | 97.5 | 83.3 |
| *M. mattheyi* | 93.7 | 86.1 |
| *M. minutoides* | 91.1 | 86.1 |

## Within *M. minutoides*

| | | |
|---|---|---|
| Stellenbosch | 100 | 100 |
| Caledon | 100 | 100 |
| Kuruman | 98.7 | 100 |
| Mutanda | 94.9 | 100 |
| West Africa | 97.5 | 100 |

→ In the pygmy mice, and in *M. minutoides* in particular: very high proportion of Amino Acid replacement

→ More surprisingly, in *M. minutoides*, all these mutations occurred within the HMG !

Veyrunes et al. 2013 *Sex Dev*

→ This clearly indicates a rapid sequence evolution that has affected the African pygmy mouse SRY protein, and unexpectedly the HMG box in particular.

→ This is probably facilitated by the presence of multiple copies of the gene, but may also be related to the evolution of the novel sex determination system (**changes in sex chromosome transmission and selective forces)**

↓

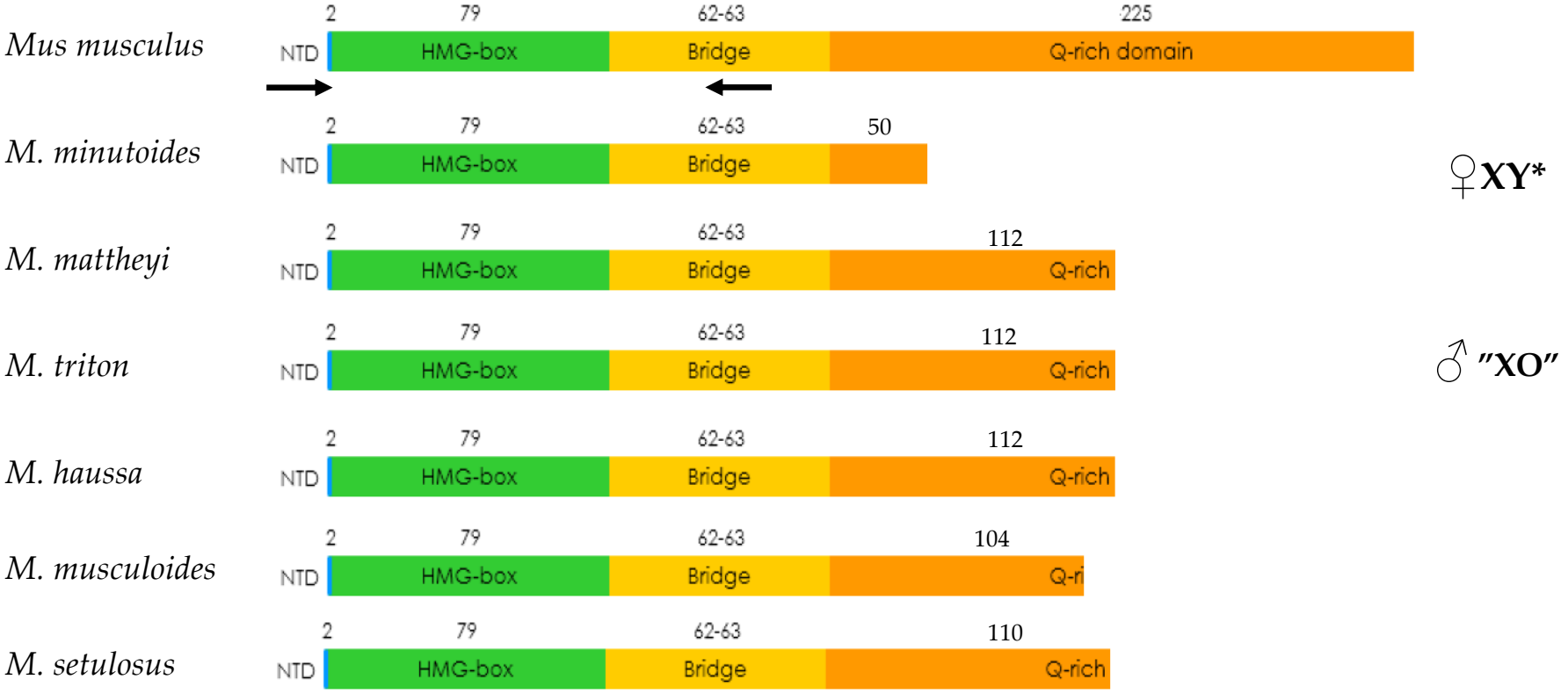
**Master Thesis of Andrés G. de la Filia Molina (Meme 2)**

- **Novel sequence analyses (trace of selection) involving the complete ORF and more species**

- **Link between rapid evolution of the gene and emergence of atypical sex determination systems ?**

  → *M. minutoides* (X*Y females) and *M. triton* ("XO" males)

| Species | HMG-box | Complete ORF | Single clones | 5'UF |
|---|---|---|---|---|
| *M. minutoides* | 23 | 22 | 0 | 19 |
| *M. mattheyi* | 12 | 9 | 23 (6) | 12 |
| *M. triton* | 16 | 1* | 19 (9) | 10 |
| *M. haussa* | 3 | 1 | 4 (1) | 3 |
| *M. musculoides* | 5 | 1 | 1 (1) | 4 |
| *M. setulosus* | 1 | 1 | 1 (1) | 1 |
| *M. gerbillus* | 1 | 1 | 1 (1) | 1 |
| *M. baoulei* | 2 | 0 | 0 | 2 |
| *M. indutus* | 1 | 0 | 0 | 1 |



## Sequence comparison

| HMG-box | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 *M. minutoides* | – | 96.45 | 82.96 | 96.32 | 96.32 | 93.02 | 98.81 | 98.12 | 96.32 | 91.01 | 92.37 | 91.01 | 90.89 | 85.60 |
| 2 *M. mattheyi* | 96.96 | – | 83.86 | 99.58 | 99.58 | 95.67 | 97.49 | 96.21 | 99.58 | 93.39 | 94.70 | 93.39 | 93.30 | 88.07 |
| 3 *M. triton* | 92.03 | 92.25 | – | 83.99 | 83.99 | 88.12 | 83.89 | 82.44 | 83.99 | 85.69 | 86.48 | 85.63 | 85.41 | 82.86 |
| 4 *M. musculoides* | 97.83 | 98.65 | 92.92 | – | 100.00 | 95.80 | 97.50 | 96.23 | 100.00 | 93.63 | 94.94 | 93.63 | 93.55 | 88.22 |
| 5 *M. haussa* | 97.09 | 99.42 | 92.53 | 99.09 | – | 95.80 | 97.50 | 96.23 | 100.00 | 93.63 | 94.94 | 93.63 | 93.55 | 88.22 |
| 6 *M. setulosus* | 96.14 | 97.02 | 93.69 | 97.79 | 97.45 | – | 94.36 | 92.90 | 95.80 | 98.62 | 100.00 | 98.62 | 98.60 | 92.90 |
| 7 *M. baoulei* | 99.25 | 96.98 | 91.96 | 97.86 | 97.12 | 96.17 | – | 98.76 | 97.50 | 92.30 | 93.63 | 92.30 | 92.20 | 86.82 |
| 8 *M. indutus* | 99.34 | 96.94 | 92.02 | 97.86 | 97.12 | 96.17 | 99.18 | – | 96.23 | 90.96 | 92.30 | 90.96 | 90.84 | 85.40 |
| 9 *M. gerbillus* | 97.91 | 98.73 | 93.01 | 99.92 | 99.18 | 97.87 | 97.94 | 97.94 | – | 93.63 | 94.94 | 93.63 | 93.55 | 88.22 |
| 10 *M. m. musculus* | 95.02 | 95.85 | 92.76 | 96.63 | 96.30 | 97.87 | 95.06 | 95.06 | 96.71 | – | 98.76 | 97.50 | 97.47 | 92.30 |
| 11 *M. m. domesticus* | 95.43 | 96.26 | 93.19 | 97.04 | 96.71 | 98.30 | 95.47 | 95.47 | 97.12 | 99.59 | – | 98.76 | 98.74 | 93.63 |
| 12 *M. pahari* | 95.02 | 95.09 | 93.02 | 95.80 | 95.47 | 97.02 | 95.06 | 95.06 | 95.88 | 97.53 | 97.94 | – | 100.00 | 92.30 |
| 13 *M. crociduriodes* | 95.83 | 95.90 | 93.04 | 96.61 | 96.28 | 97.86 | 95.87 | 95.87 | 96.69 | 98.35 | 98.76 | 99.17 | – | 92.20 |
| 14 *R. norvegicus* | 89.68 | 90.80 | 89.74 | 91.28 | 90.95 | 92.34 | 89.71 | 89.71 | 91.36 | 93.00 | 93.42 | 93.00 | 92.98 | – |

| CTD | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|---|---|---|---|
| 15 *M. minutoides* | – | 69.37 | 63.64 | 66.67 | 71.17 | 67.31 | 66.67 | 61.26 | 60.91 | 59.30 |
| 16 *M. mattheyi* | 87.34 | – | 79.43 | 82.18 | 95.71 | 84.13 | 82.69 | 70.06 | 73.76 | 75.00 |
| 17 *M. triton* | 81.50 | 90.32 | – | 79.19 | 81.12 | 83.20 | 74.66 | 71.81 | 69.78 | 72.41 |
| 18 *M. musculoides* | 87.19 | 88.64 | 88.57 | – | 86.31 | 84.92 | 76.92 | 66.67 | 70.67 | 75.00 |
| 19 *M. haussa* | 87.00 | 97.84 | 90.04 | 92.49 | – | 84.92 | 82.39 | 72.51 | 72.11 | 76.14 |
| 20 *M. setulosus* | 85.45 | 93.19 | 91.01 | 93.35 | 93.40 | – | 81.60 | 74.60 | 72.22 | 74.71 |
| 21 *M. gerbillus* | 87.00 | 89.86 | 89.24 | 91.97 | 90.85 | 92.08 | – | 68.64 | 65.54 | 76.14 |
| 22 *M. m. musculus* | 79.06 | 84.17 | 83.84 | 83.50 | 86.10 | 86.84 | 85.74 | – | 84.67 | 71.59 |
| 23 *M. m. domesticus* | 81.11 | 86.81 | 83.90 | 84.95 | 87.92 | 87.80 | 85.82 | 95.64 | – | 69.32 |
| 24 *R. norvegicus* | 79.17 | 85.77 | 83.90 | 86.14 | 85.77 | 85.71 | 86.52 | 81.27 | 82.02 | – |

**TABLE 3.** Pairwise nucleotide (in black, below the diagonal) and predicted amino acid sequences (in blue, above the diagonal) identities between murid species. . Distances for the HMG-box and CTD regions were calculated separately.

## Multiple polymorphic *Sry* copies in *M. mattheyi* and *M. triton,* <u>but not *M. haussa*</u>

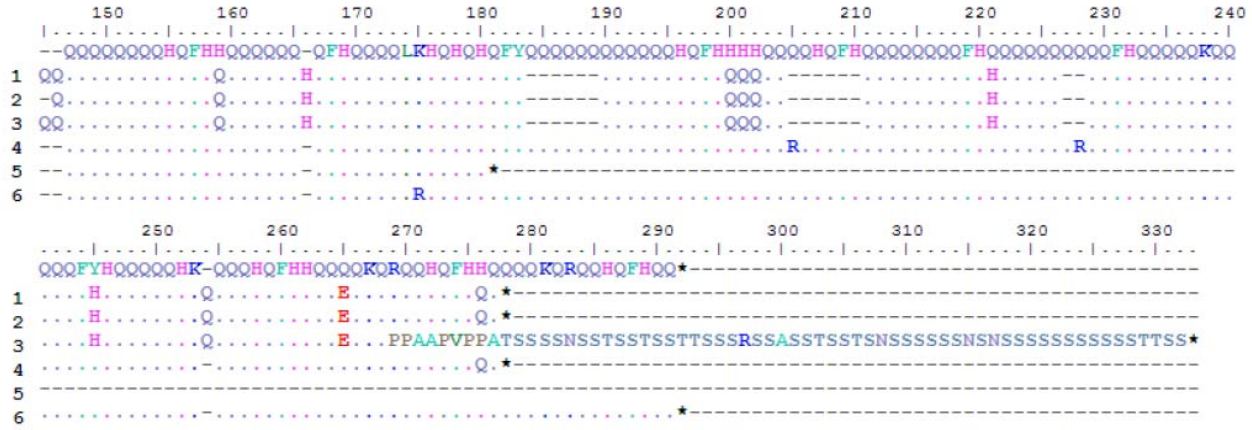
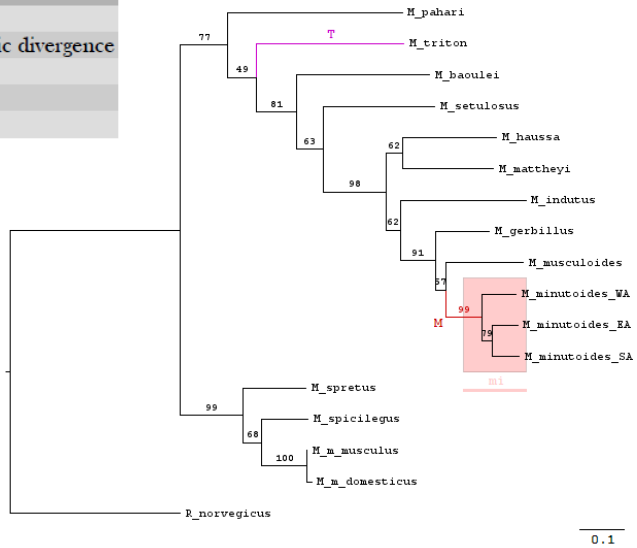


**FIGURE 5.** *(Next page, right)* Alignment of Q- rich regions of 6 *Sry* clones from a *M. mattheyi* individual, *mmat256.* The first sequence was obtained by direct sequencing of PCR product amplified by *SRY* primers. HMG-box and bridge are identical in all clones (not shown).

## dN/dS ratios and estimation of selection pressure in the *Nannomys* phylogeny (PAML)

| Model | Assumptions | Hypothesis |
|---|---|---|
| A: One ratio | $\omega_0 = \omega_M = \omega_{mi} = \omega_T$ | Selection pressure is the same across all lineages in the phylogeny |
| B: Two ratios | $\omega_0 = \omega_{mi} = \omega_T \neq \omega_M$ | Episodic change in selection pressure in *M. minutoides* prior to intraspecific divergence |
| C: Two ratios | $\omega_0 = \omega_T \neq \omega_M = \omega_{mi}$ | Permanent change in selection pressure in *M. minutoides* |
| D: Two ratios | $\omega_0 = \omega_M = \omega_{mi} \neq \omega_T$ | Change in selection pressure in *M. triton* |
| E: Free ratio | $\omega_1 \neq \omega_i \neq \ldots \neq \omega_n$ | Selection pressure varies across all lineages in the phylogeny |

| Region | Model | n. p. | $-\ell$ | $\omega$ | LRT | d.f. | $2\Delta\ell$ | P |
|---|---|---|---|---|---|---|---|---|
| ORF | A | 19 | 1651.6091 | $\omega_0 = 0.856$ | - | - | - | - |
| | B | 20 | 1651.1608 | $\omega_0 = 0.765$, $\omega_M = 1.127$ | B vs A | 1 | 0.897 | NS |
| | C | 20 | 1651.2487 | $\omega_0 = 0.762$, $\omega_M = \omega_{mi} = 1.056$ | C vs A | 1 | 0.721 | NS |
| | D | 20 | 1651.1764 | $\omega_0 = 0.807$, $\omega_T = 1.455$ | D vs A | 1 | 0.865 | NS |
| | E | 35 | 1639.2229 | See FIGURE 7 | E vs A | 16 | 24.772 | NS |
| CTD | A | 19 | 1120.0575 | $\omega_0 = 1.243$ | - | - | - | - |
| | B | 20 | 1119.9850 | $\omega_0 = 1.178$, $\omega_M = 1.457$ | B vs A | 1 | 0.145 | NS |
| | C | 20 | 1120.0575 | $\omega_0 = 1.243$, $\omega_M = \omega_{mi} = 1.245$ | C vs A | 1 | 0.000 | NS |
| | D | 20 | 1119.1934 | $\omega_0 = 1.130$, $\omega_T = 3.604$ | D vs A | 1 | 1.728 | NS |
| | E | 35 | 1109.0058 | Not shown | E vs A | 16 | 22.103 | NS |
| HMG-box | A | 33 | 872.0413 | $\omega_0 = 0.421$ | - | - | - | - |
| | B | 34 | 871.9971 | $\omega_0 = 0.413$, $\omega_M = 0.549$ | B vs A | 1 | 0.088 | NS |
| | C | 34 | 871.5068 | $\omega_0 = 0.386$, $\omega_M = \omega_{mi} = 0.936$ | C vs A | 1 | 1.070 | NS |
| | D | 34 | 870.2745 | $\omega_0 = 0.311$, $\omega_T = 0.922$ | D vs A | 1 | 3.534 | NS |
| | E | 63 | 858.2532 | Not shown | E vs A | 30 | 27.576 | NS |



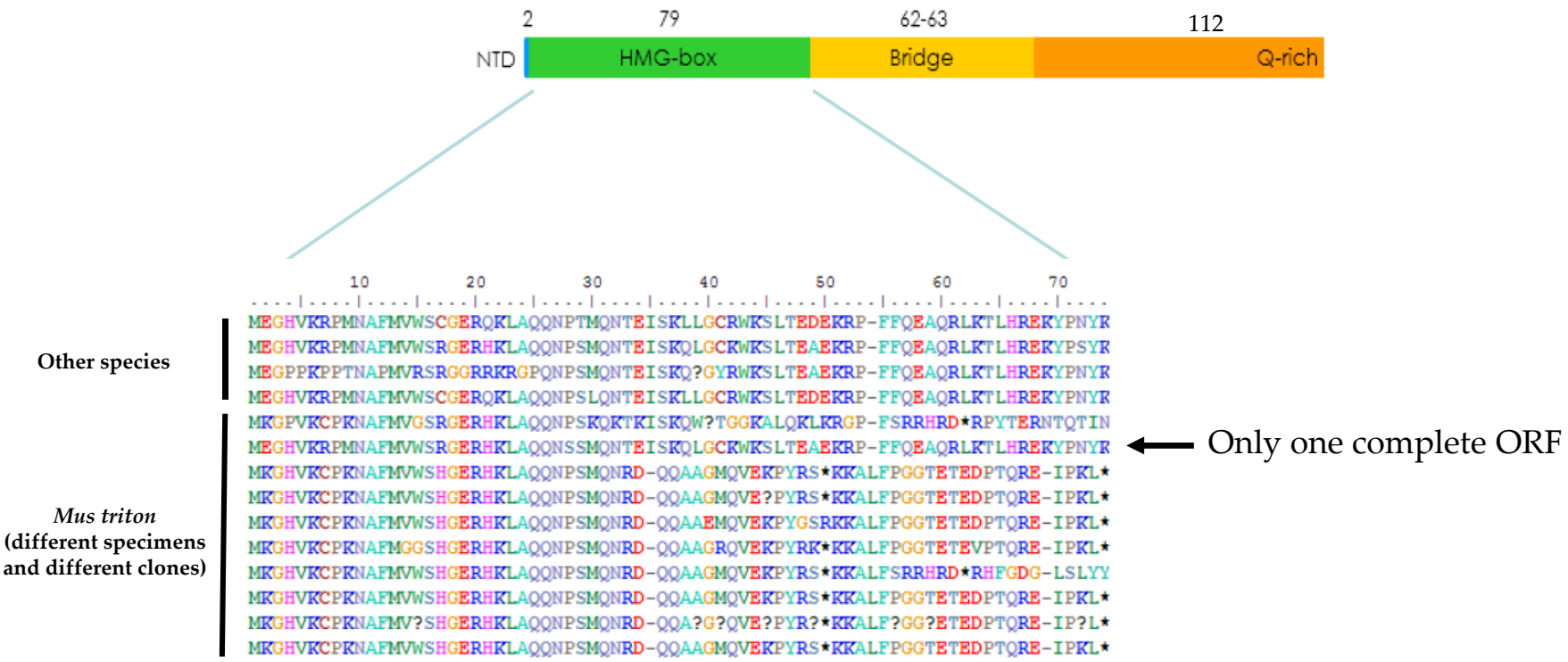

→ No significative differences, best model is the null one-ratio which assumes an identical $\acute{\omega}$ for the whole phylogy (Similar results as in Felidae and in Cebidae)

→ Purifying selection acts upon HMG-box, while CTD (particularly Q-rich domain) evolves more freely (relaxation of selective pressure?)

→ CTD evolves faster, even if this region is indispensable for SRY function

→ Regarding *M. minutoides* and *triton*: hints of increased of dN/dS ratio suggest ralaxation of selection or adaptive evolution (cause or consequence to novel SDS ?)

- **ORF of *Mus triton* ("XO" males)**



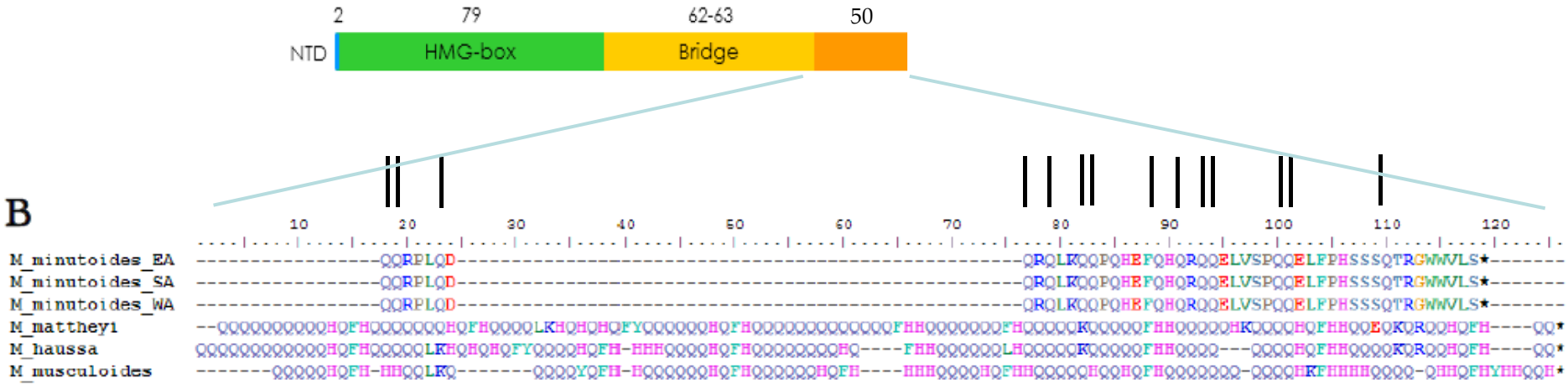

→ At least there is some Y-specific material.

→ All analyzed copies, except one, were pseudogenes.

→ For 3 males, we were unabled to find a single functional copy. Since one copy is enough, we need to do more cloning work

→ Systematic revision (Bryja et al., 2014)

- **ORF of *Mus minutoides* (X*Y females)**

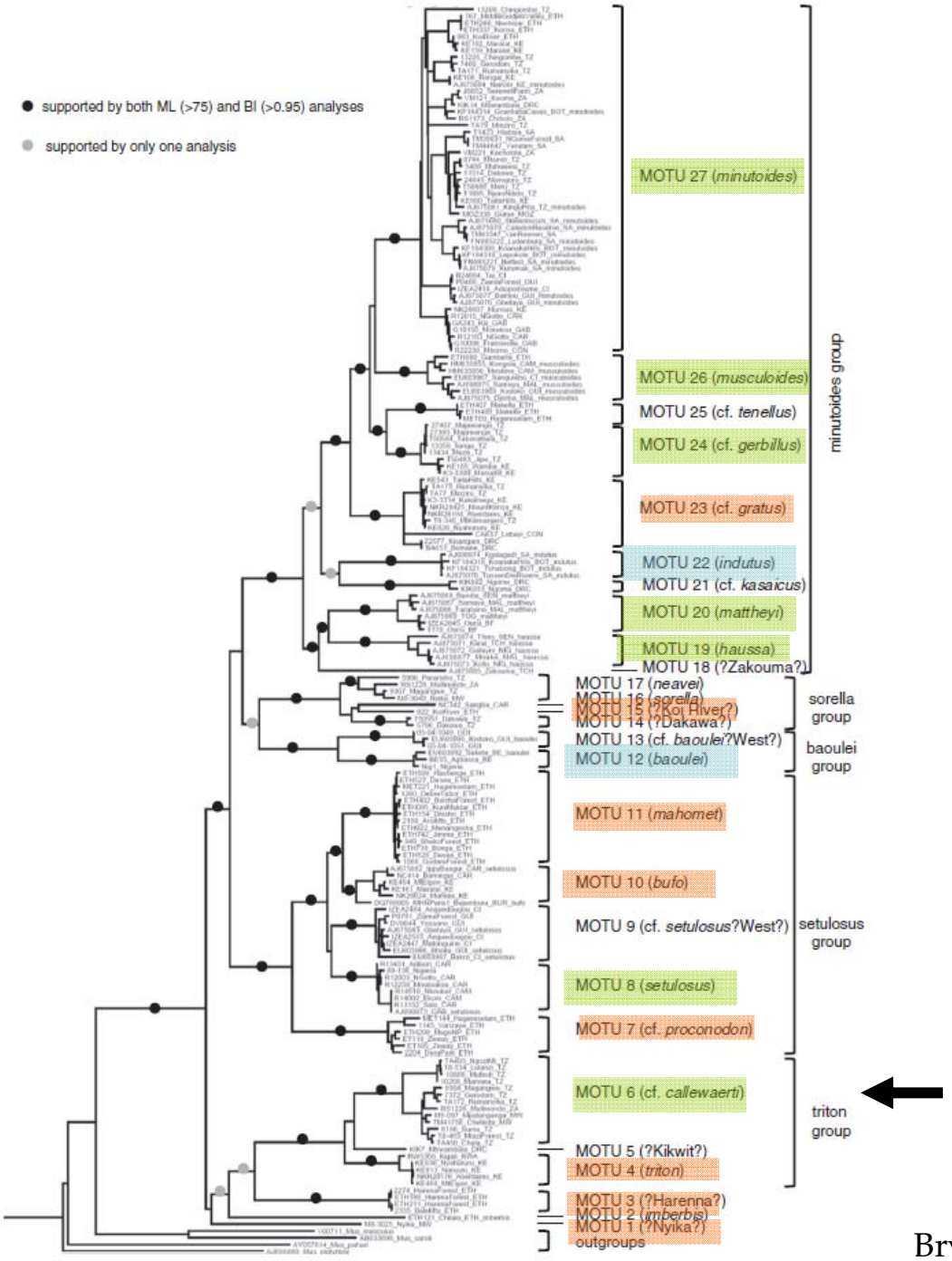



→ 14 Q residues, in contrast with 60-80 in the other pygmy mice, >100 in *M. domesticus*

→ *minutoides* CTD does not carry the number of 3 poly-Q blocks that was defined by Chen et al. 2013 (*PNAS*) as the minimal requirement to trigger male sex determination in rodents.

→ Poly-Q blocks are essential for stabilising SRY protein and transcriptionally activating TESCO (Zhao et al., 2014 *PNAS*)

→ **Hyp**: changes in CTD have intensively weakened *Sry* activity in *M. minutoides*, which would be enough to determine maleness in normal conditions but not in presence of a feminizing gene on the X*




Bryja et al., 2014 *BMC Evol Biol*

# Reduced activity of SRY and its target enhancer *Sox9*-TESCO in a mouse species with X chromosome-dependent XY sex reversal

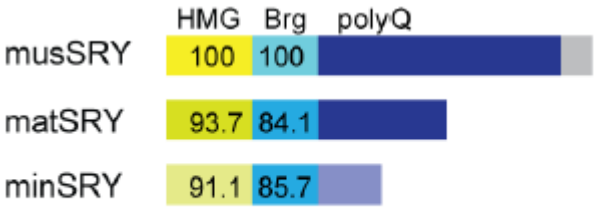
Liang Zhao, Ee Ting Ng, Frederic Veyrunes and Peter Koopman

In revision for *Genome Research*

- **Structure and activity of SRY and its target, the testis-specific enhancer element TESCO of *Sox9***
- **Lack of poly-Q track that is essential to male sex determination**

→ Stability and Transactivation activity of *M. minutoides* Sry

→ Sequence variation of TESCO



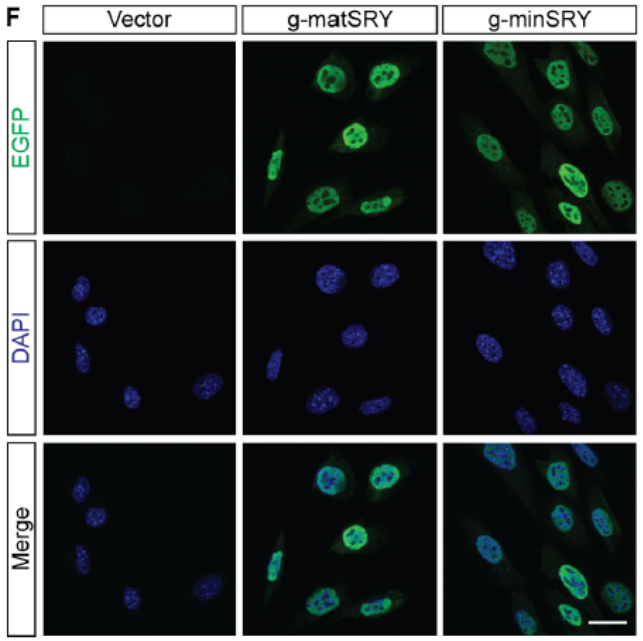

```
                  150       160       170       180       190       200       210
Consensus         QQQQQQQQQXQFHXHXQQQQ--XXXHQQQQLKXQXQXQFXXXQQQQHQFHXQXQQQQ------XXXQQQQFH
M. musculus       QQQQQQQQQQQFHNHHQQQQQFYDHHQQQQQQQQQQQQFHDHHQQKQQFHDHHQQQQQFHDHHHHHQEQQFH
M. mattheyi       QQQQQQQQQHQFHHHQQQQQ--HQFHQQQQLKHQHQHQFYQQQQQQHQFHQQQQQQ------QQQQQQQFH
M. minutoides     QQRLLQ--------------------DQRQLK----------QQLQHEFQHQRQQEL------VSPQQELFP

                  220       230       240       250       260       270       280
Consensus         XXXQQQQ---XHQQQQQXQQQQQFHXXHQQX-------HXQQQXQFHXXQQQX------QRQQHQFHXX---
M. musculus       DHHQQQQQFHDHQQQQQQQQQQQFHDHHQQKQQFHDHHHHQQQQQFHDHQQQQQQFHDHQQQQQHQFHDHPQQ
M. mattheyi       QQQQQQQ---FHQQQQQKQQQQQFHHQHQQQ-------HKQQQHQFHHQQQQK------QRQQHQFHQQ
M. minutoides     HS--S----------SQ-----------------------------------------TRGWWVFS

                  290       300       310       320       330       340       350       360
Consensus         ------------------------------------------------------------------------
M. musculus       KQQFHDHPQQQQQFHDHHHQQQQKQQFHDHHQQKQQFHDHHQQKQQFHDHHQQQQQFHDHHQQQQQQQQQQQ
M. mattheyi
M. minutoides

                            370       380       390
Consensus         ----------------------------------
M. musculus       QQFHDQQLTYLLTADITGEHTPYQEHLSTALWLAVS
M. mattheyi
M. minutoides
```

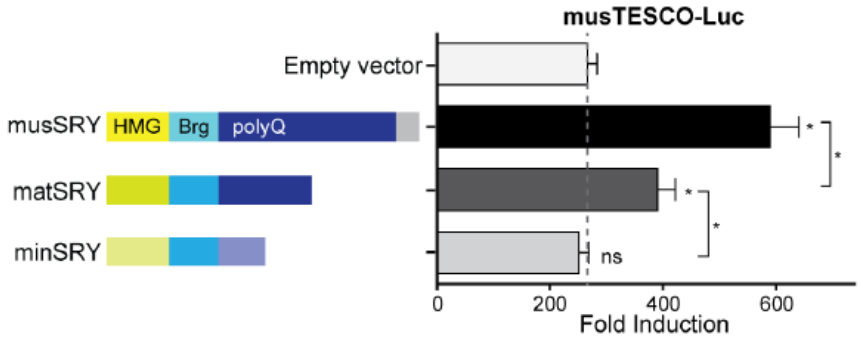
- **Stability?**


Immunofluorescence

→ Ok, not affected by its degraded polyQ tract

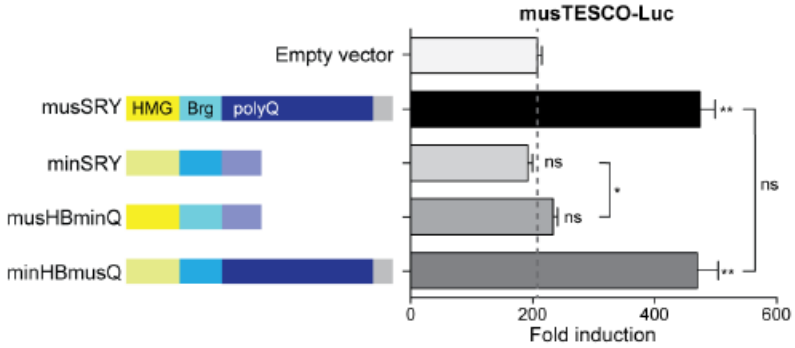
- **Transactivation? (Luciferase reporter Assay)**




→ **MinSRY failed to activate Mus TESCO**

*Due to degraded polyQ domain that lost transactivation potential?*
*or to variations in the HMG/bridge domains that impaired binding to TESCO?*

↓ **Two new mutant constructs**

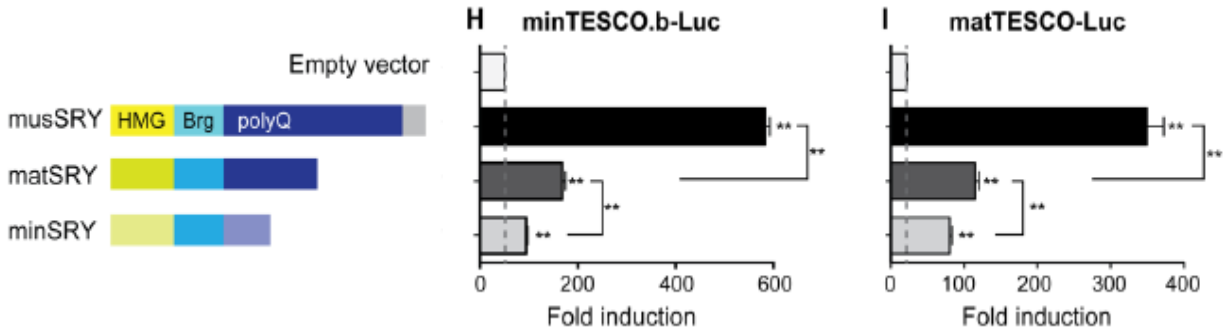



→ **Dicreased transactivation activity of MinSRY is due to the loss of a typical polyQ tract**

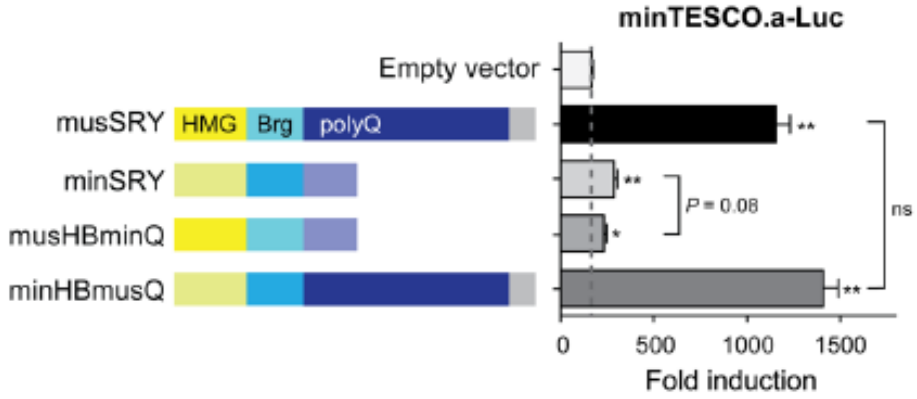
- **TESCO sequence of *M. minutoides*?**

→ 7% of divergence with Mus TESCO

→ TESCO and SRY in *M. minutoides* may have co-evolved such that Min SRY remains able to activate its cognate TESCO




→ Functional compatibilities: MinSRY activates Min TESCO (albeit more weakly than Mat/Mus SRY with longer polyQ)




→ SRY capacity to activate TESCO relies primarily on its polyQ tract

## Conclusions

- Ability of Min SRY to activate transcription of TESCO is diminished due to degeneration of polyQ tract
- TESCO shows impaired basal transcriptional activity
- These differences together rendered the male sex determining pathway vulnerable to the invasion of a sex reversal mutation