Approche génomique de l'impact des déversements de truites (*Salmo trutta*) domestiques dans les populations sauvages d'origine méditerranéenne

Maëva Leitwein

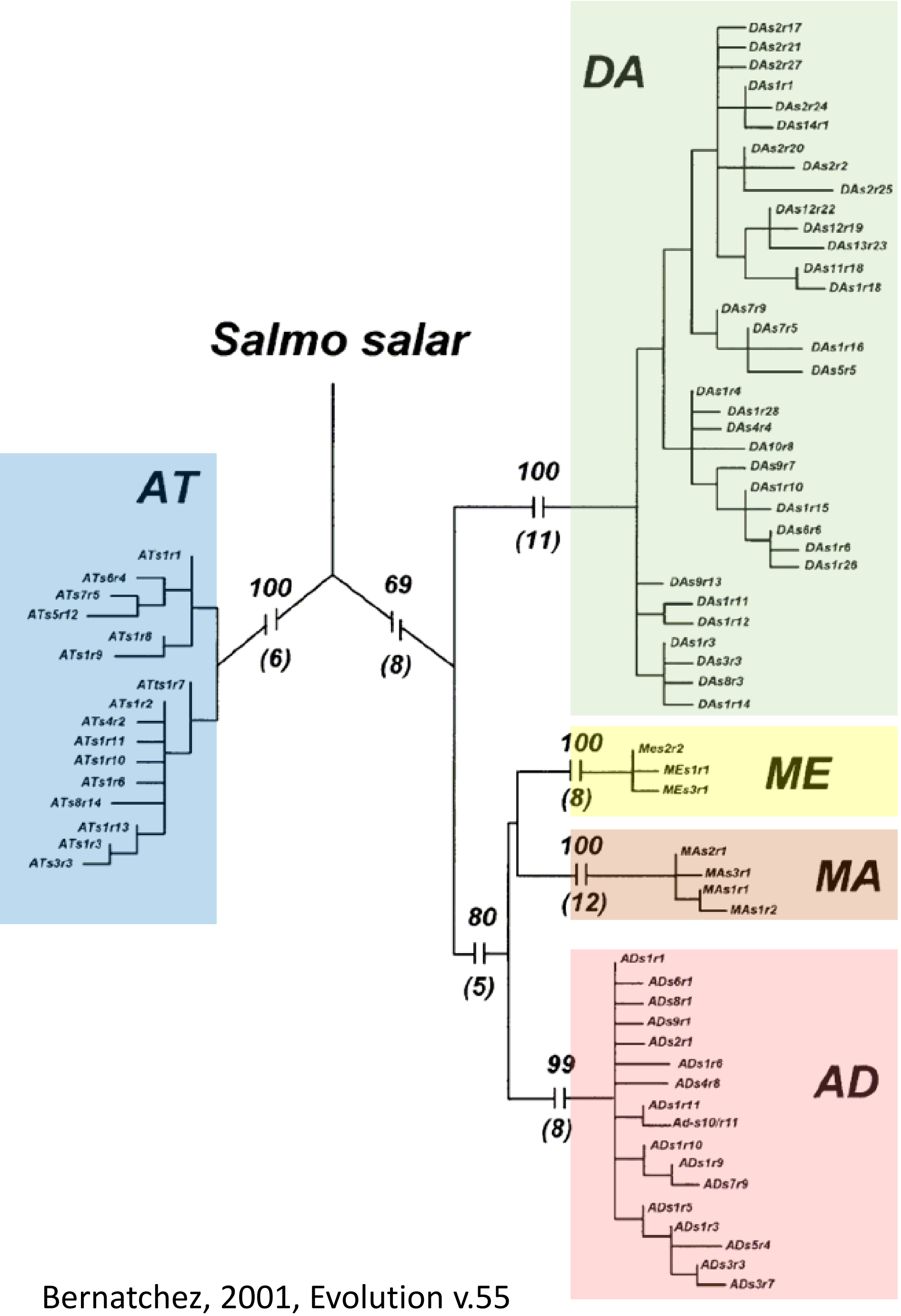
isem
Institut des Sciences de l'Evolution-Montpellier

CeMEB

Université Montpellier

34
FÉDÉRATION
DÉPARTEMENTALE
PÊCHE

Séminaire de l'UMR CBGP 13/12/2016

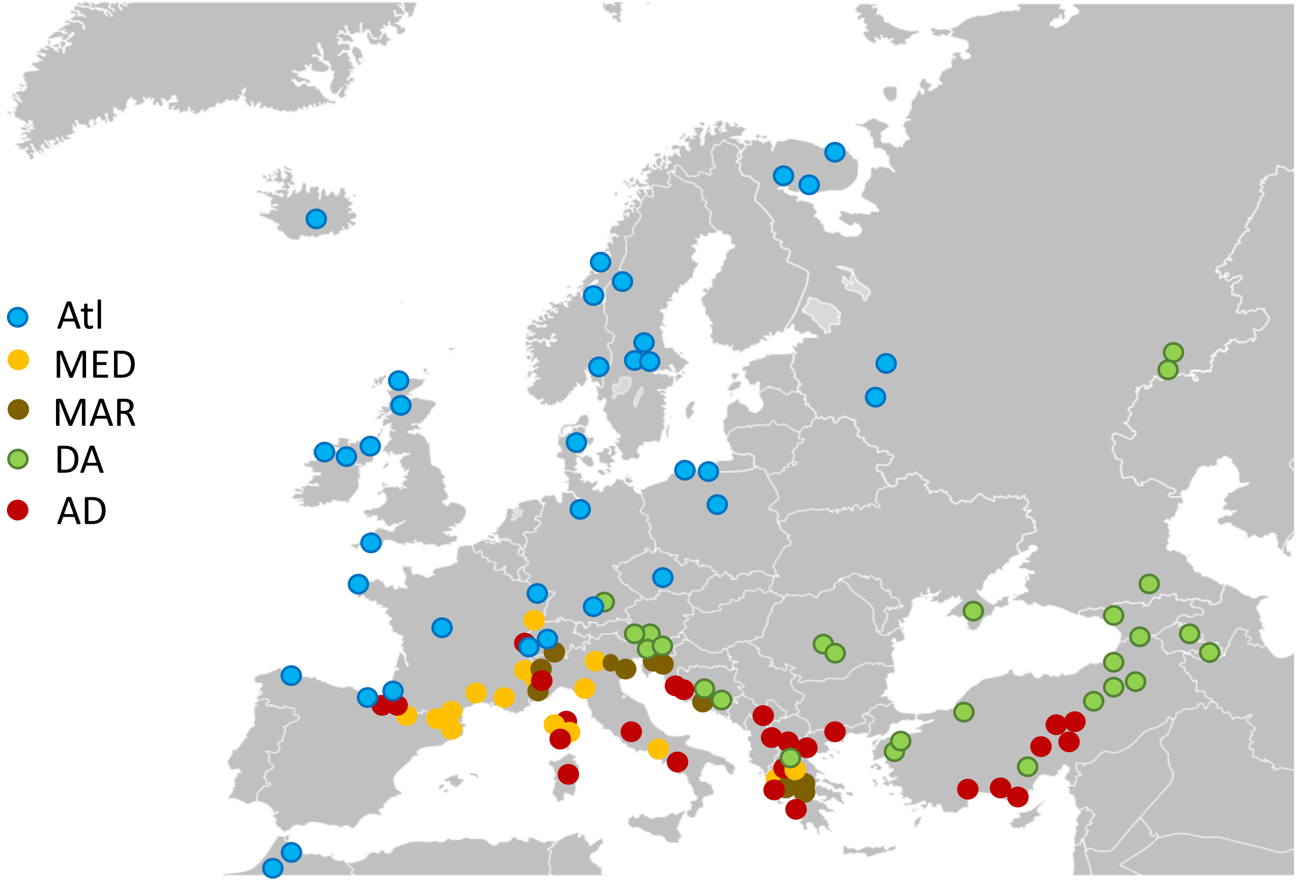


Bernatchez, 2001, Evolution v.55

Marbled

Mediterranean

Atlantic



Trout stocking

Enhancement of Atlantic and Mediterranean domestic strains into wild Mediterranean populations

That's a wild one !

Main Questions

What are the impacts of restocking on wild population structure and genetic diversity ?

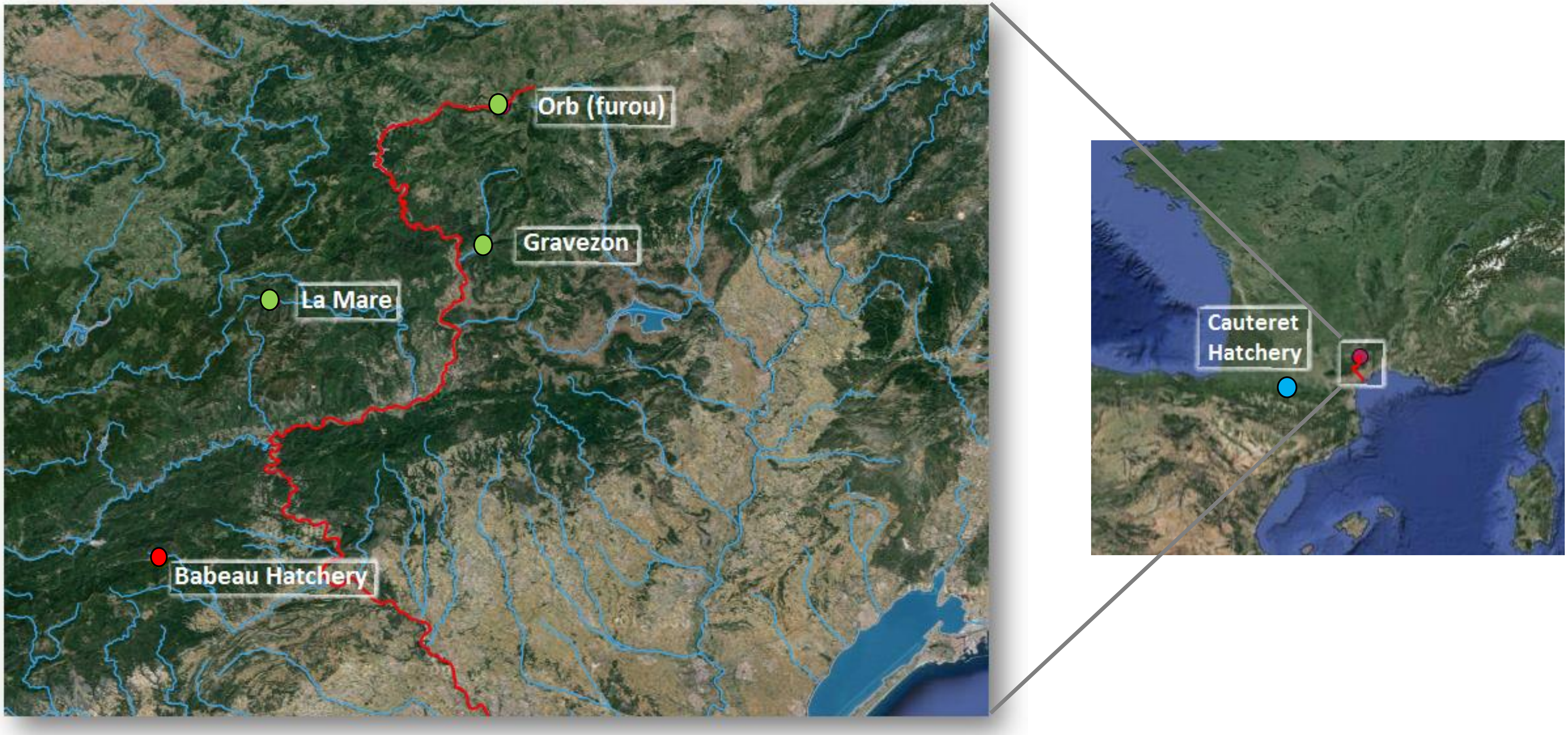
Main Questions

What are the impacts of restocking on wild population structure and genetic diversity ?

Do domestic alleles introgress into wild populations ?
What is the fitness of the introgressed alleles ?

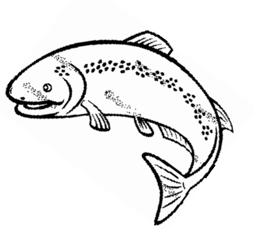
Part 1

→ Development of a high density SNPs array by dd-RAD sequencing



🟢 Wild populations
🔵 Atlantic hatchery strains
🔴 Mediterranean hatchery strains

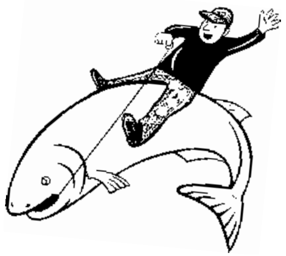
Wild



3 populations

30 individuals (3*10)

Domestic



2 populations

30 individuals (1*10 + 2*10)

Double digest RAD seq

(Paired End, 2*125 bp)

<u>Average number of reads per individual :</u>

12 millions

Workflow

→ *De novo*

→ Reference mapping

→ Comparison between both methods

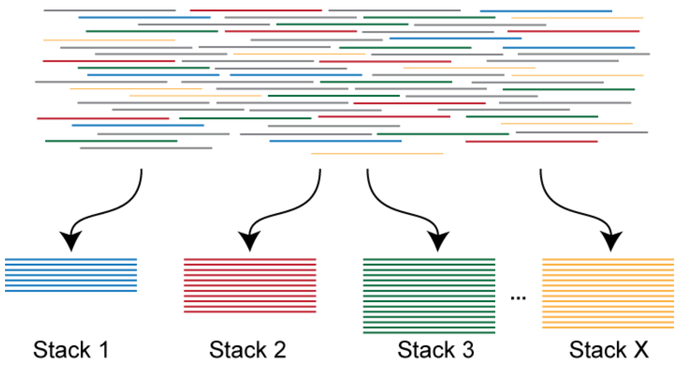
→ Estimation of nucleotide diversity

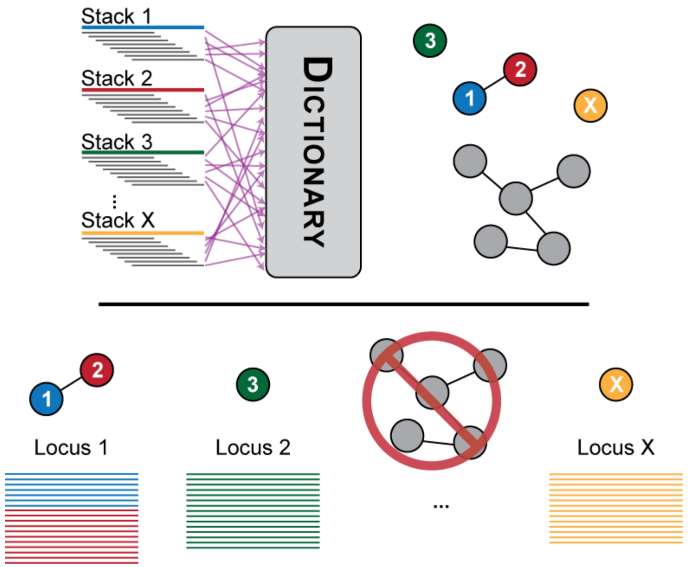
→ Population structure

Stacks Catchen et al., 2013

De novo assembly

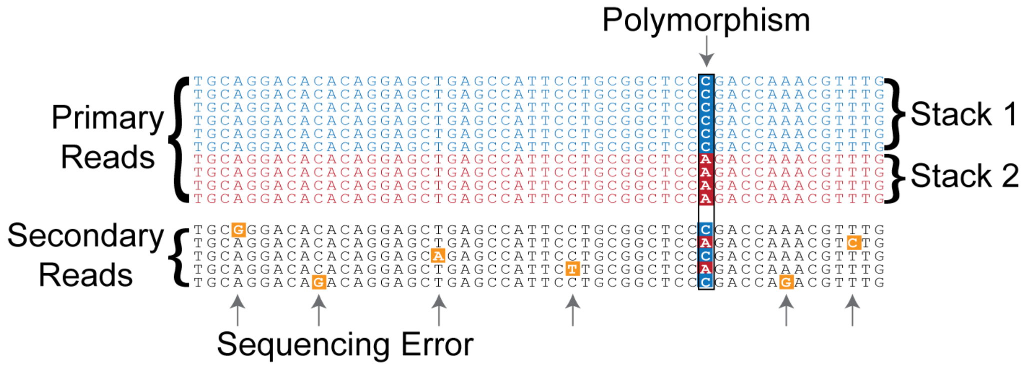
Ustacks





m= Minimum depth of coverage required to create a stack

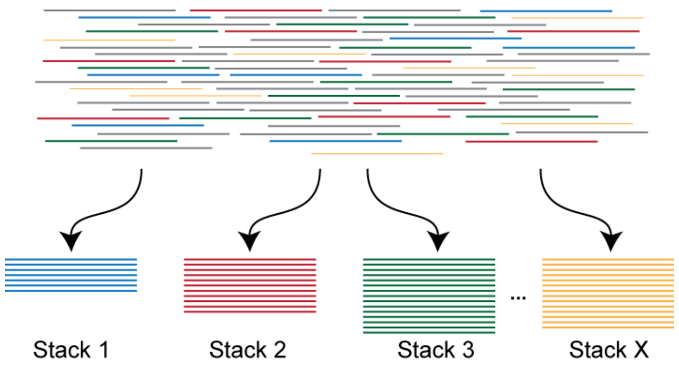
M= Maximum distance allowed between stacks

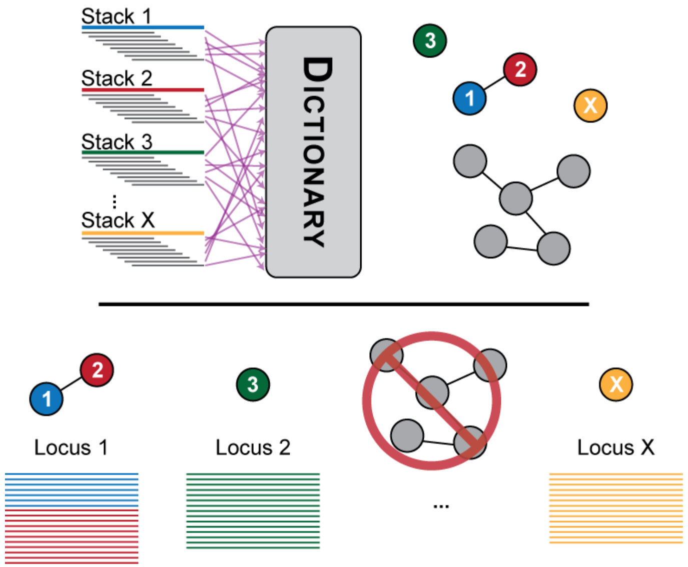


SNP= Single Nucleotide polymorphism

De novo assembly

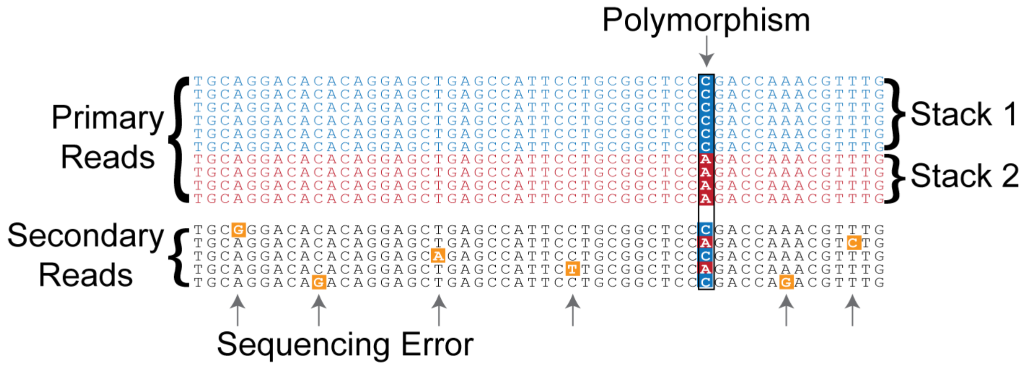
Ustacks





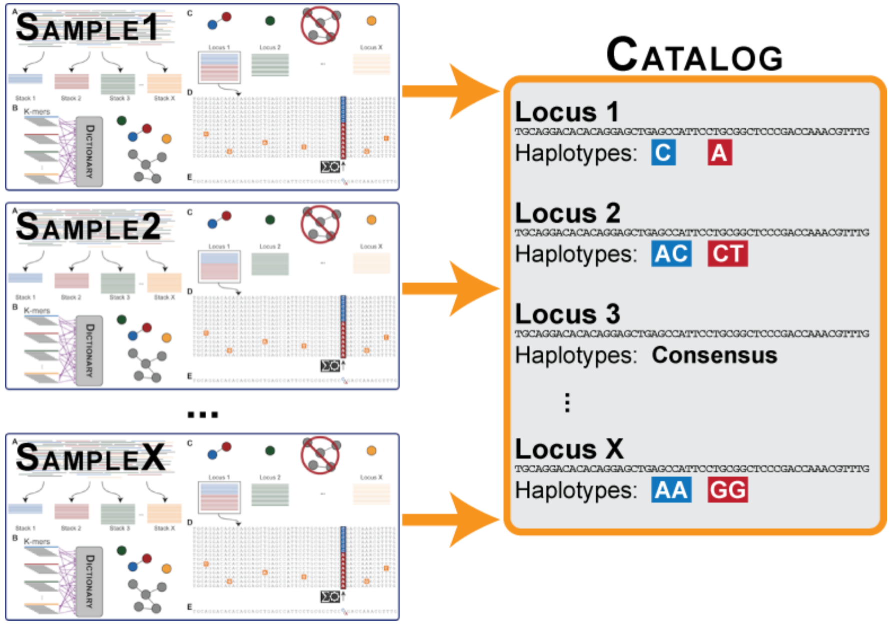
m= Minimum depth of coverage required to create a stack

M= Maximum distance allowed between stacks



SNP= Single Nucleotide polymorphism

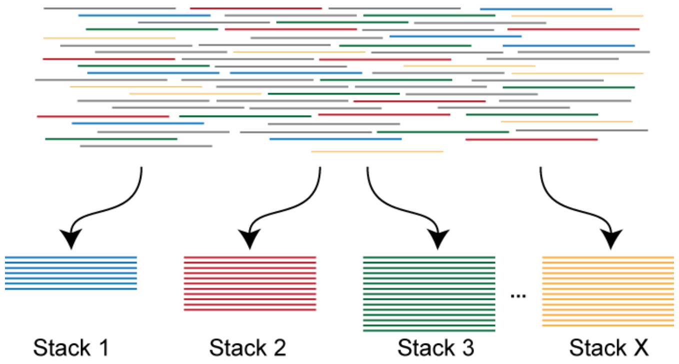
Cstacks

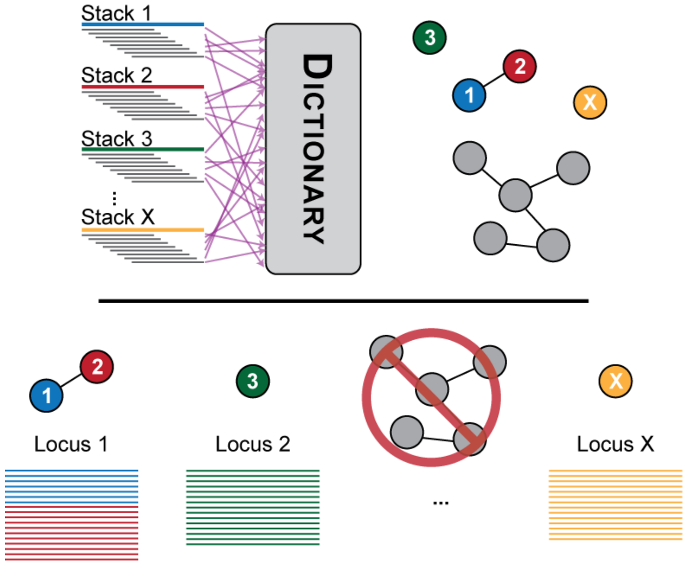


n= number of mismatches allowed between loci when generating the catalog

De novo assembly

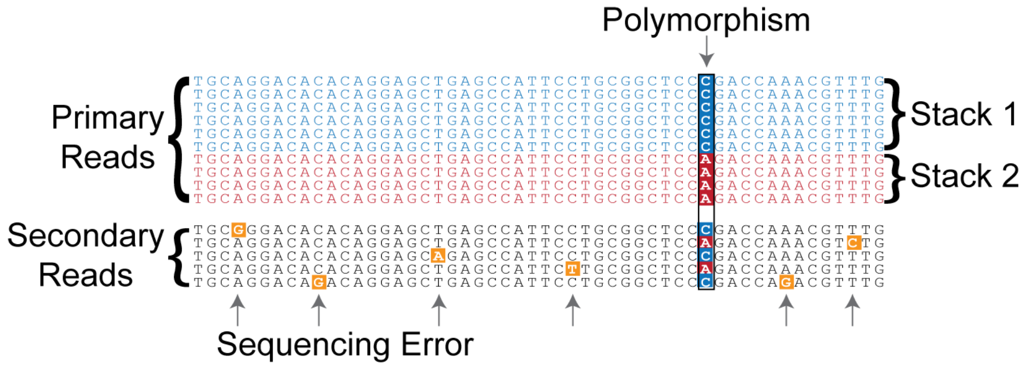
Ustacks





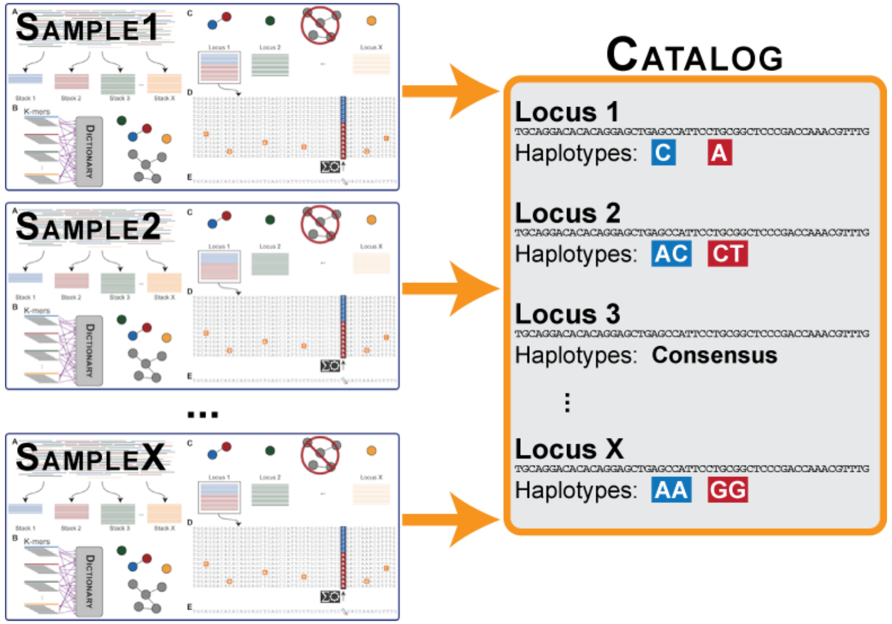
m= Minimum depth of coverage required to create a stack

M= Maximum distance allowed between stacks



SNP= Single Nucleotide polymorphism

Cstacks

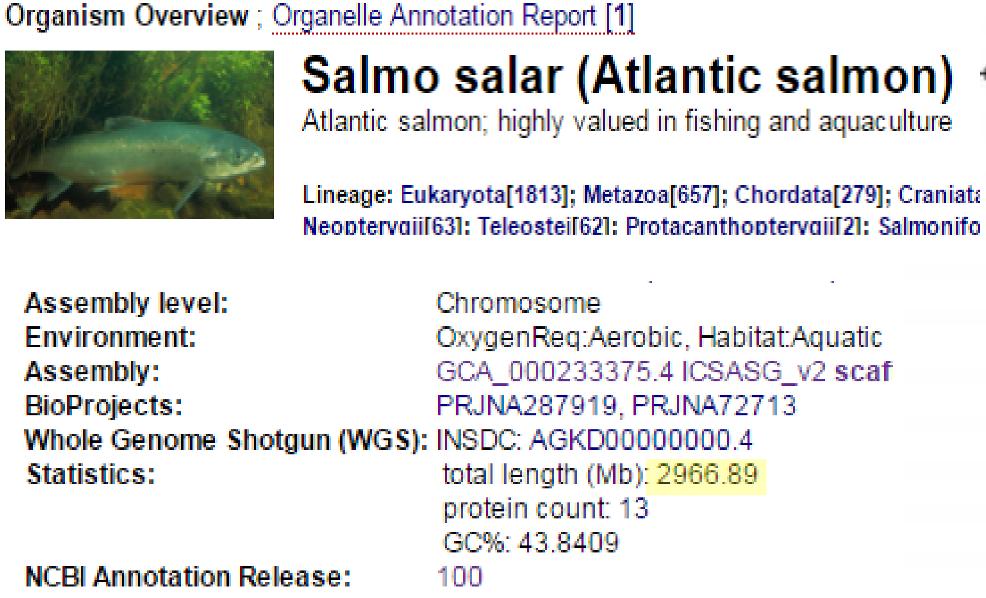


n= number of mismatches allowed between loci when generating the catalog

Sstacks

Match to the catalog

Reference mapping



Organism Overview ; Organelle Annotation Report [1]

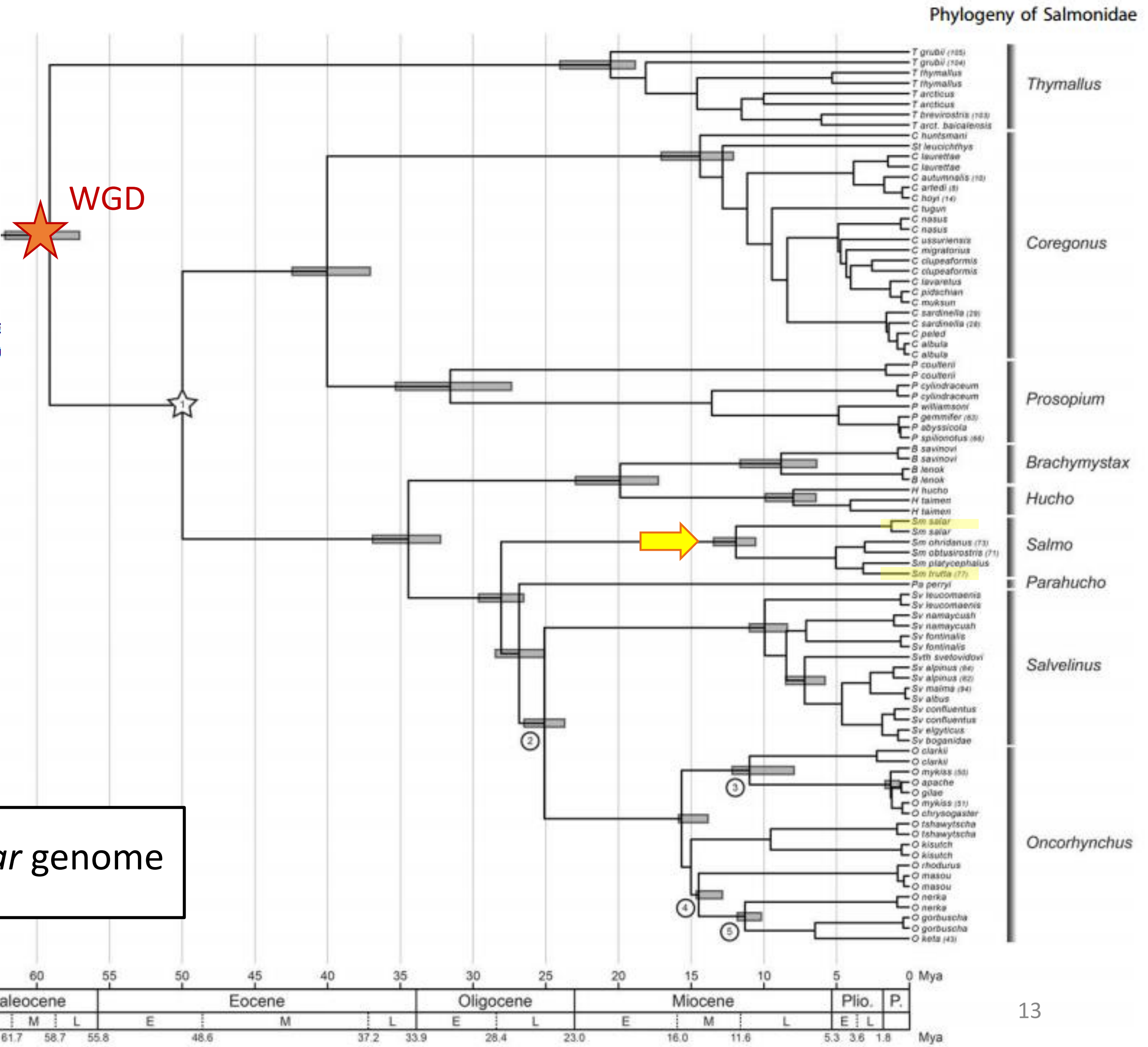
Salmo salar (Atlantic salmon)

Atlantic salmon; highly valued in fishing and aquaculture

Lineage: Eukaryota[1813]; Metazoa[657]; Chordata[279]; Craniata
Neopterygii[63]; Teleostei[62]; Protacanthopterygii[2]; Salmonifo

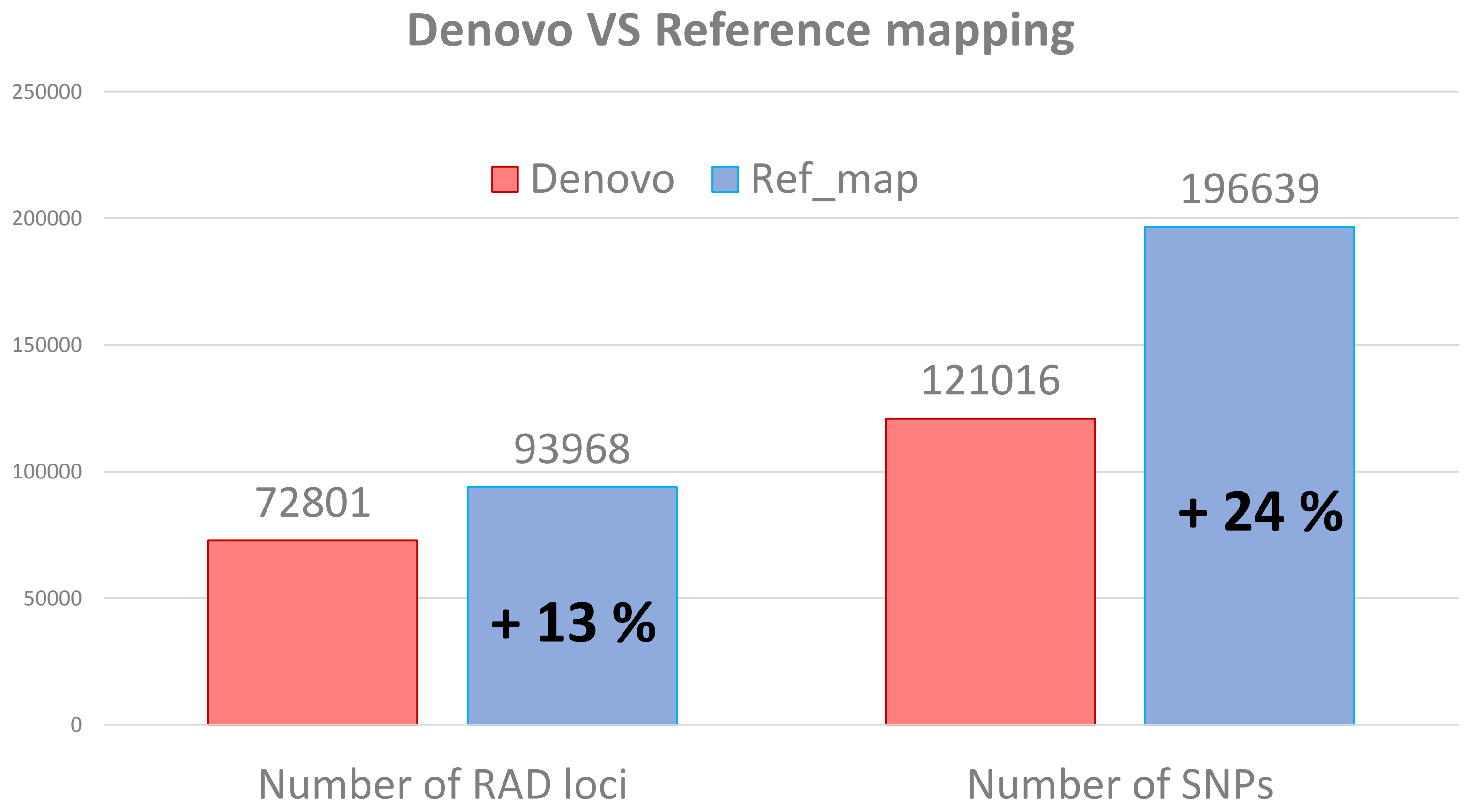
Assembly level:	Chromosome
Environment:	OxygenReq:Aerobic, Habitat:Aquatic
Assembly:	GCA_000233375.4 ICSASG_v2 scaf
BioProjects:	PRJNA287919, PRJNA72713
Whole Genome Shotgun (WGS):	INSDC: AGKD00000000.4
Statistics:	total length (Mb): 2966.89
	protein count: 13
	GC%: 43.8409
NCBI Annotation Release:	100

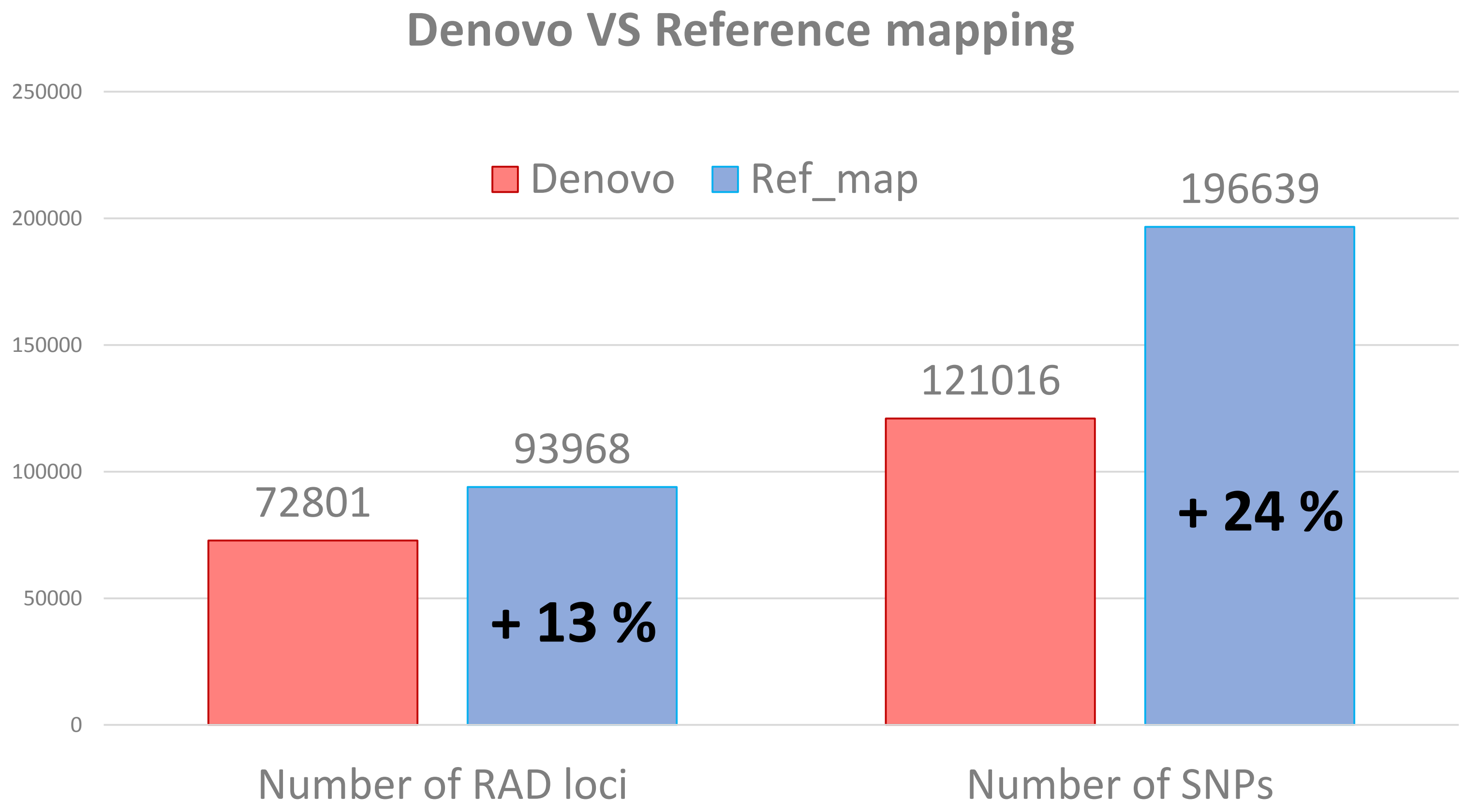
Lien *et al*. 2016


Phylogeny of Salmonidae

99.6 % of the reads mapped to *Salmo salar* genome

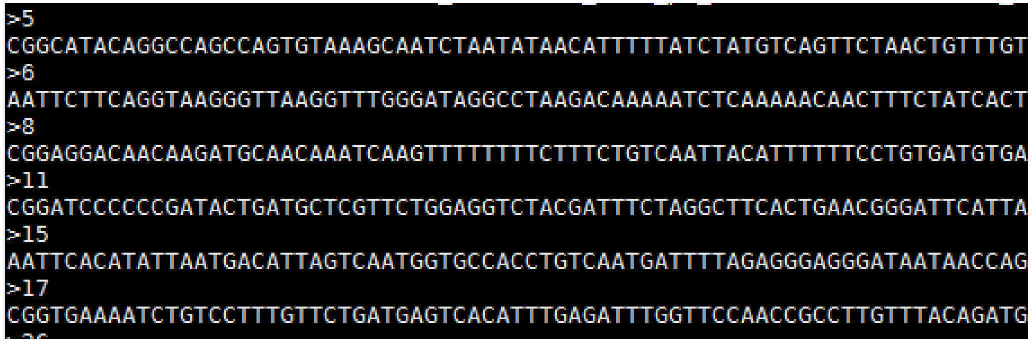
Crete-Lafreniere *et al*. 2012


Denovo VS Reference mapping
250000
200000
150000
100000
50000
0
🟥 Denovo
🟦 Ref_map
72801
93968
+ 13 %
121016
196639
+ 24 %
Number of RAD loci
Number of SNPs

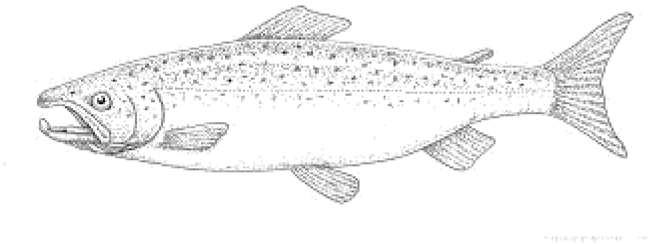


How many loci are common ?

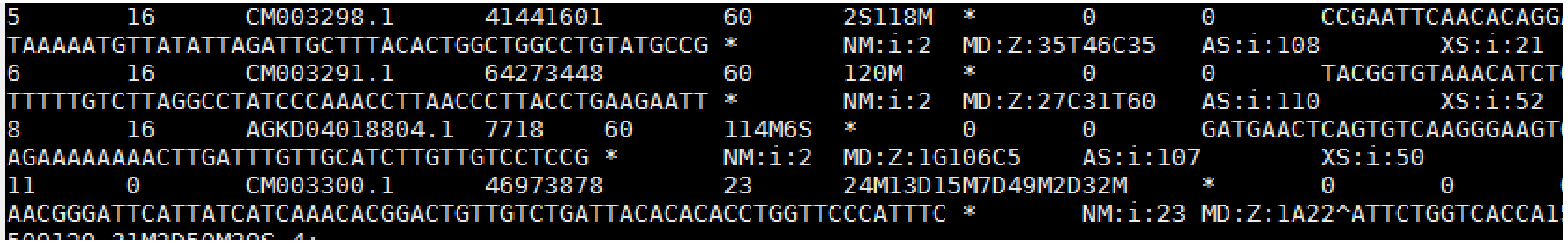
De novo loci from catalog



```
>5
CGGCATACAGGCCAGCCAGTGTAAAGCAATCTAATATAACATTTTTATCTATGTCAGTTCTAACTGTTTGT
>6
AATTCTTCAGGTAAGGGTTAAGGTTTGGGATAGGCCTAAGACAAAAATCTCAAAAACAACTTTCTATCACT
>8
CGGAGGACAACAAGATGCAACAAATCAAGTTTTTTTTTCTTTCTGTCAATTACATTTTTTCCTGTGATGTGA
>11
CGGATCCCCCCGATACTGATGCTCGTTCTGGAGGTCTACGATTTCTAGGCTTCACTGAACGGGATTCATTA
>15
AATTCACATATTAATGACATTAGTCAATGGTGCCACCTGTCAATGATTTTAGAGGGAGGGATAATAACCAG
>17
CGGTGAAAATCTGTCCTTTGTTCTGATGAGTCACATTTGAGATTTGGTTCCAACCGCCTTGTTTACAGATG
```



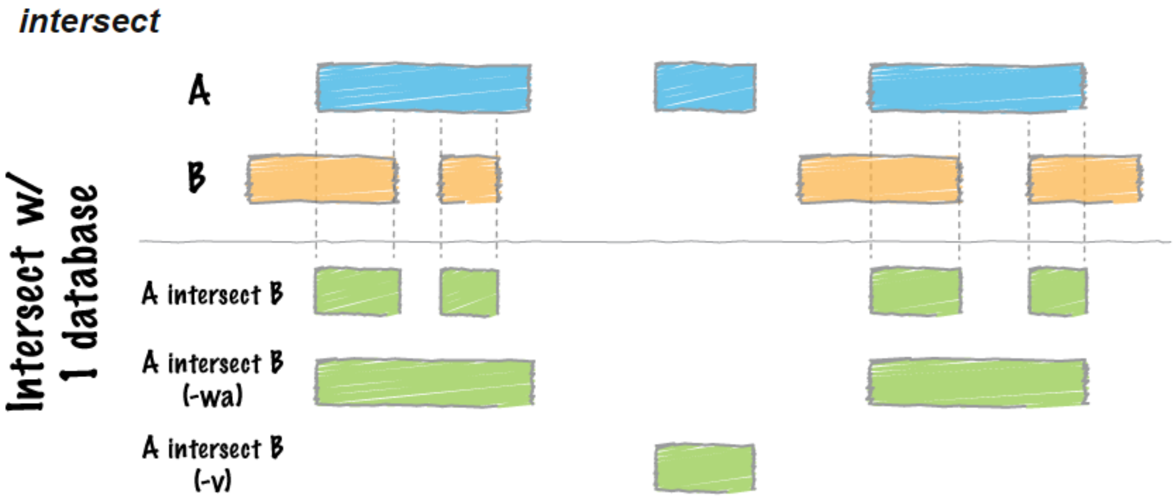
Sam alignement file

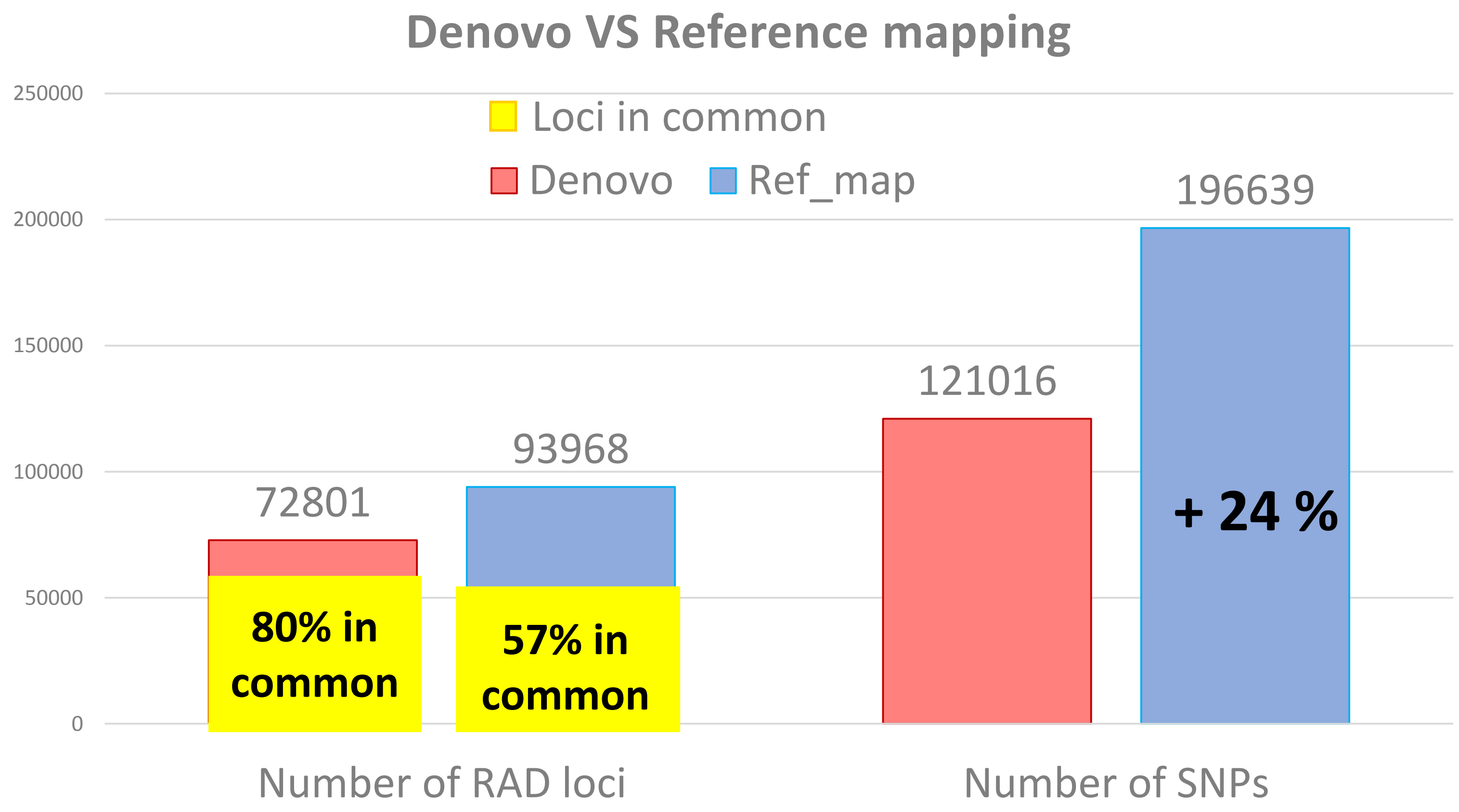


```
5       16      CM003298.1      41441601        60      2S118M  *       0       0       CCGAATTCAACACAGG
TAAAAATGTTATATTAGATTGCTTTACACTGGCTGGCCTGTATGCCG *       NM:i:2  MD:Z:35T46C35   AS:i:108        XS:i:21
6       16      CM003291.1      64273448        60      120M    *       0       0       TACGGTGTAAACATCT
TTTTTGTCTTAGGCCTATCCCAAACCTTAACCCTTACCTGAAGAATT *       NM:i:2  MD:Z:27C31T60   AS:i:110        XS:i:52
8       16      AGKD04018804.1  7718    60      114M6S  *       0       0       GATGAACTCAGTGTCAAGGGAAGT
AGAAAAAAAACTTGATTTGTTGCATCTTGTTGTCCTCCG *       NM:i:2  MD:Z:1G106C5    AS:i:107        XS:i:50
11      0       CM003300.1      46973878        23      24M13D15M7D49M2D32M     *       0       0
AACGGGATTCATTATCATCAAACACGGACTGTTGTCTGATTACACACACCTGGTTCCCATTTC *       NM:i:23 MD:Z:1A22^ATTCTGGTCACCA1
```

87% (63 164) loci mapped
(104 139 SNPs)

Identifying loci in common



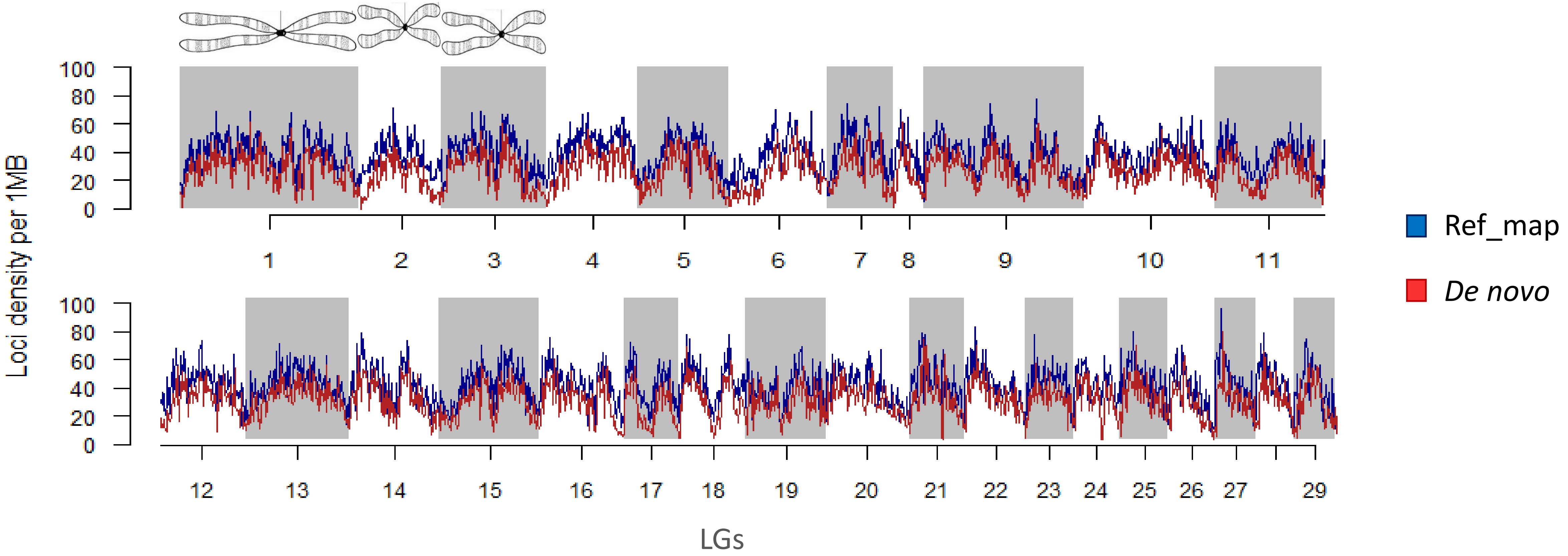


How are they distributed on chromosomes ?

Distribution of brown trout loci along *S. salar* LGs

S. trutta

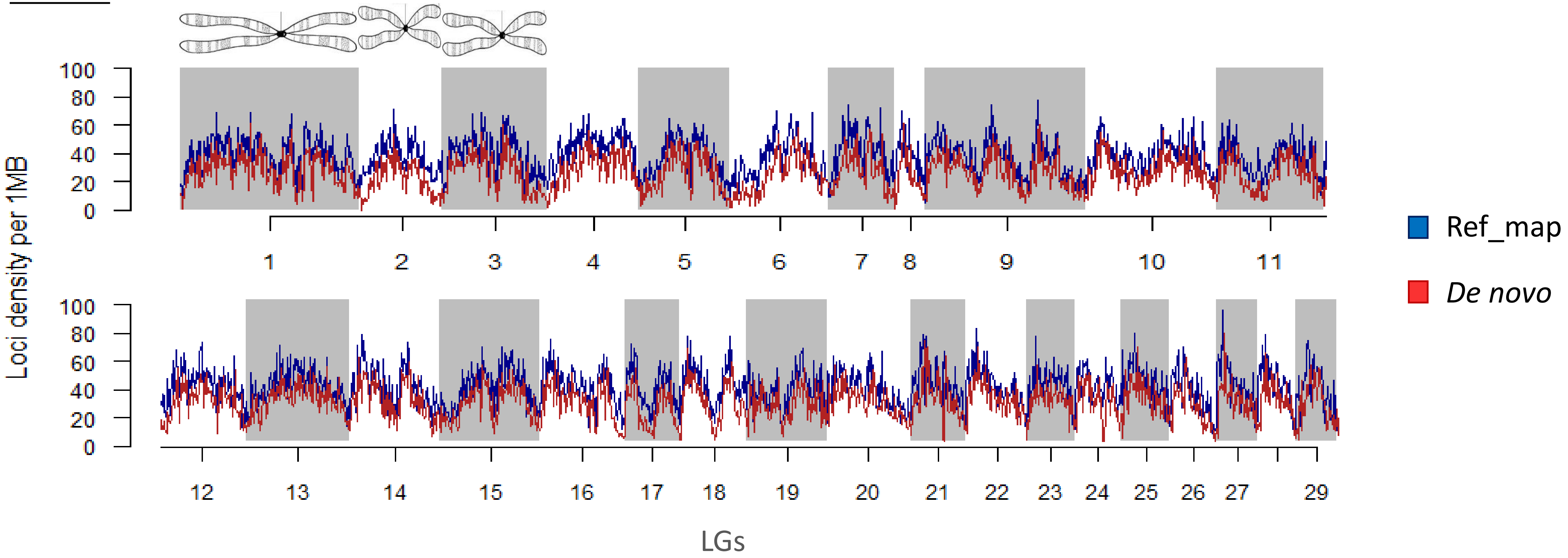
S. Salar



Distribution of Brown trout loci along *S. salar* LGs

S. trutta

S. Salar



→ Homogeneous distribution

→ Lower density at LGs extremities : paralogous loci ?

Workflow

→ *De novo*

→ Reference mapping

→ Comparison between both methods

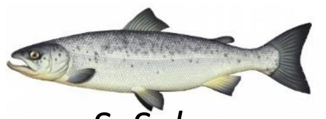
→ Estimation of nucleotide diversity

→ Population structure

Stacks Catchen et al., 2013

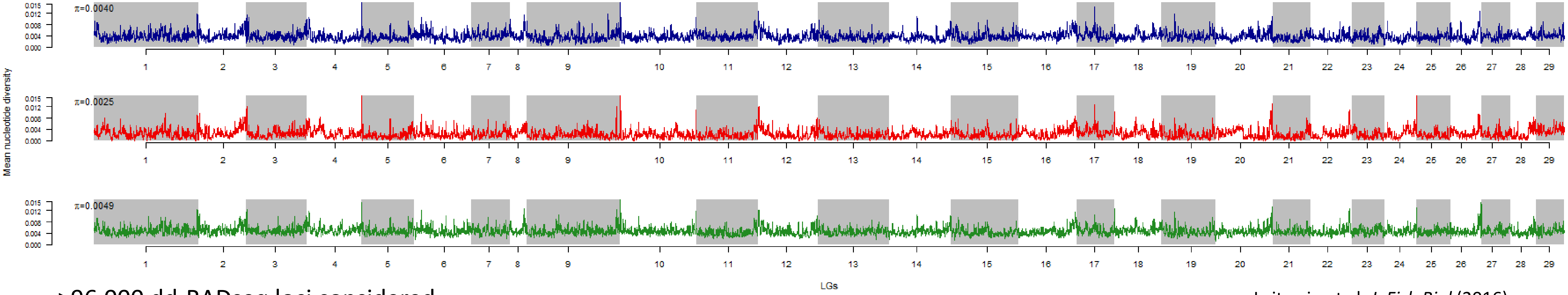
S. trutta

Distribution of brown trout nucleotide diversity along Atlantic salmon (*S. salar)* linkage groups



S. Salar

🟦 Atlantic strain ,	mean π = 0.0040
🟥 Mediterranean strain,	mean π = 0.0025
🟩 Wild Mediterranean trout,	mean π = 0.0049



>96,000 dd-RADseq loci considered

Leitwein et al. *J. Fish Biol* (2016)

Distribution of brown trout nucleotide diversity along Atlantic salmon (*S. salar)* linkage groups

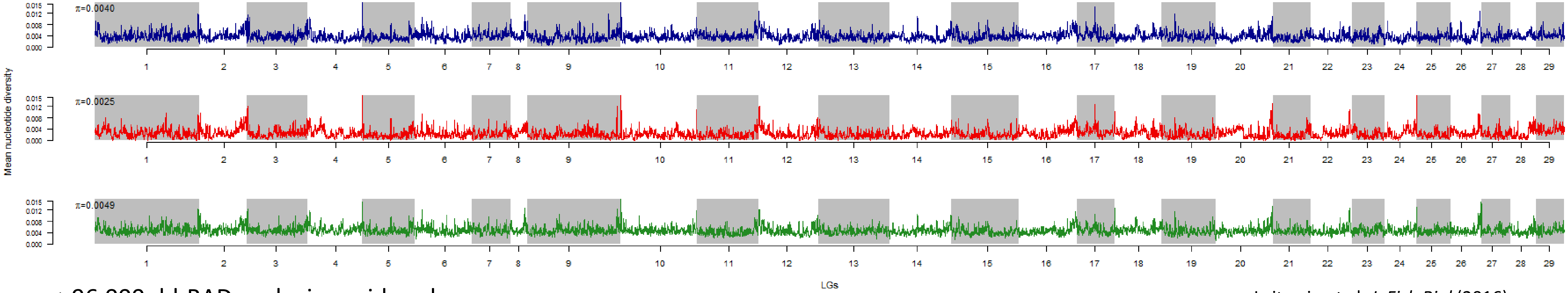
S. trutta

S. Salar

🟦 Atlantic strain , mean π = 0.0040

🟥 Mediterranean strain, mean π = 0.0025

🟩 Wild Mediterranean trout, mean π = 0.0049



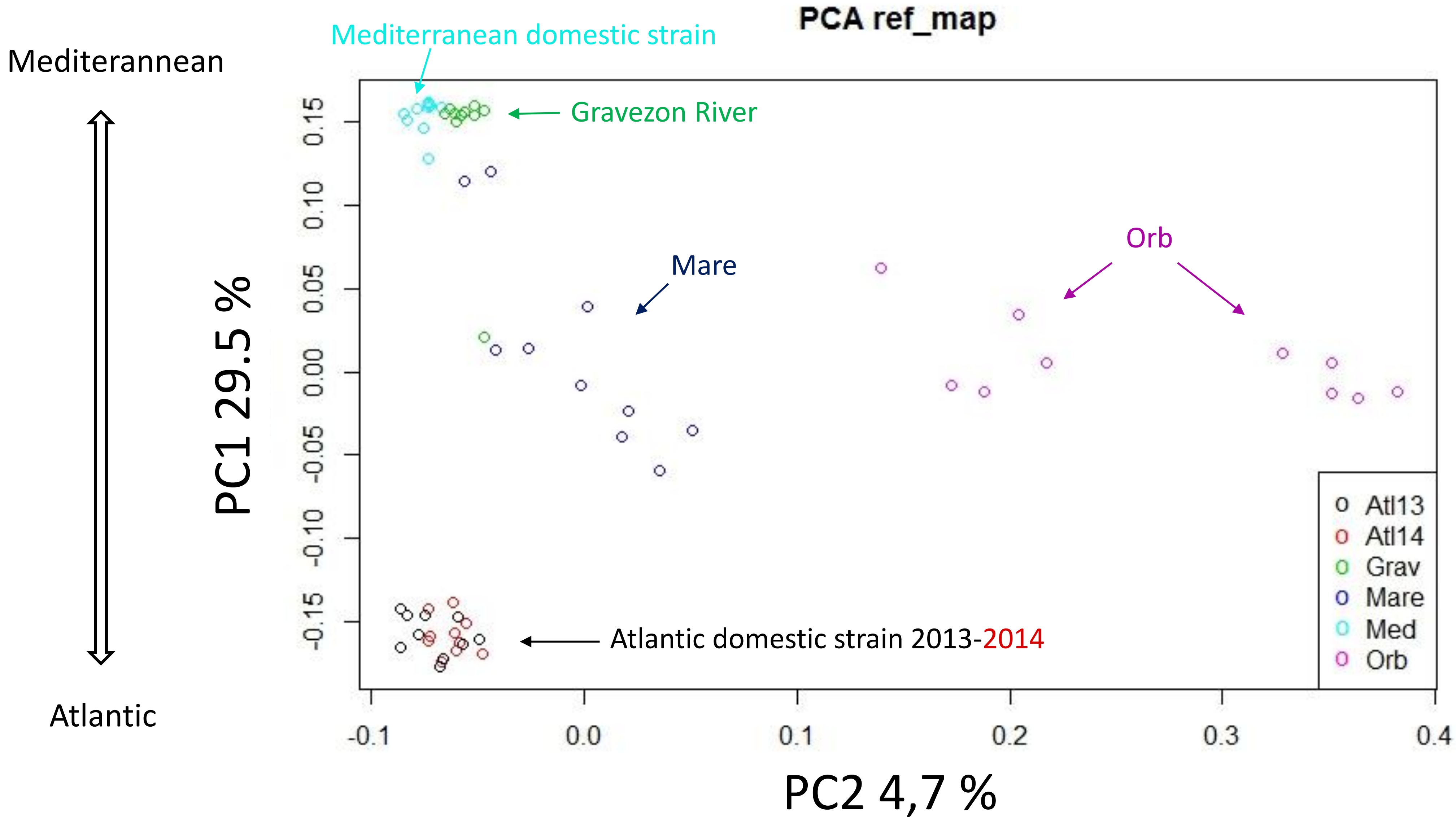
>96,000 dd-RADseq loci considered

Leitwein et al. *J. Fish Biol* (2016)

→ Lower average diversity for the Mediterranean strain

→'Classical' higher diversity at LGs extremities

Principal Component Analysis (PCA)



Mediterannean

Atlantic

PC1 29.5 %

PC2 4,7 %

Mediterranean domestic strain

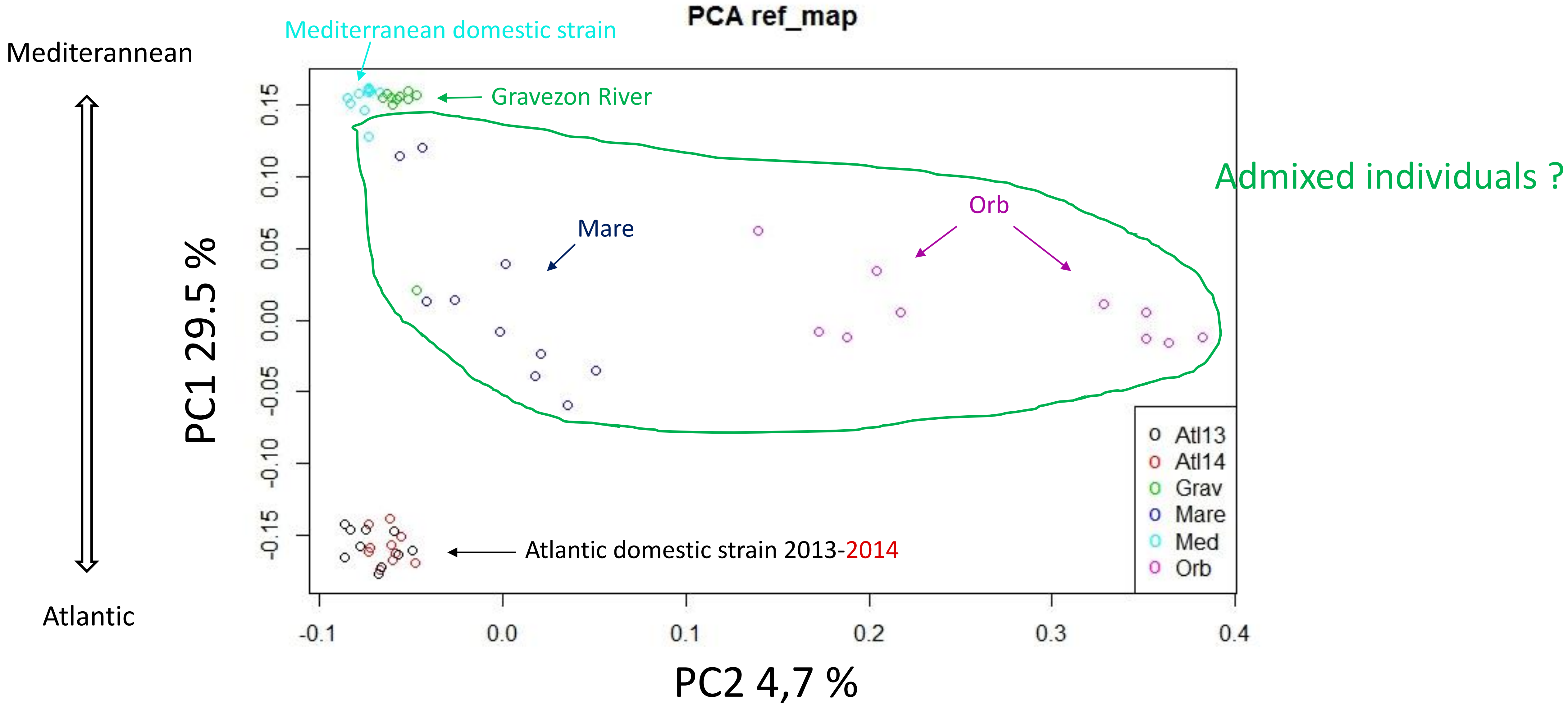
Gravezon River

Mare

Orb

Atlantic domestic strain 2013-2014

Principal Component Analysis (PCA)



Part 1

Concluding remarks

→ **Large number of SNPs** discovered; **similar genome-wide distribution** along LGs using both approaches

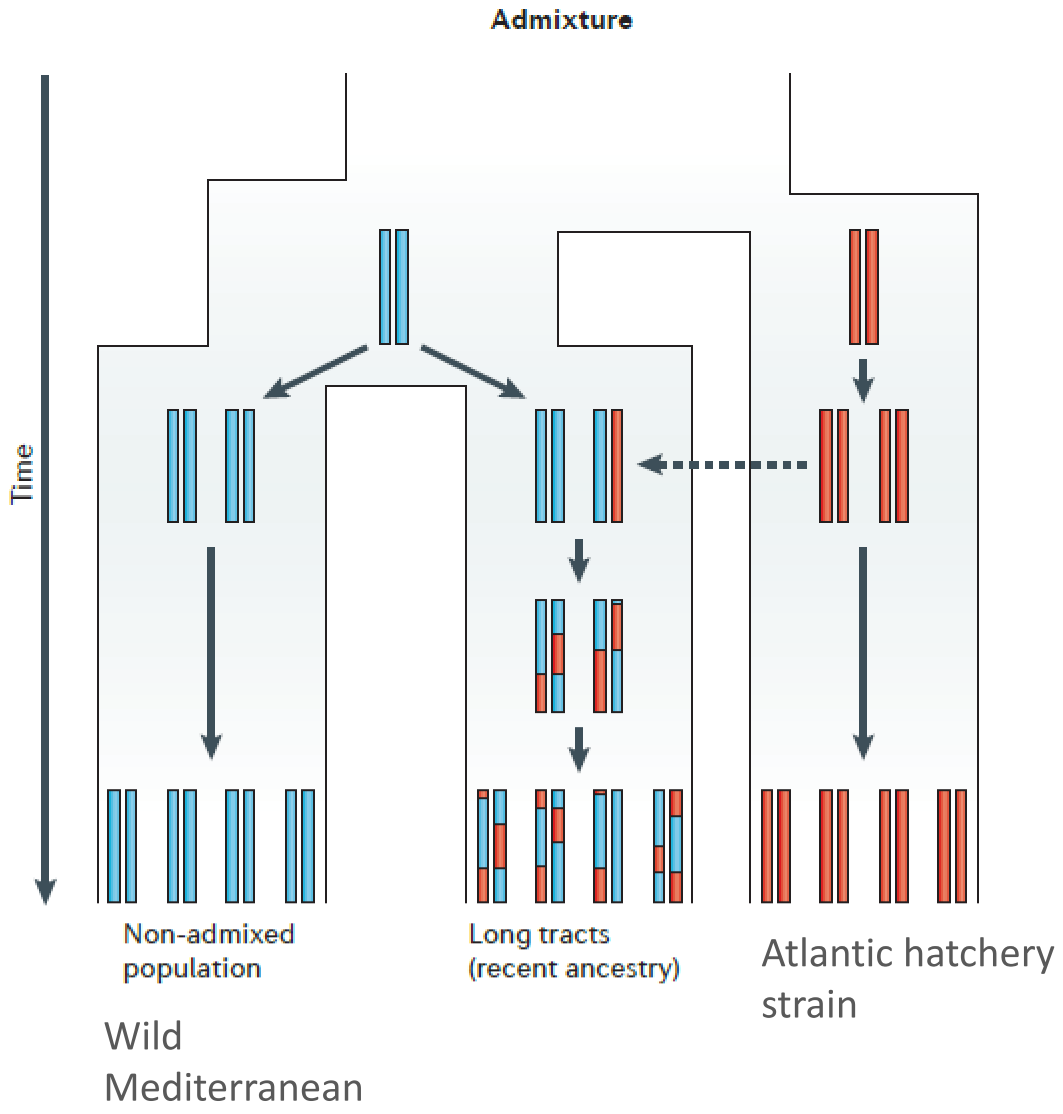
→ **Lower mapping density** and **higher nucleotide diversity** at LGs extremities → footprint of **residual tetrasomy** in these regions ?

→The **Mediterranean hatchery strain** is twice less diversified than **wild Mediterranean** samples

→ **Presence of admixed individuals in wild populations**

→ Presence of admixed individuals :

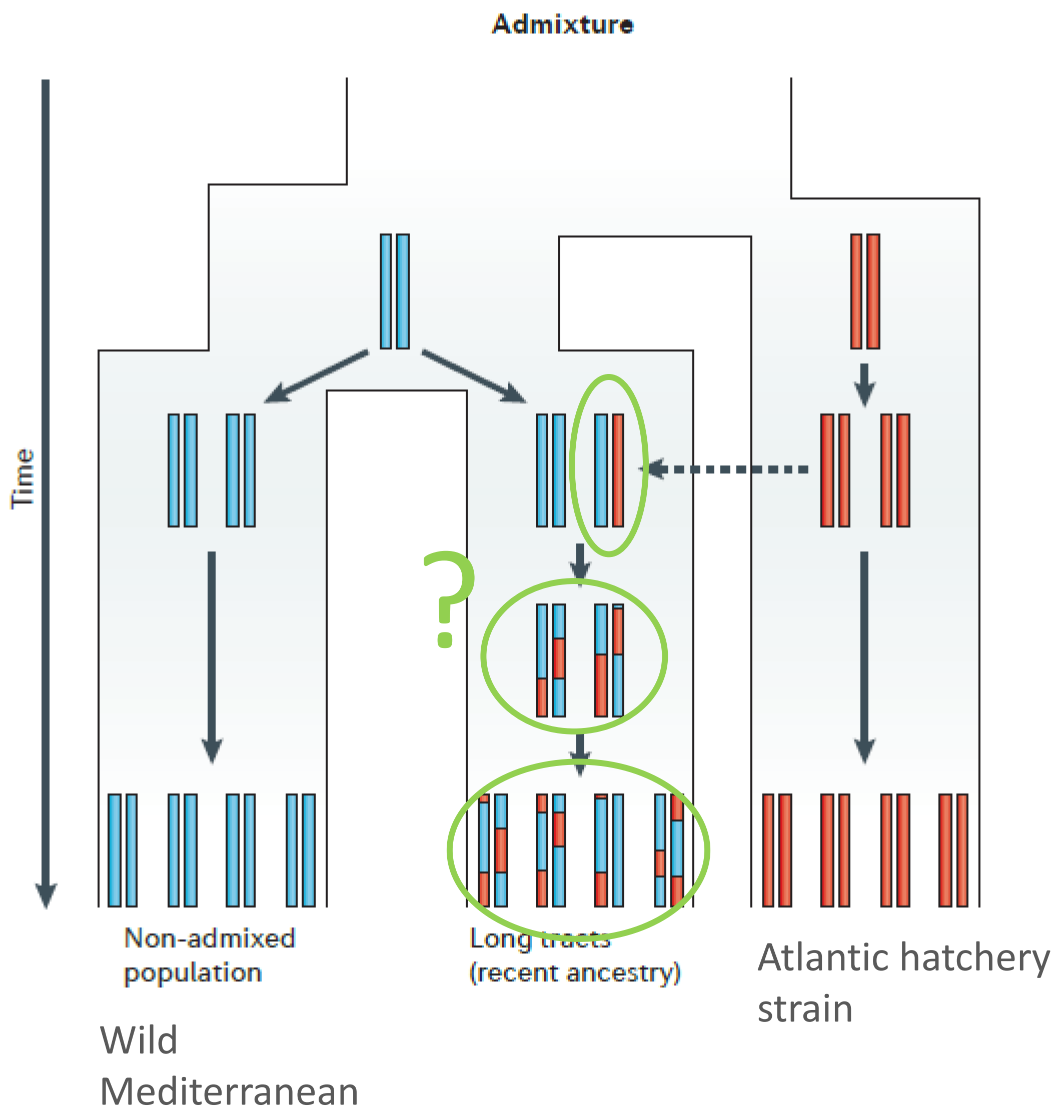
What are the distribution patterns of admixture within the genome ?



Racimo et al. (2015)

→ Presence of admixed individuals :

What are the distribution patterns of admixture within the genome ?



Racimo et al. (2015)

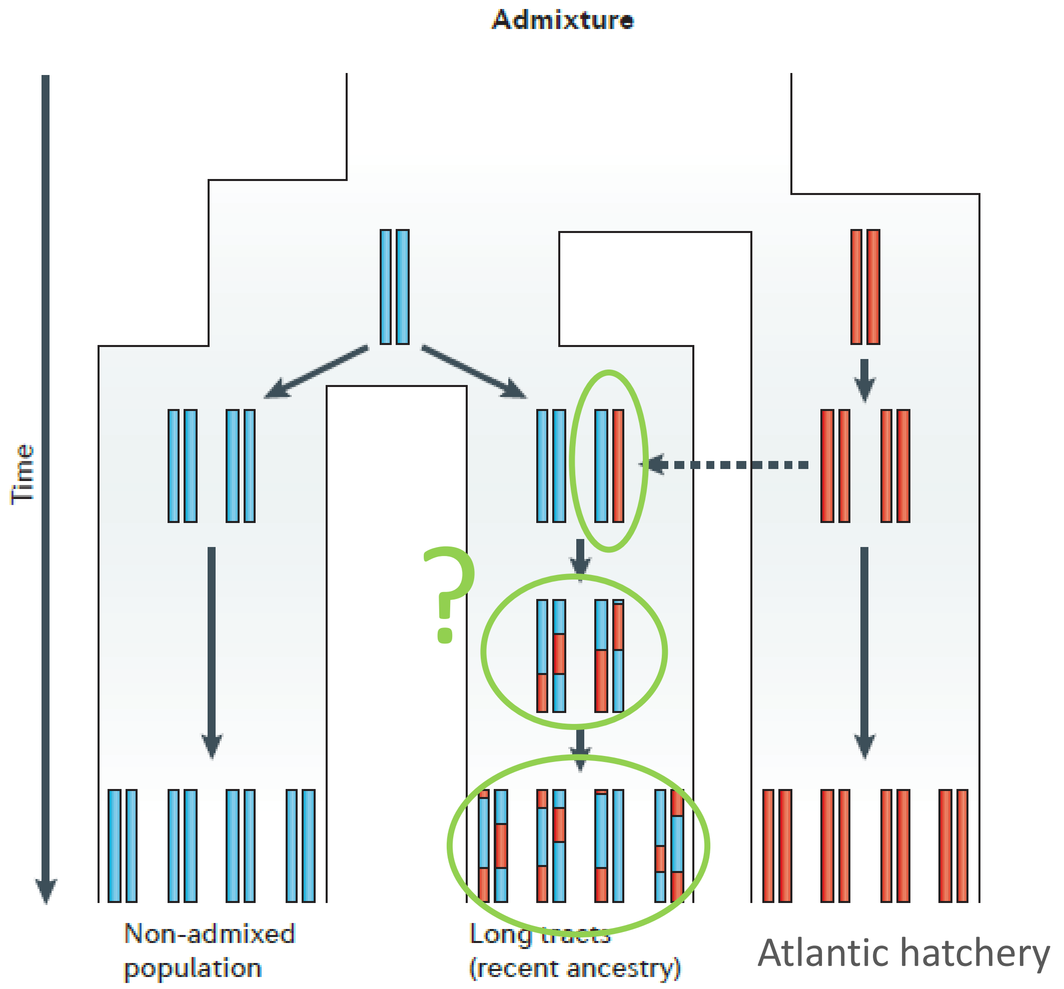
→ Presence of admixed individuals :

What are the distribution patterns of admixture within the genome ?

To estimate the introgression rate along the genome we need :

-> to assess the introgression at haplotype level

-> to infer the recombination rate landscape



Atlantic hatchery strain

Wild Mediterranean

Racimo et al. (2015)

→ Presence of admixed individuals :

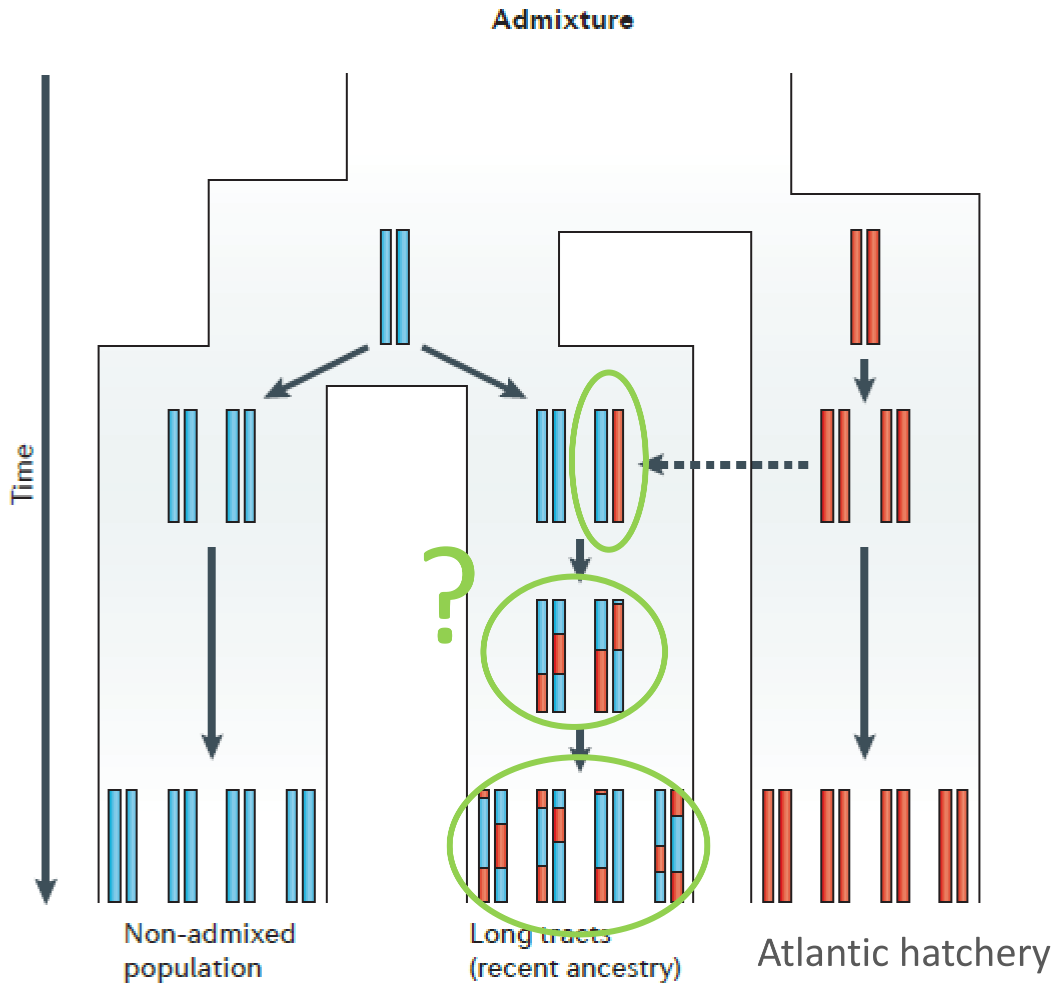
What are the distribution patterns of admixture within the genome ?

To estimate the introgression rate along the genome we need :

-> to assess the introgression at haplotype level

-> to infer the recombination rate landscape

Need for ***S. trutta*** linkage groups



Atlantic hatchery strain

Wild Mediterranean

Racimo et al. (2015)

Part 2

High density *S. trutta* linkage map

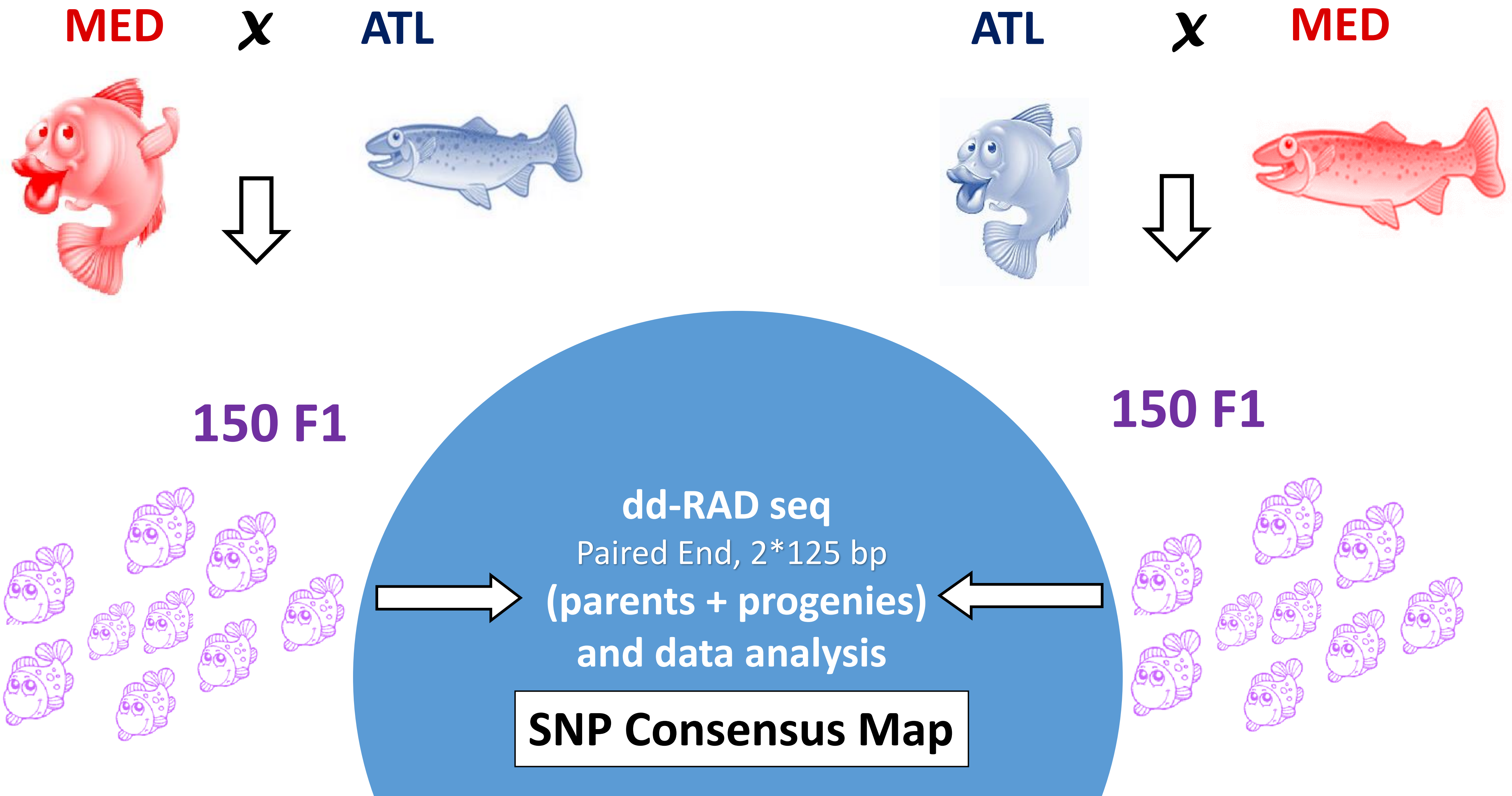
Gharbi et al.'s (2006) brown trout linkage map

→ Based on microsatellite (N = 288) + allozyme (N= 13) markers

→ **Incomplete: 37** LGs found but **40** LGs expected from karyotypic studies

→ Might be improved by high-throughput genomic techniques

The hybrid linkage map design



MED *x* **ATL**

ATL *x* **MED**

150 F1

150 F1

dd-RAD seq

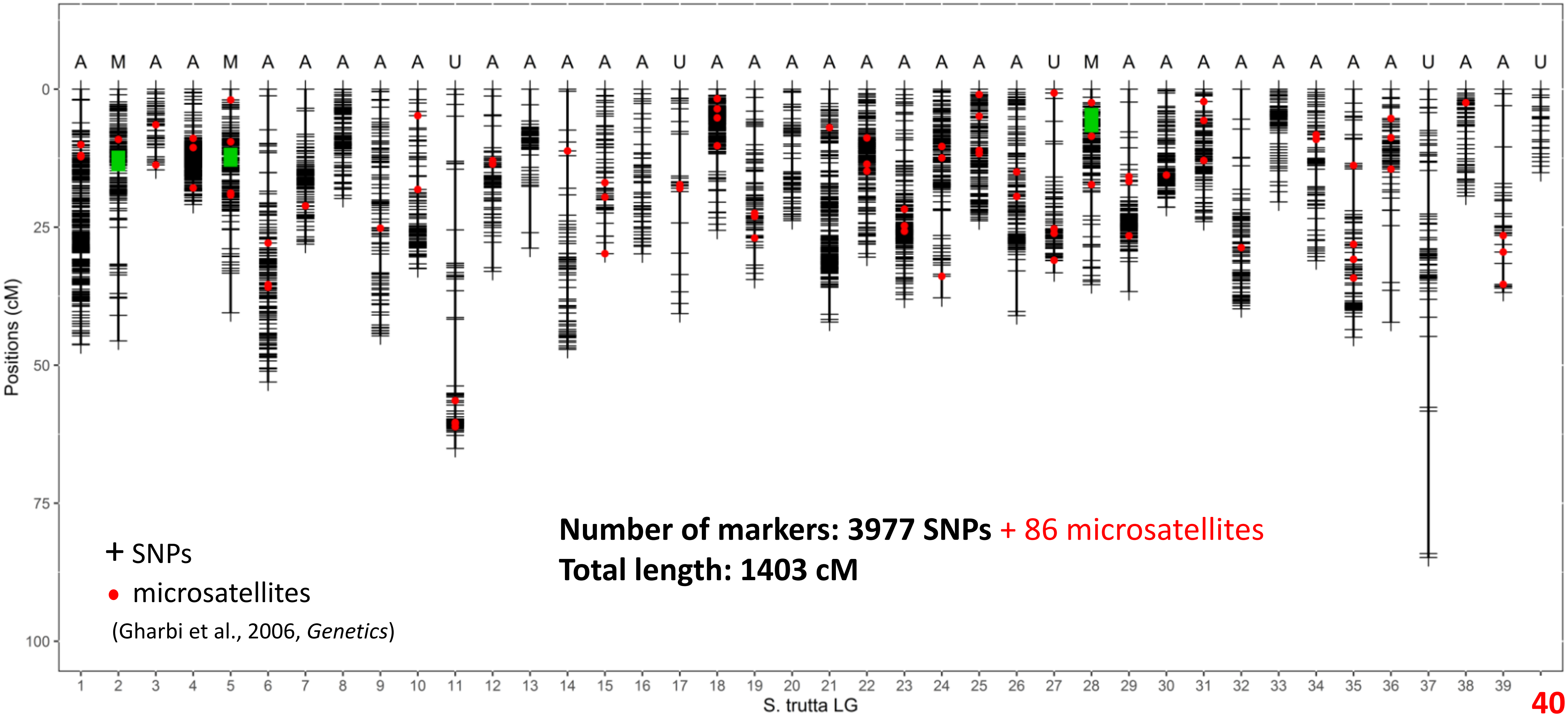
Paired End, 2*125 bp

(parents + progenies)

and data analysis

SNP Consensus Map

A new *Salmo trutta* linkage map



+ SNPs

● microsatellites

(Gharbi et al., 2006, *Genetics*)

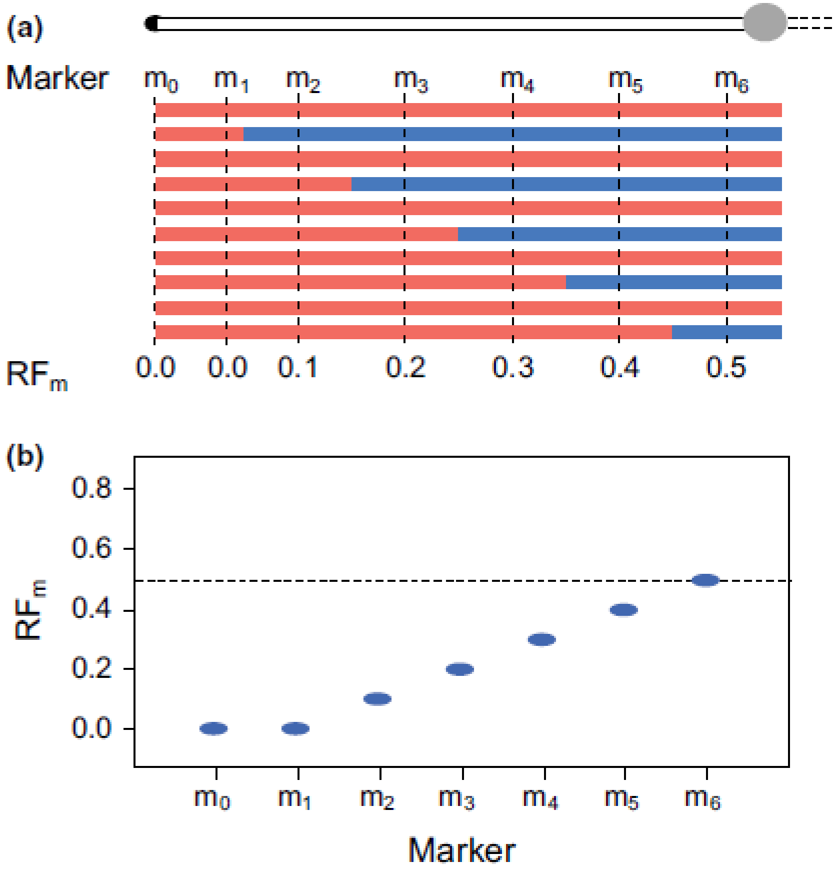
Number of markers: 3977 SNPs + 86 microsatellites

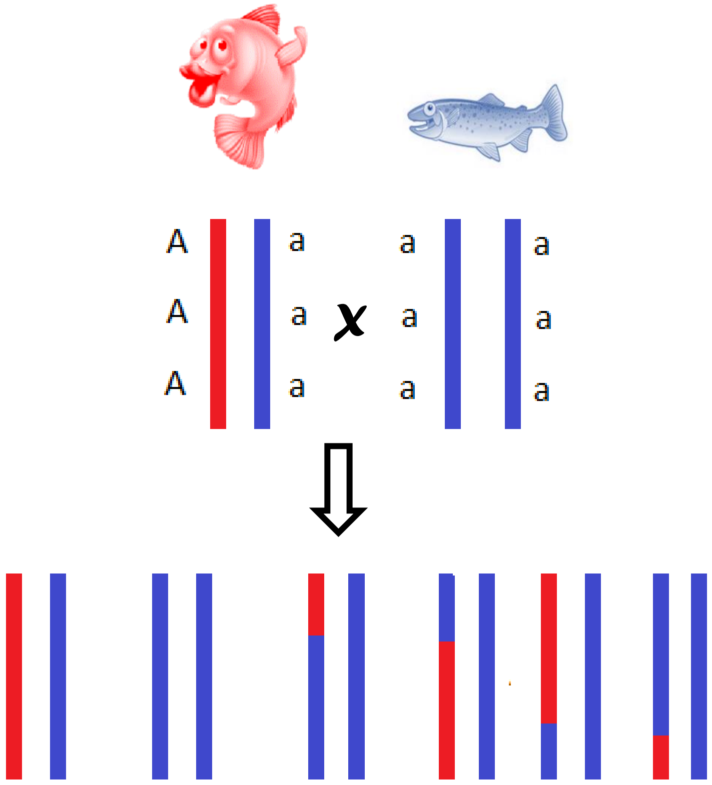
Total length: 1403 cM

40

We got them!

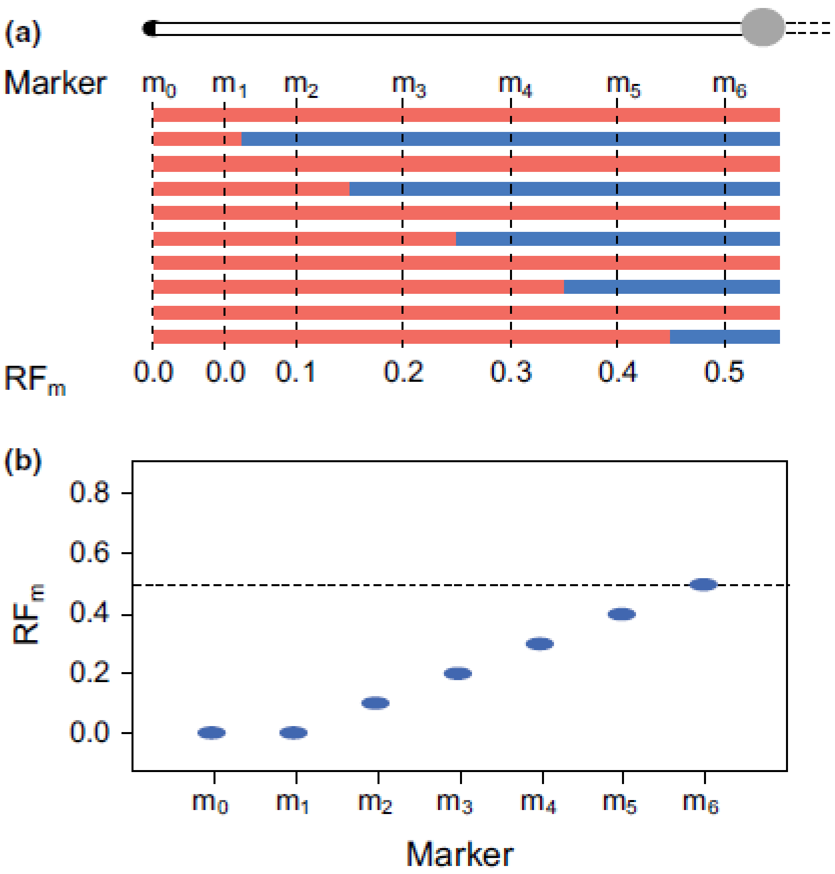
Centromere location and chromosome type

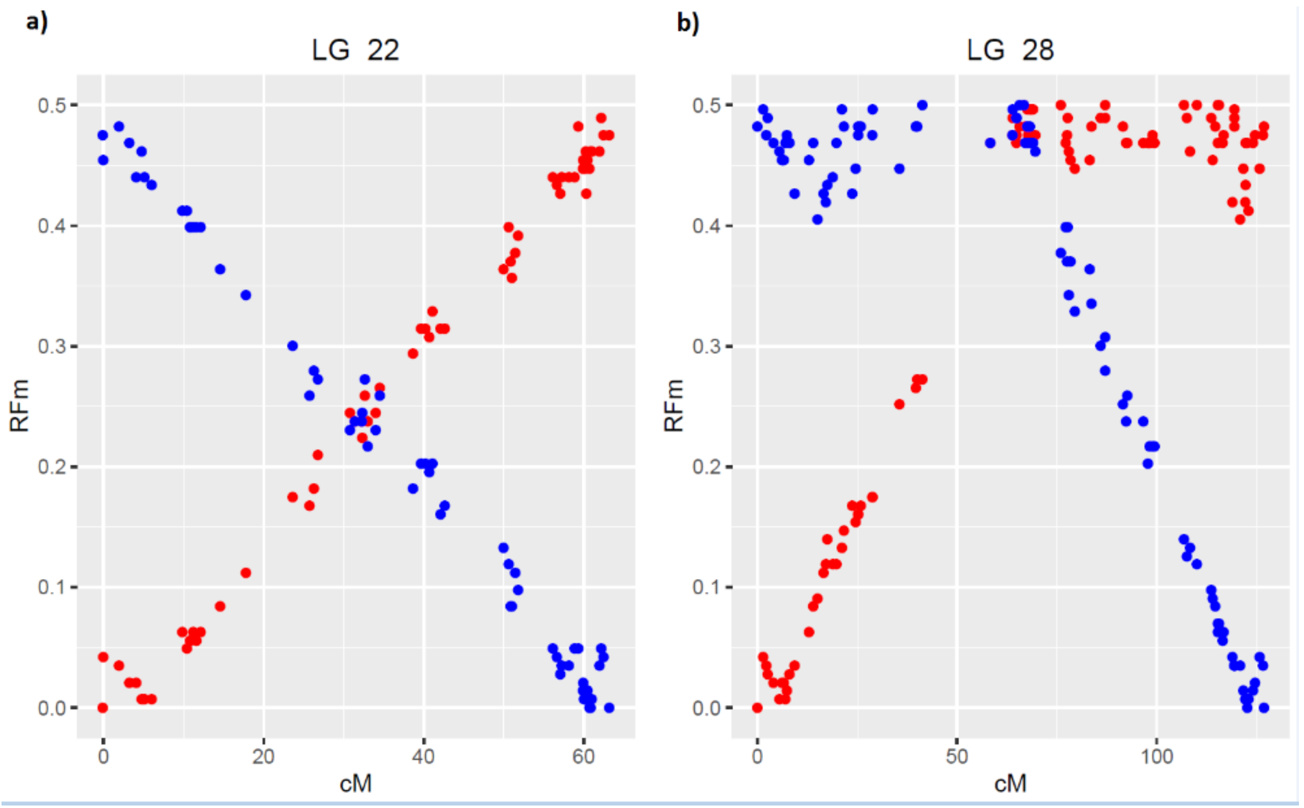




Limborg et al., 2015

Centromere location and chromosome type



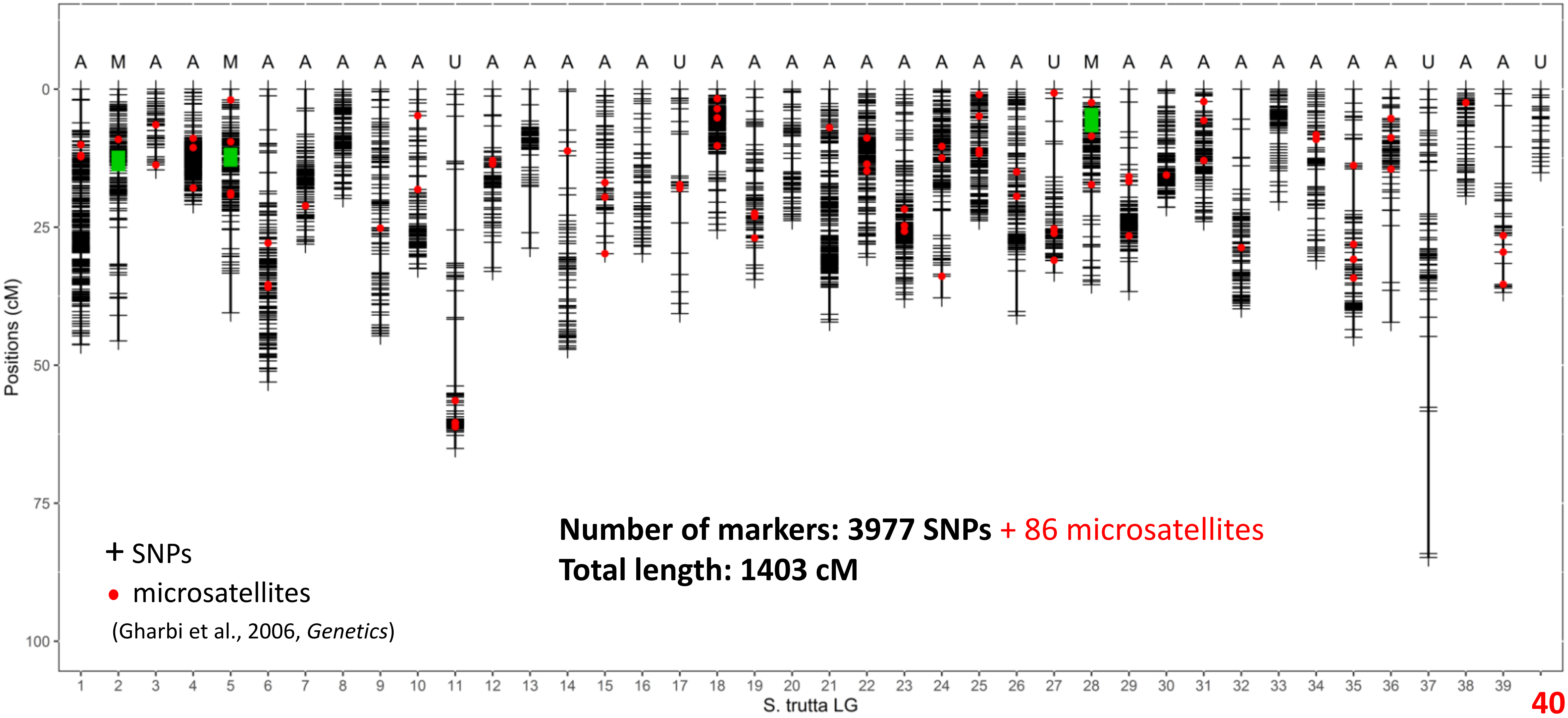


Acrocentric

Metacentric

Limborg et al., 2015

A new *Salmo trutta* linkage map



+ SNPs

● microsatellites

(Gharbi et al., 2006, *Genetics*)

Number of markers: 3977 SNPs + 86 microsatellites

Total length: 1403 cM

40

We got them!

MapComp

Sutherland et al. (2016)

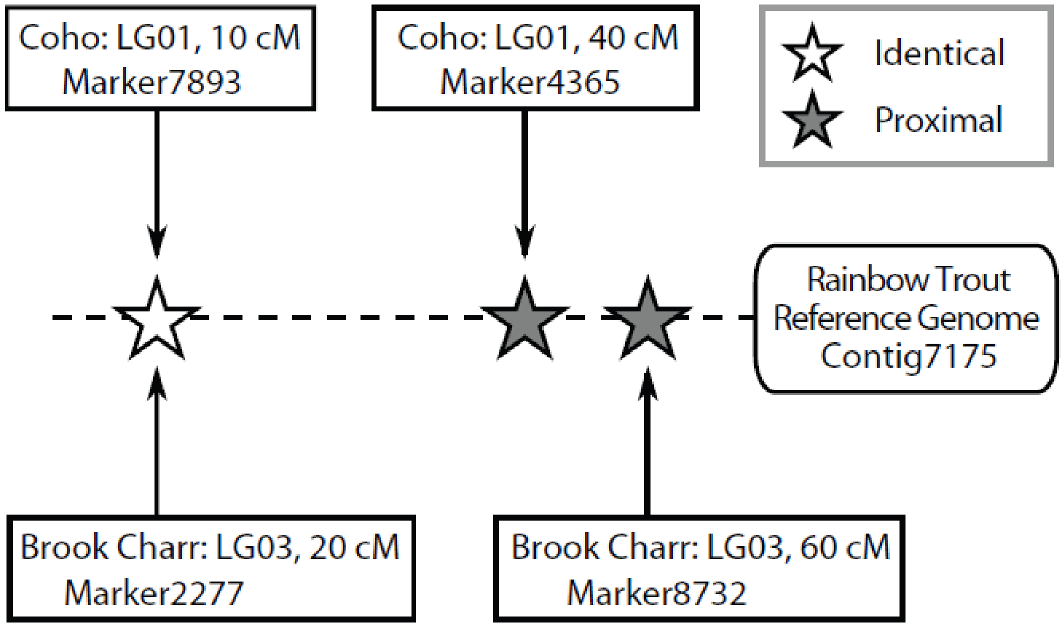
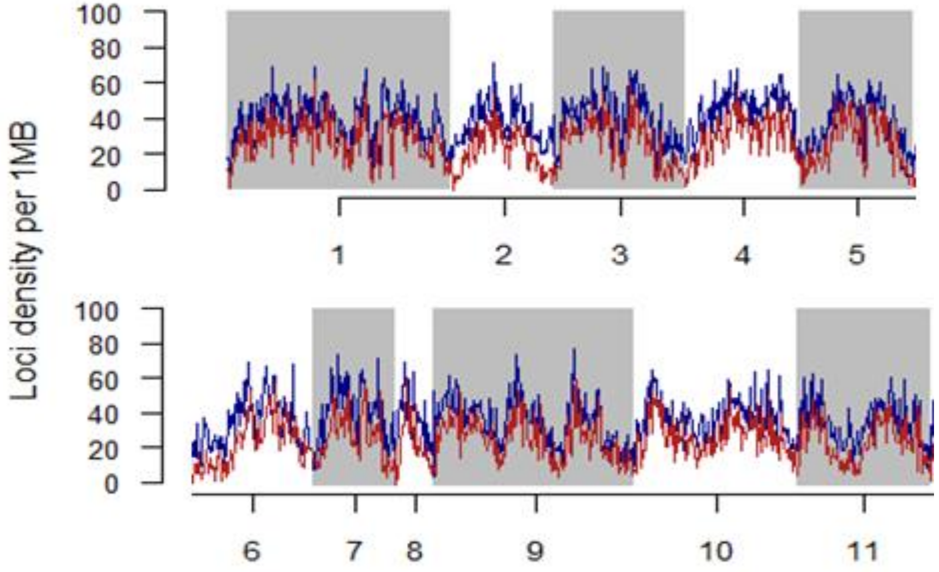
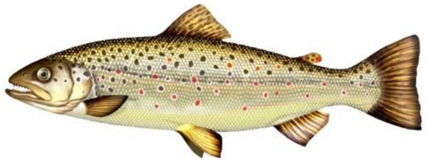


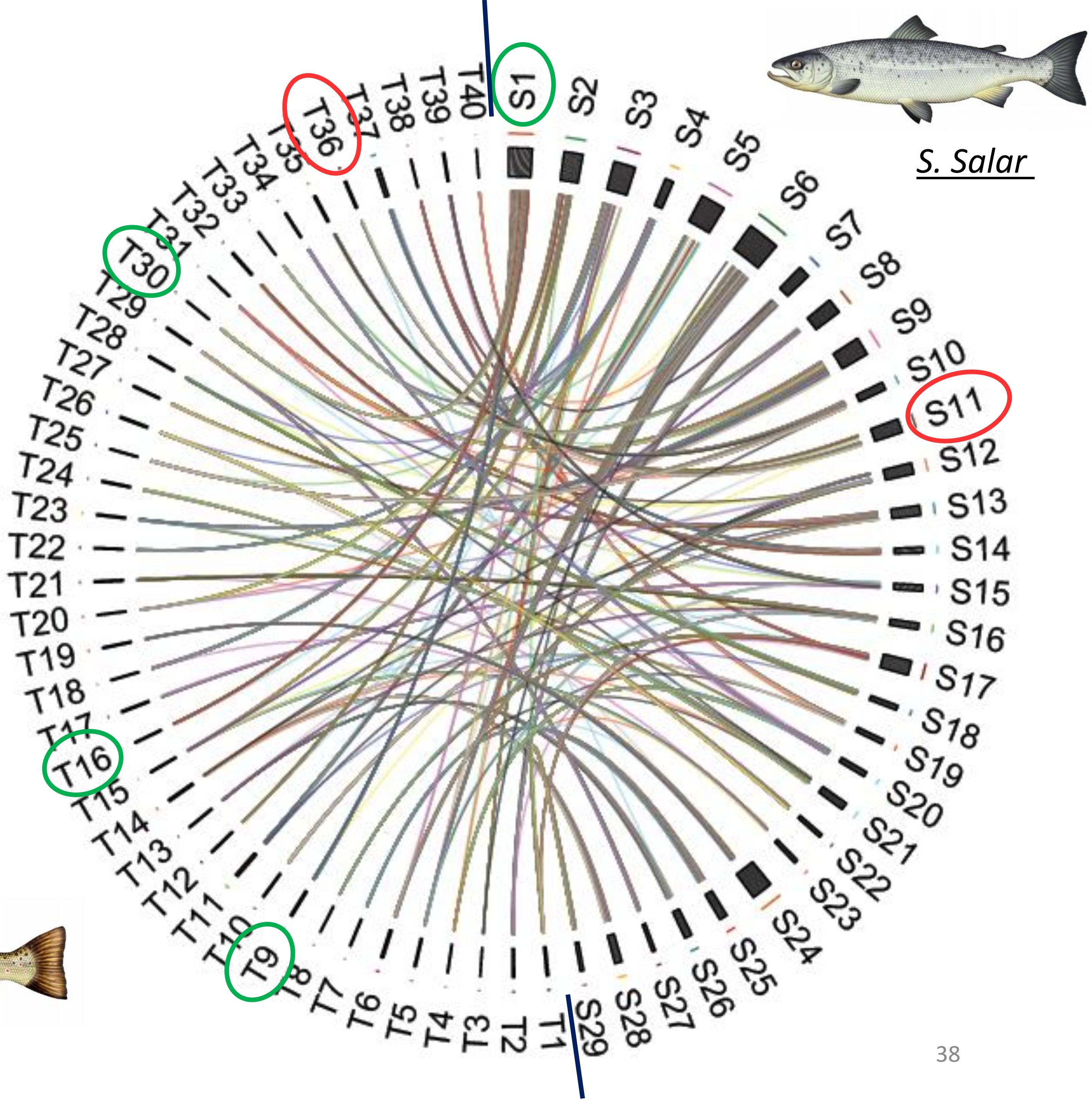
Fig. 1. Schematic of MapComp using a reference genome to pair markers. MapComp compares genetic maps from two different species by mapping marker sequences against a reference genome, then retaining high quality mappings that only hit against one locus in the genome. Markers from each species are paired if they hit against the same contig/scaffold by taking the closest two markers together as each pair. Each marker is paired without replacement, and so any other marker that was second closest to the now-paired marker is discarded. This method captures identical markers (white star in image) and non-identical markers (grey stars). Finally, the linkage group and cM position of each marker is plotted in an Oxford grid. Note that the marker names and contig ID in the schematic are for demonstration purposes only and do not reflect actual pairings.

Syntenies between *S. salar* and *trutta*

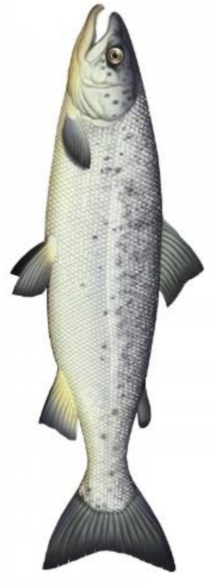




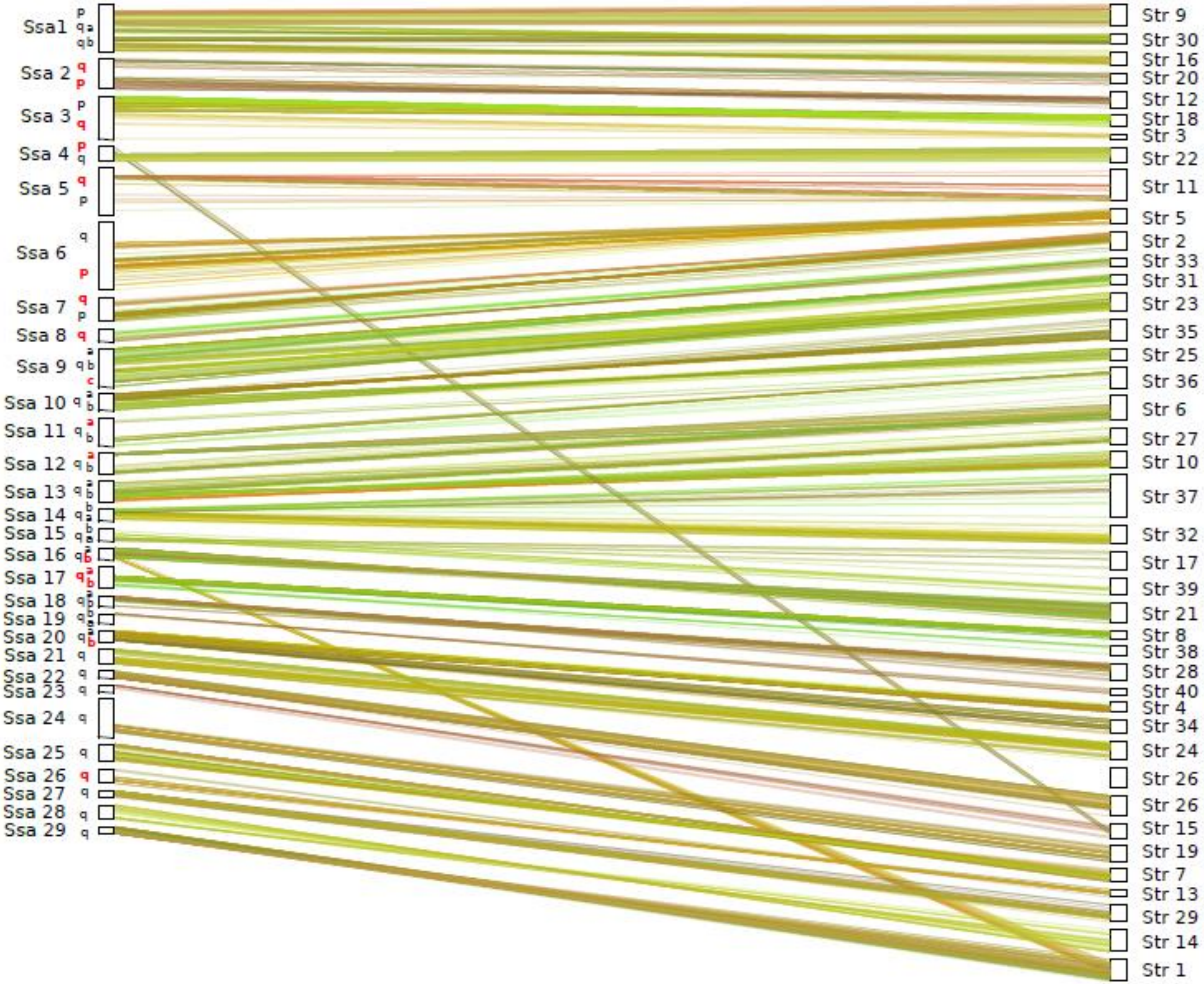
S. trutta



S. Salar



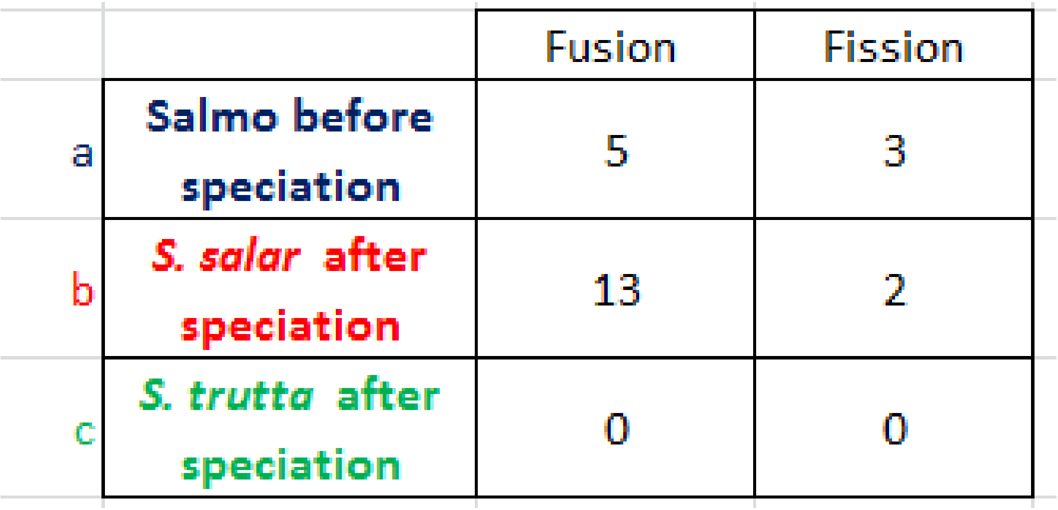
S. Salar



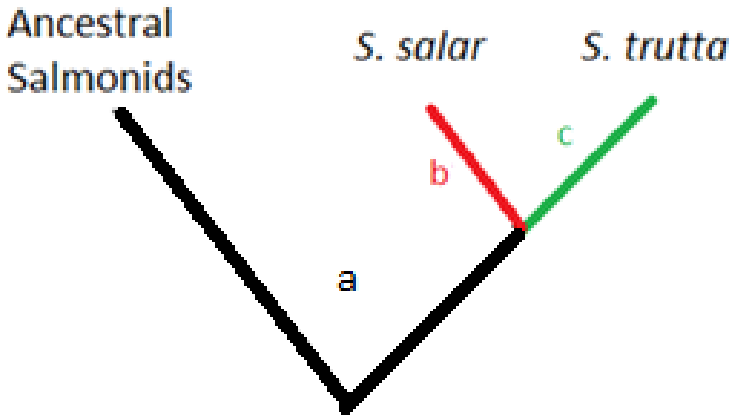
S. trutta

Tsai et al., 2016
Lien et al., 2016

Chromosomal rearrangements



		Fusion	Fission
a	**Salmo before speciation**	5	3
b	***S. salar*** **after speciation**	13	2
c	***S. trutta*** **after speciation**	0	0



Sutherland et al., 2016

Nucleotide divergence between *S. salar* and *trutta*

$$\boxed{d = \pi_b - \pi_w}$$

π_b : nucleotide diversity between *S. salar* and *S. trutta*

π_w : nucleotide diversity within *S. trutta*

Nucleotide divergence between *S. salar* and *trutta*

$$\boxed{d = \pi_b - \pi_w}$$

π_b : nucleotide diversity between *S. salar* and *S. trutta*

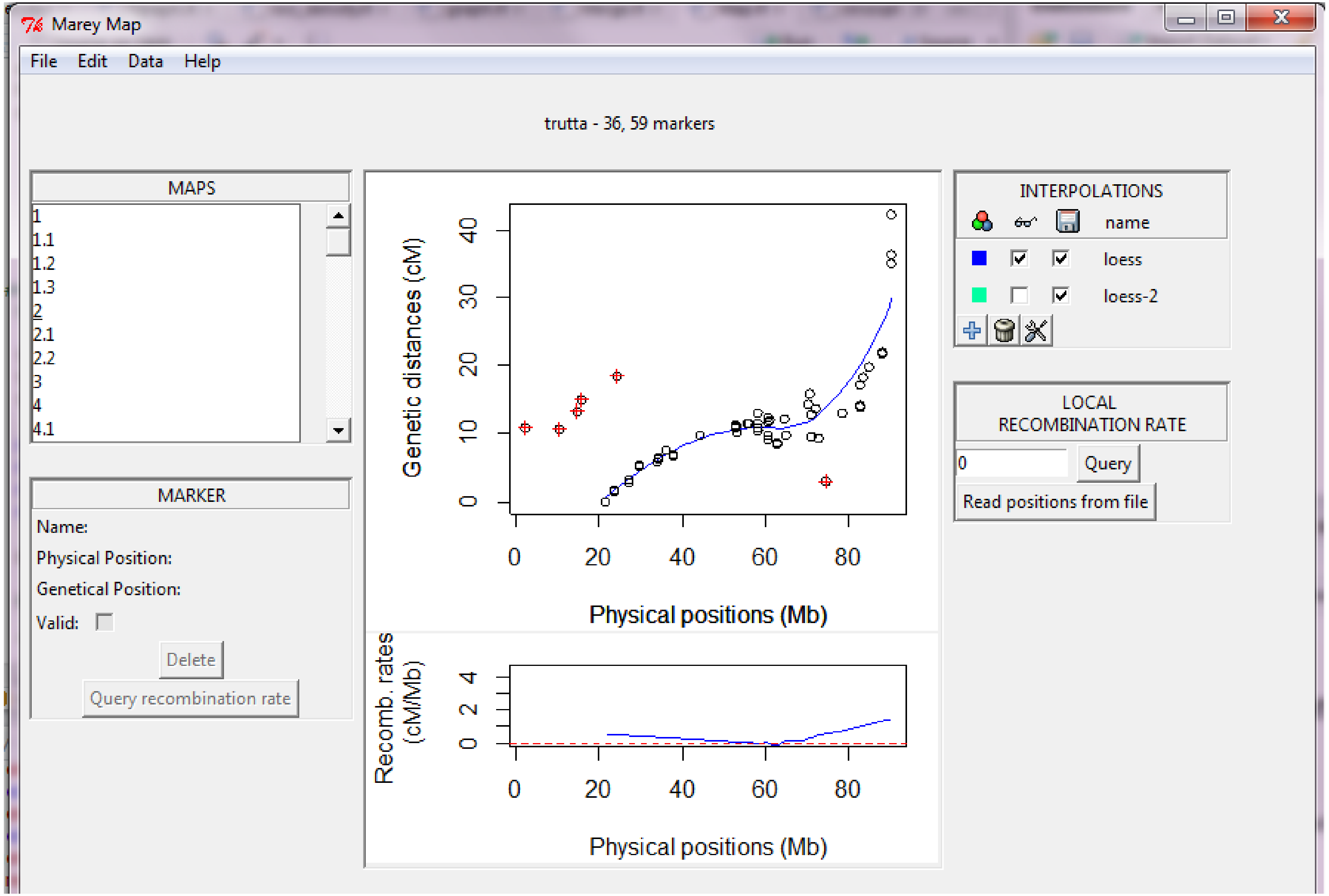
π_w : nucleotide diversity within *S. trutta*

$$d = 0.02285 - 0.0041$$

$$\boxed{\mathbf{d = 1.87\%}}$$

5.94% (Bernatchez *et al.*, 1992)

Estimate of recombination rate in the brown trout genome



trutta - 36, 59 markers

MAPS

- 1
- 1.1
- 1.2
- 1.3
- 2
- 2.1
- 2.2
- 3
- 4
- 4.1

MARKER

Name:
Physical Position:
Genetical Position:
Valid: ☐

INTERPOLATIONS

			name
🟦	☑	☑	loess
🟩	☐	☑	loess-2

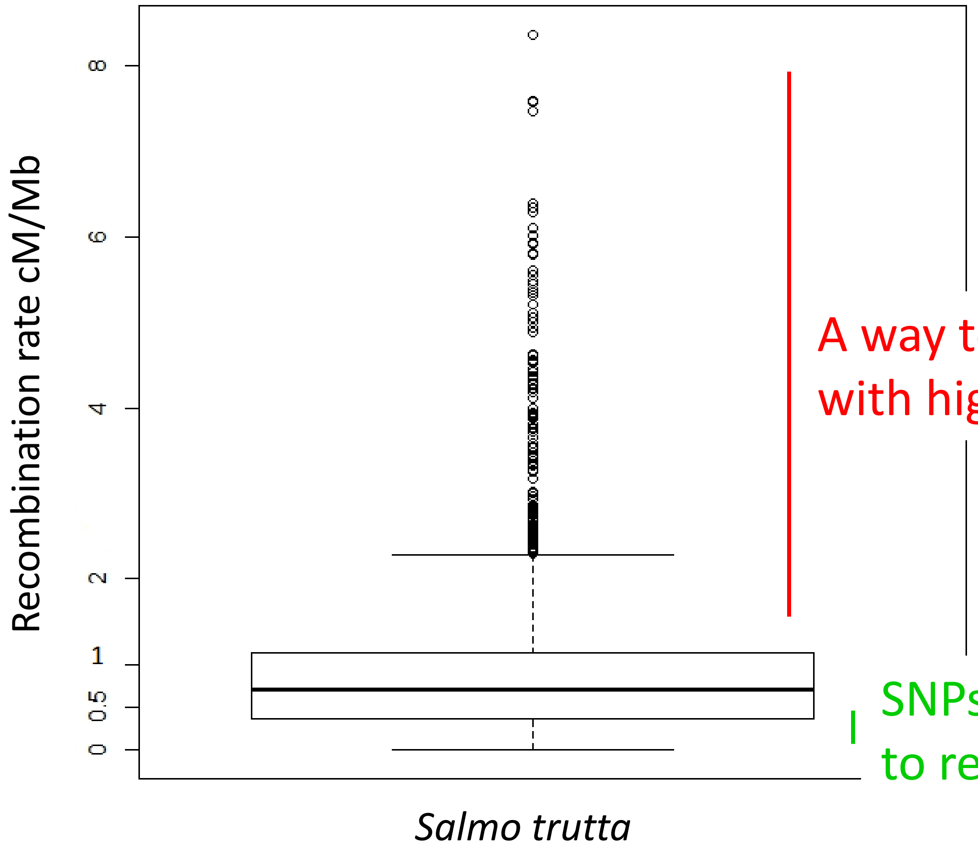
LOCAL RECOMBINATION RATE

0

Rezvoy *et al*. 2007

Estimate of recombination rate in the brown trout genome

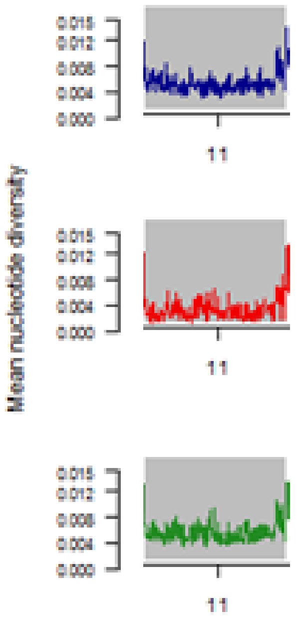
Total mean recombination rate = 0.88 cM/Mb

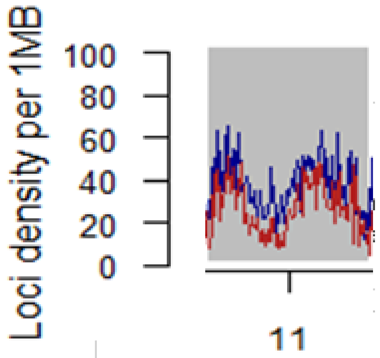


A way to identify SNPs associated with high recombination rate

SNPs associated
to regions with low recombination rate

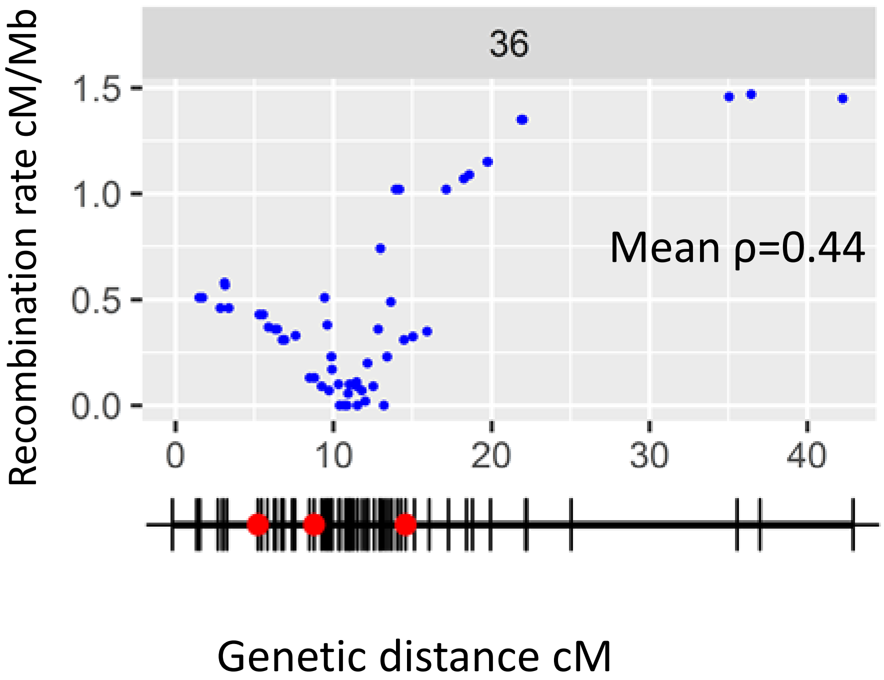
Recombination rate estimation





LG 11

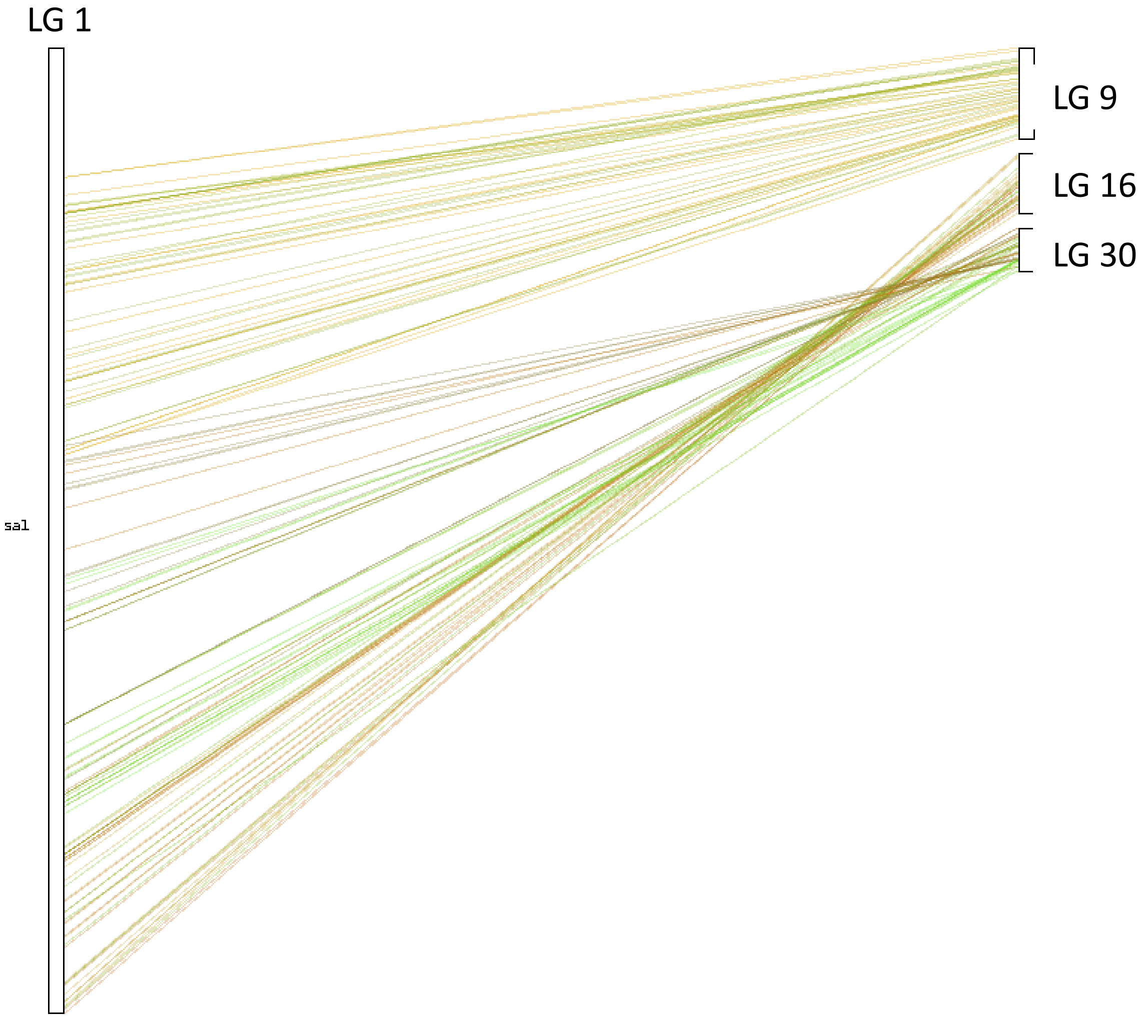
S. Salar



LG 36

S. trutta

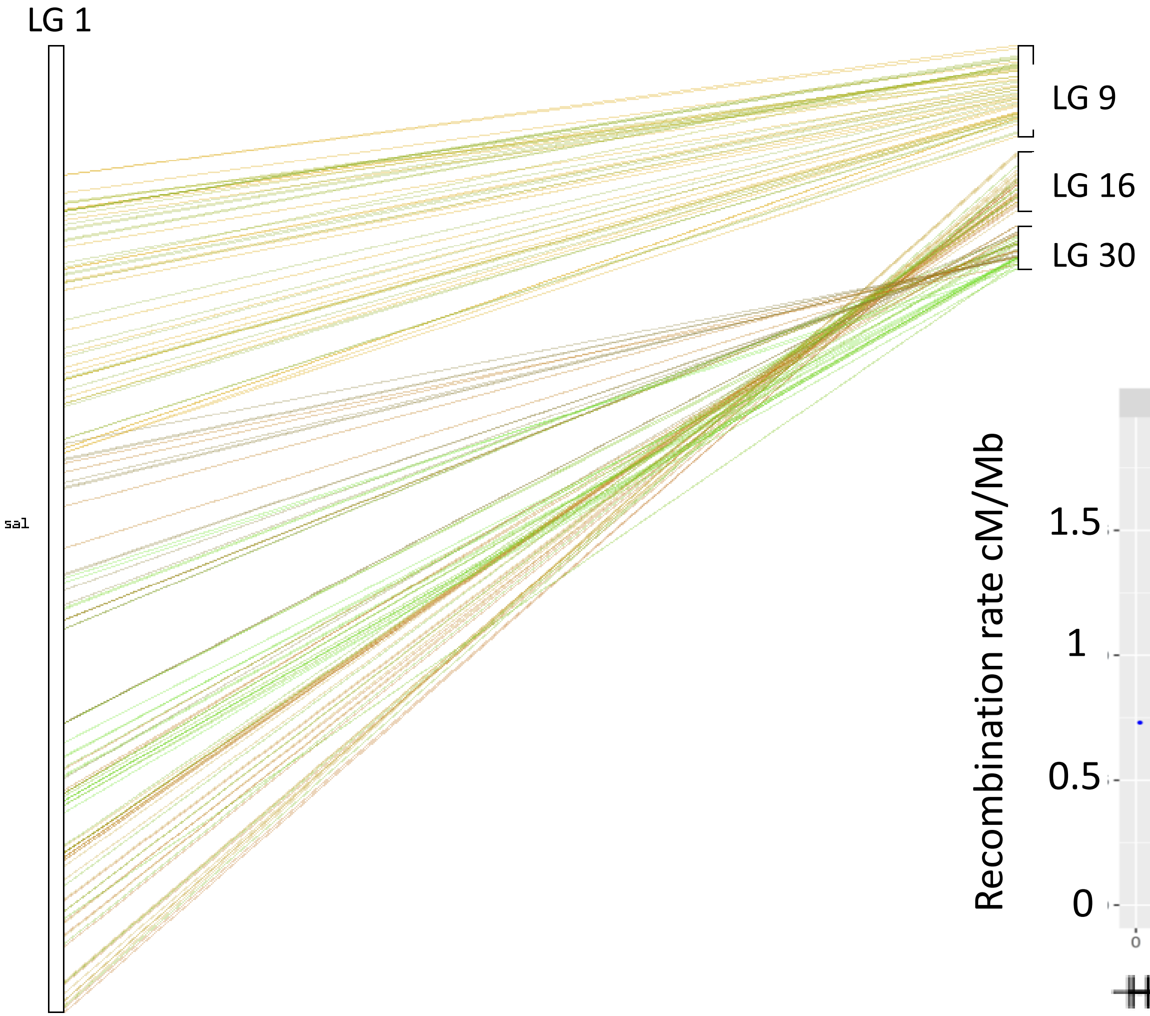
Recombination rate estimation

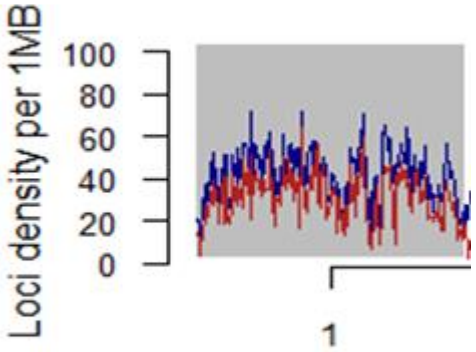


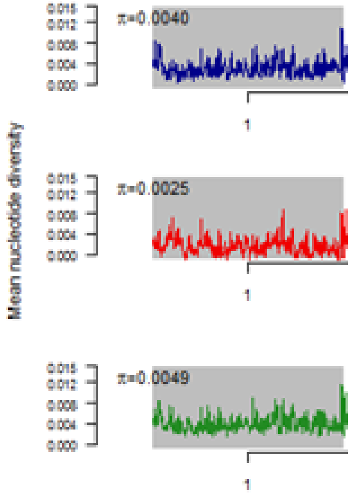
Recombination rate estimation

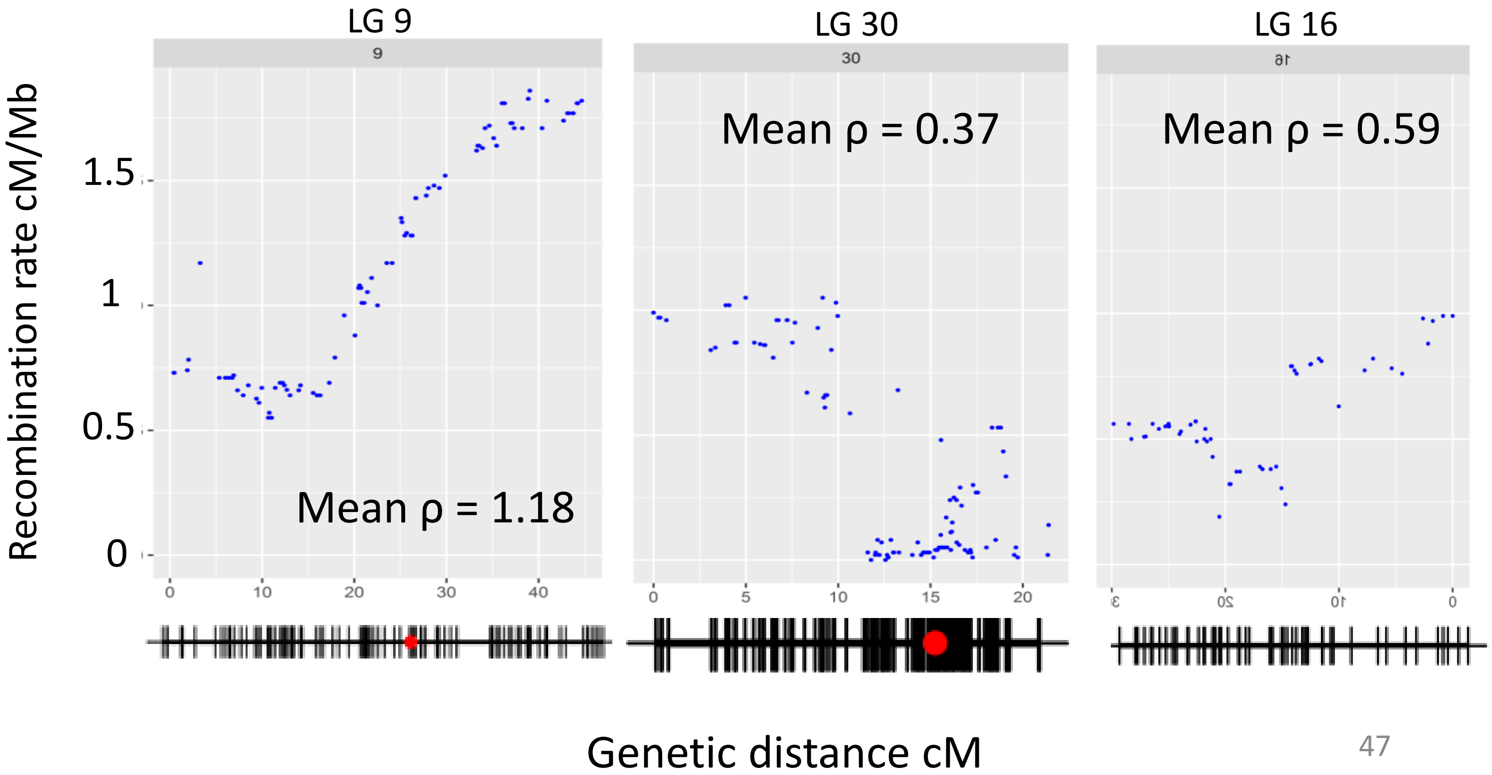
S. Salar

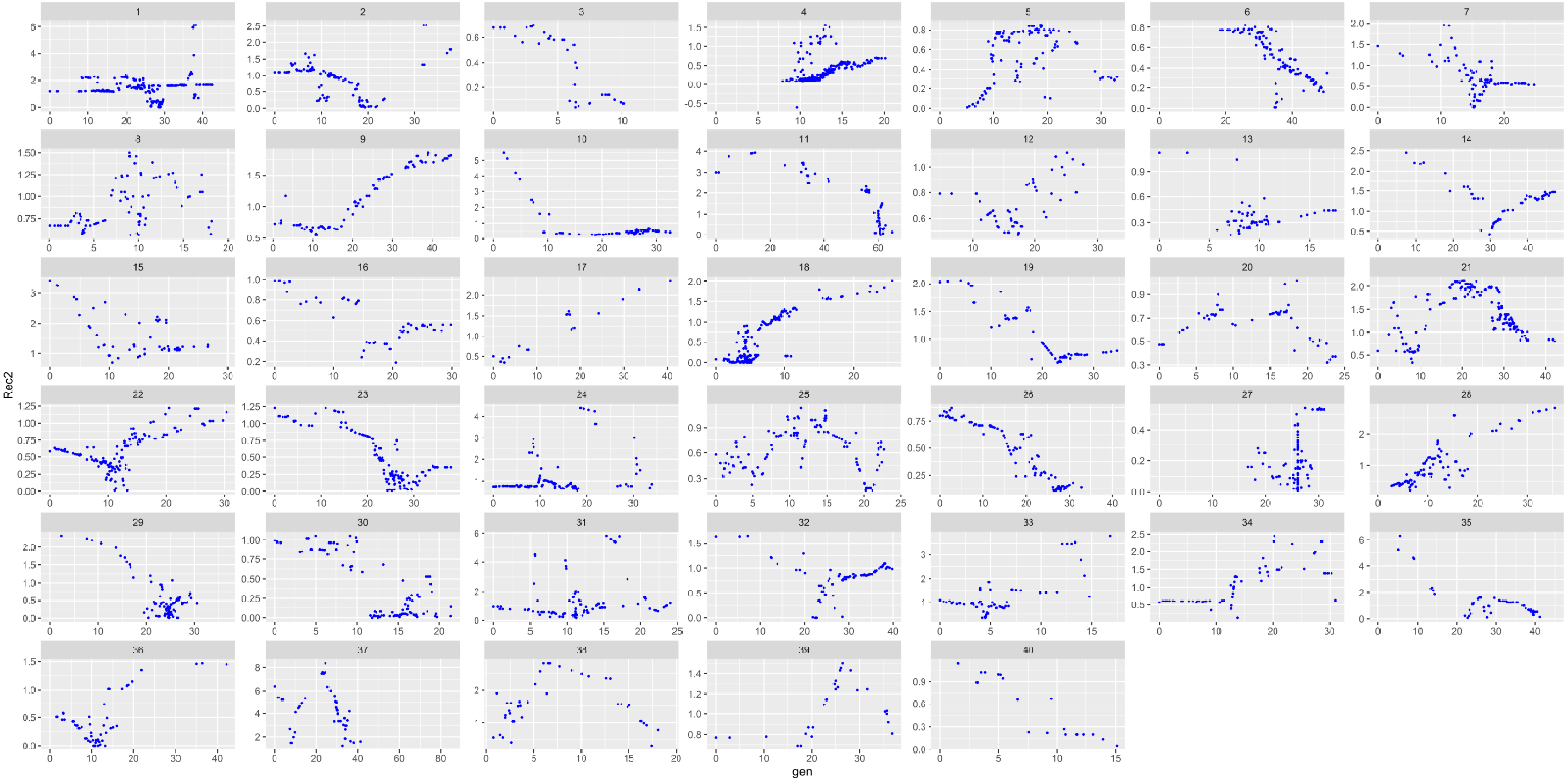
S. trutta



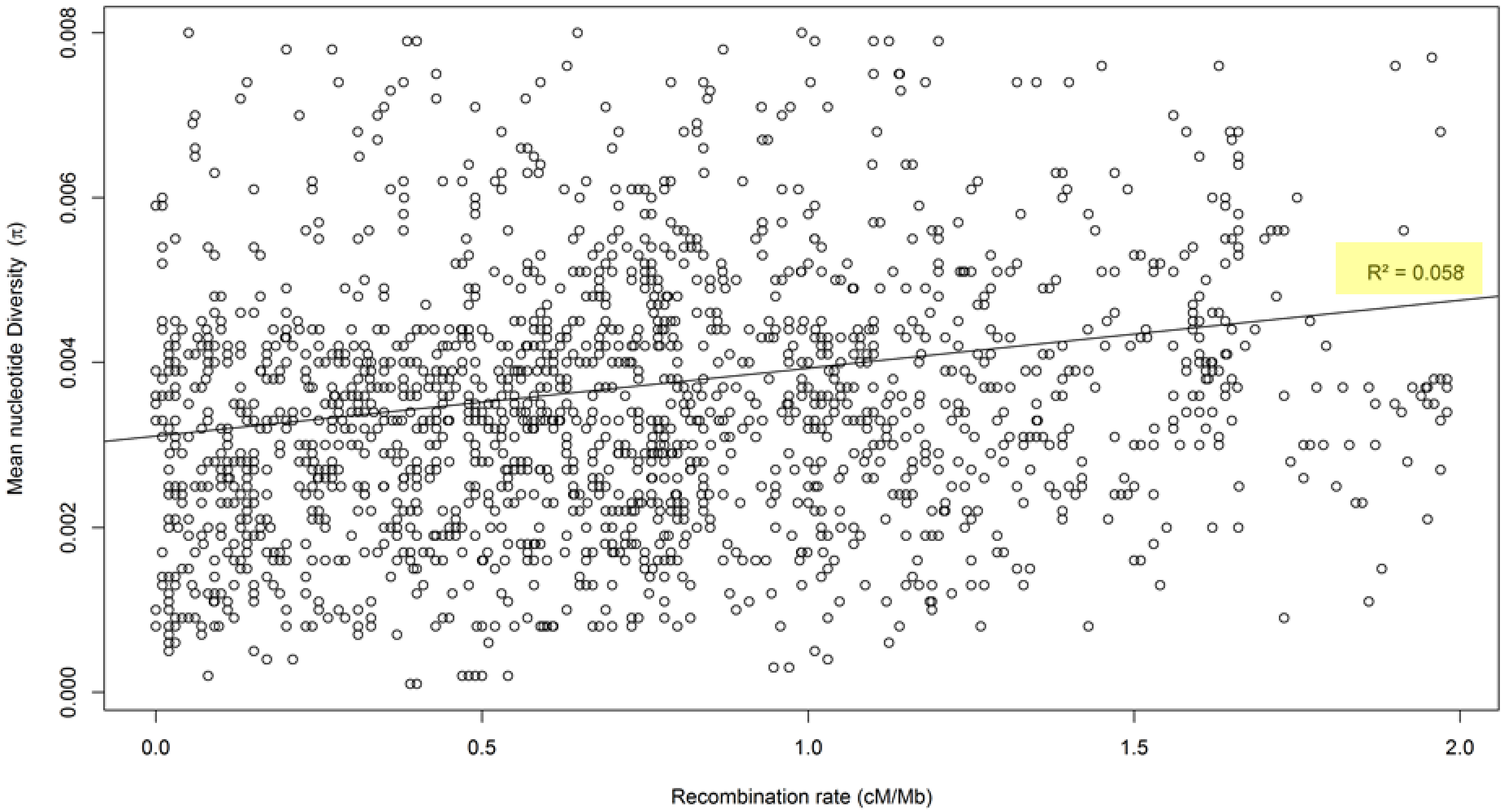







1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
Rec2
gen

Correlation between the nucleotide diversity and the recombination rate



$P < 0.001$

Part 2

Concluding remarks

→ **Strong (and expected…) synteny** between *S. salar* and *S. trutta*

→ A improved linkage map is available providing information on **chromosomal rearrangements**:

translocation, fissions, Robertsonian rearrangements between *S. salar* and *S. trutta*

→ **Positive correlation** between the **recombination rate** and the **nucleotidic diversity**

→ The estimation of the (local) **recombination rate** is accessible in *S. trutta*

→ Does the **recombination** landscape **affect local introgression rate** along the genome ?

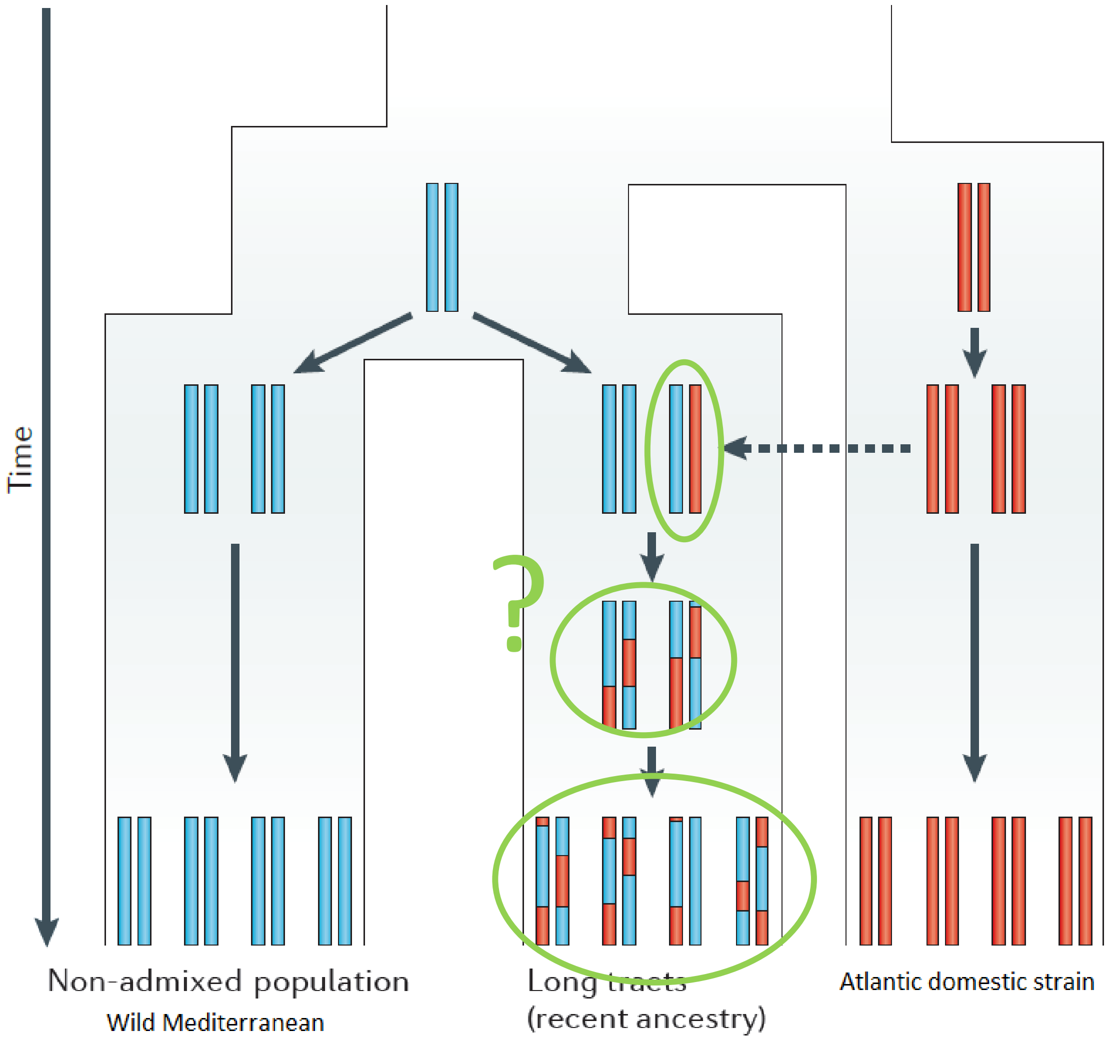
Part 3

→ To identify admixed individuals in the wild:

3 rivers : Mare N= 14
Orb N= 23
Gravezon N= 45

N = 82

(Dom ATL N= 61
Dom Med N=41)



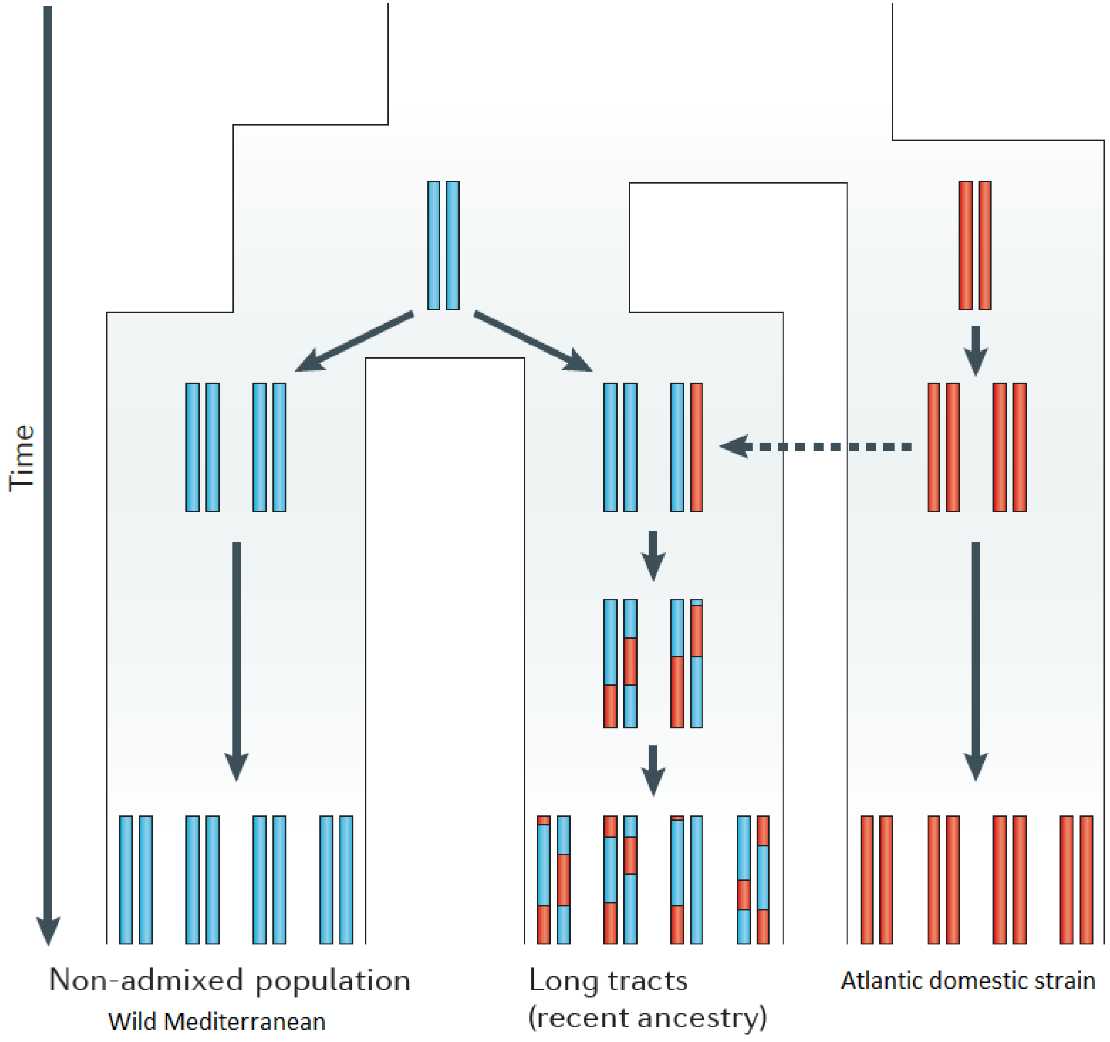
Racimo et al. (2015)

Part 3

→ To **identify introgressed haplotypes**

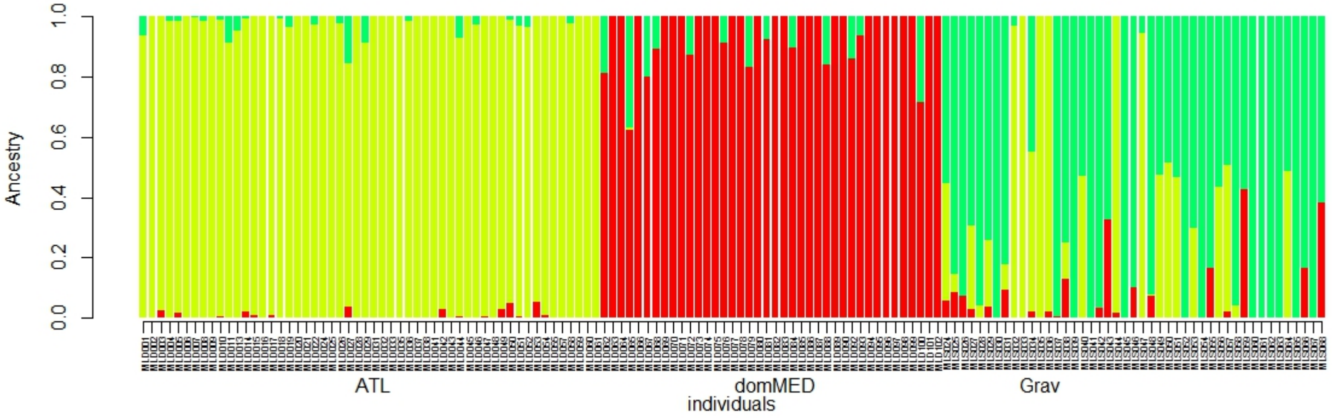
→To use the **introgressed haplotype distribution size** as a proxy of the timing of the introgression

(shorter haplotypes if more generations/recombination events)

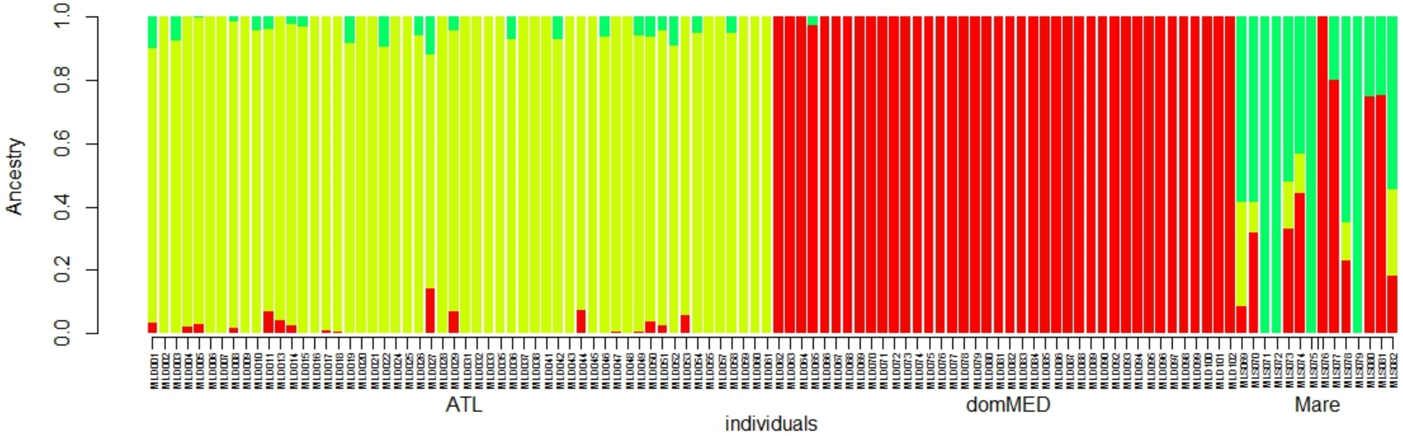


Identification of “pure” individuals for reference with Admixture

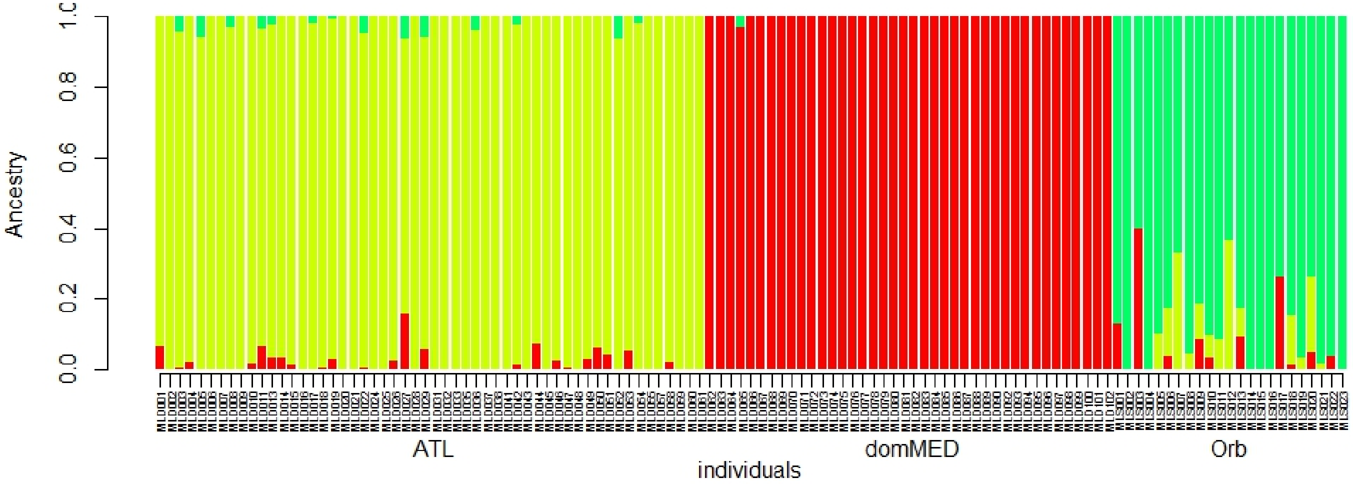
Grav



Mare



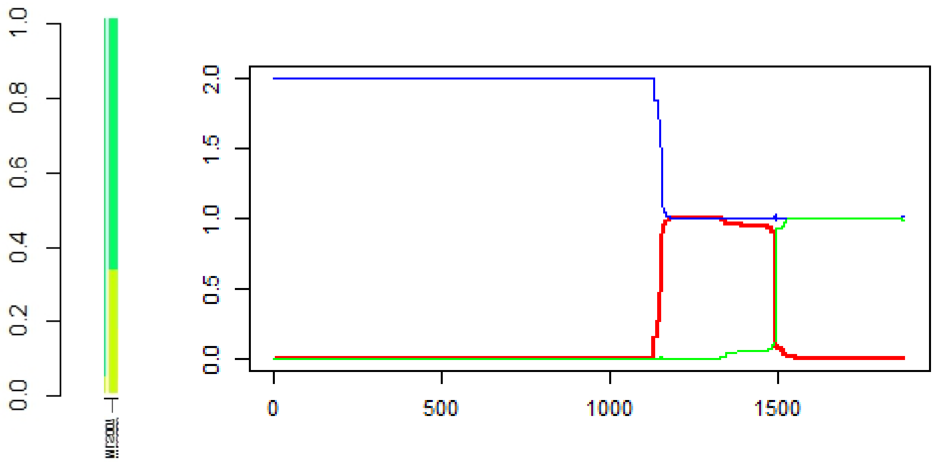
Orb

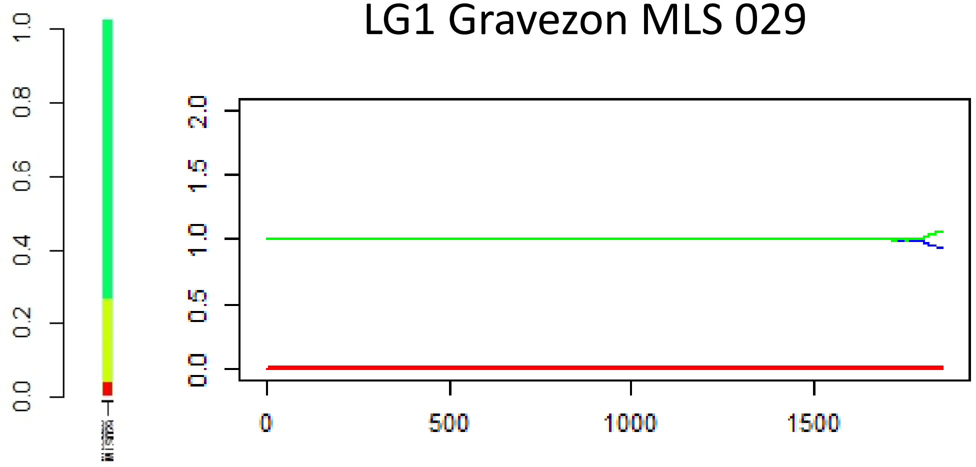


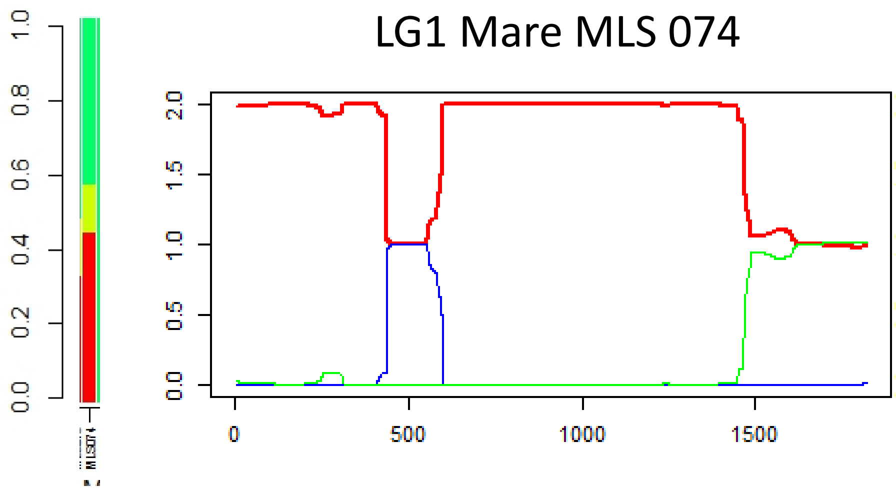
ELAI: Efficient Local Ancestry Inference

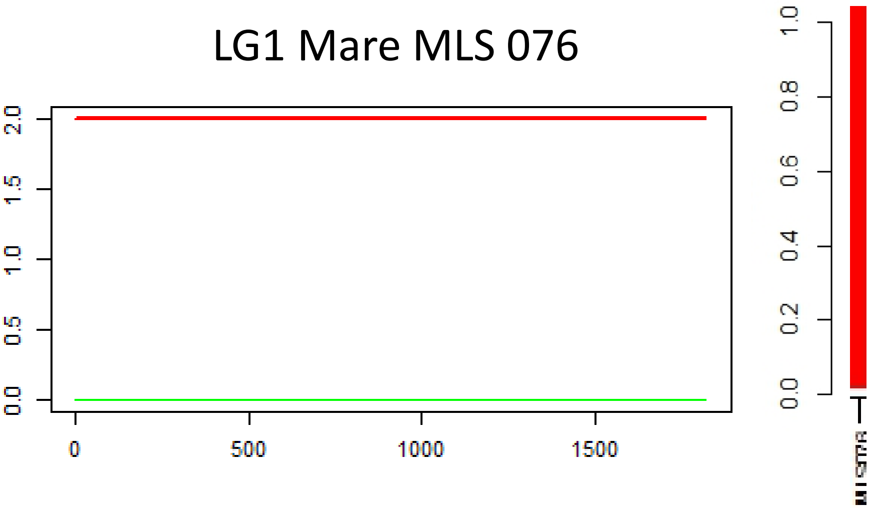
Guan Y., 2014

LG1 Orb MLS 007



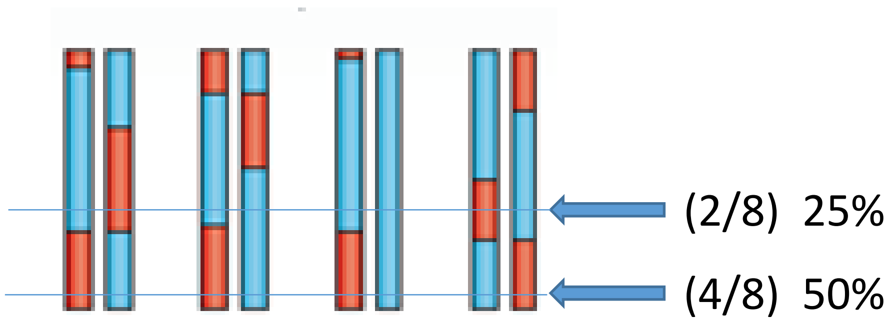

LG1 Gravezon MLS 029

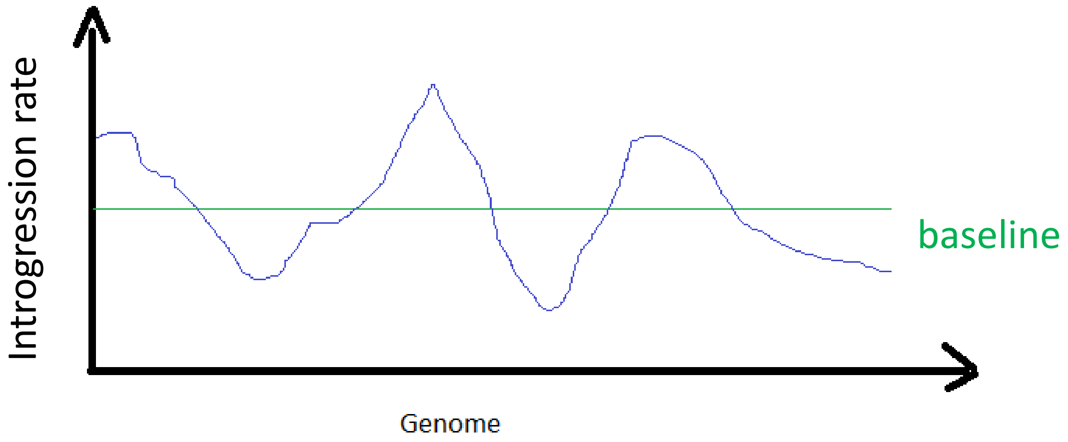

LG1 Mare MLS 074


LG1 Mare MLS 076

Part 3

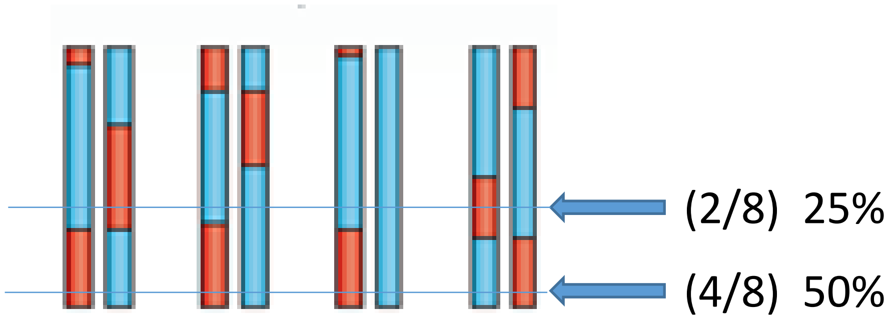
→ To determine the introgression rate along the genome

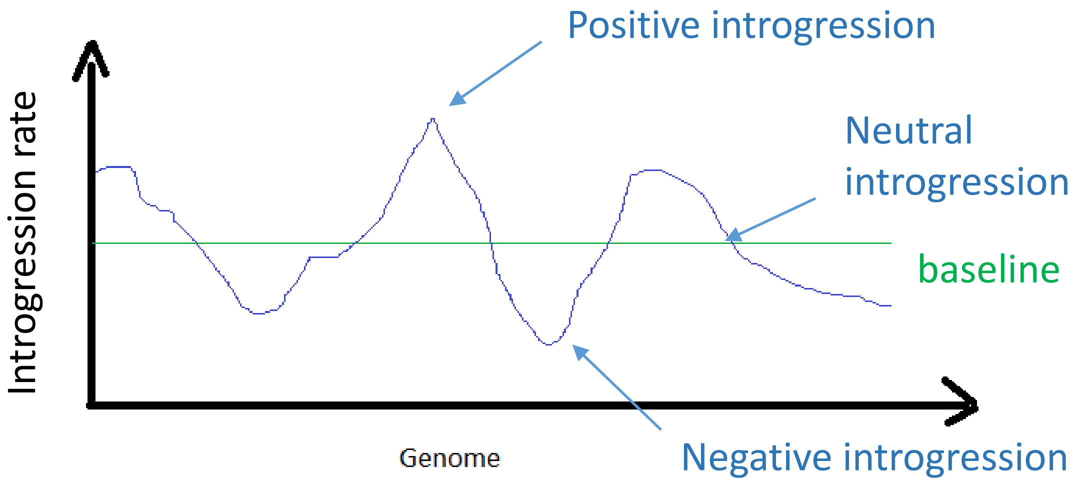




Part 3

→ To determine the introgression rate along the genome





→To find signatures of positive or negative introgression positive introgression

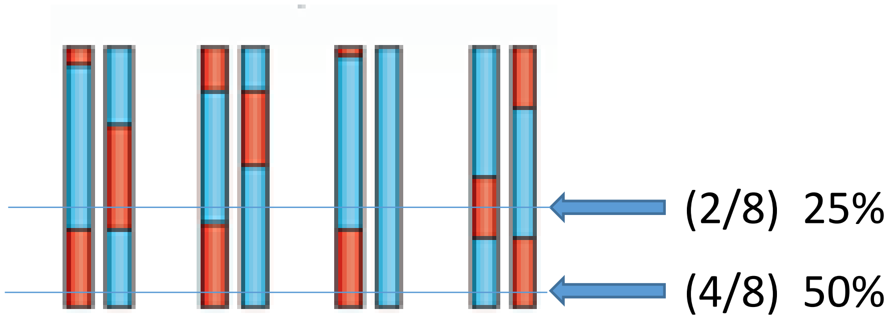
- By masking the effect of deleterious alleles
- Adaptive introgression

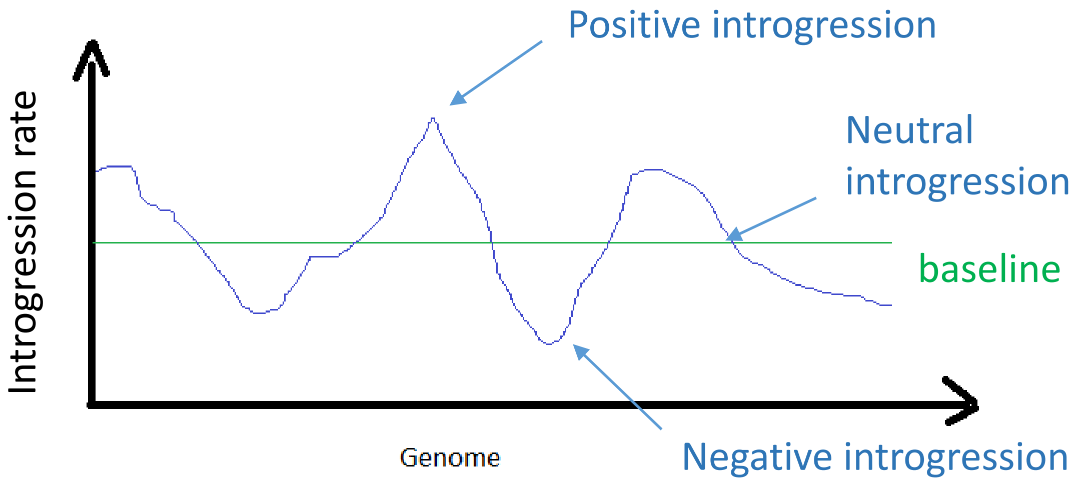
negative introgression

- Elimination of non adaptive domestic alleles
- Incompatibilities between ATL and MED lineages

Part 3

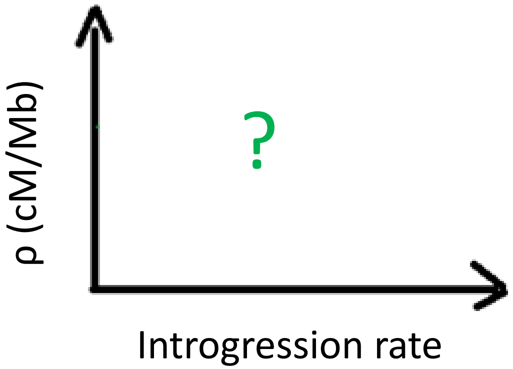
→ To determine the introgression rate along the genome





→To find signatures of positive or negative introgression

→To associate the recombination landscape to the introgression rate



Thank you !

Eric Ravel

FÉDÉRATION DÉPARTEMENTALE PÊCHE 34

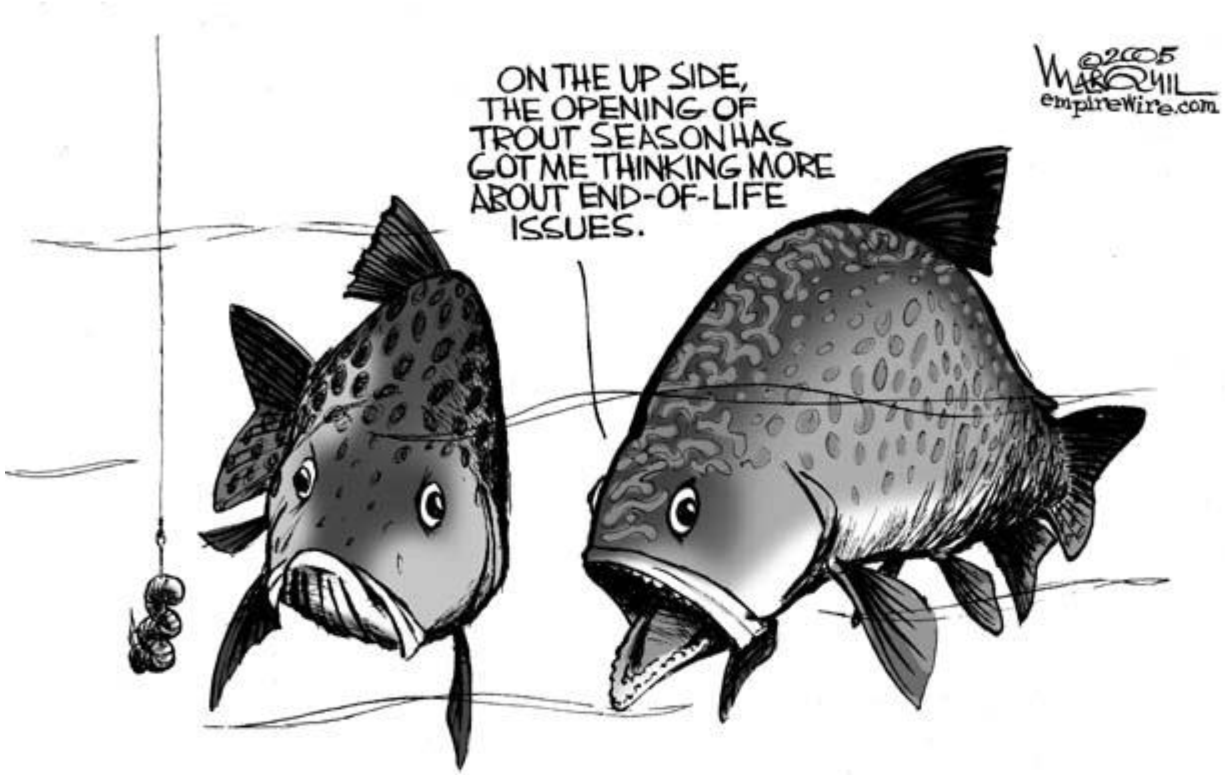
CeMEB

Patrick Berrebi

UNIVERSITÉ MONTPELLIER

Bruno Guinand

isem
Institut des Sciences de l'Evolution-Montpellier



Pierre-Alexandre Gagnaire

Juliette Pouzadoux

GÉNOTYPAGE SÉQUENÇAGE

Erick desmarais

Montpellier MGX GenomiX

Marine Rohmer

MBB

Julien Veysier, Khalid Belkir, Remy Dernat

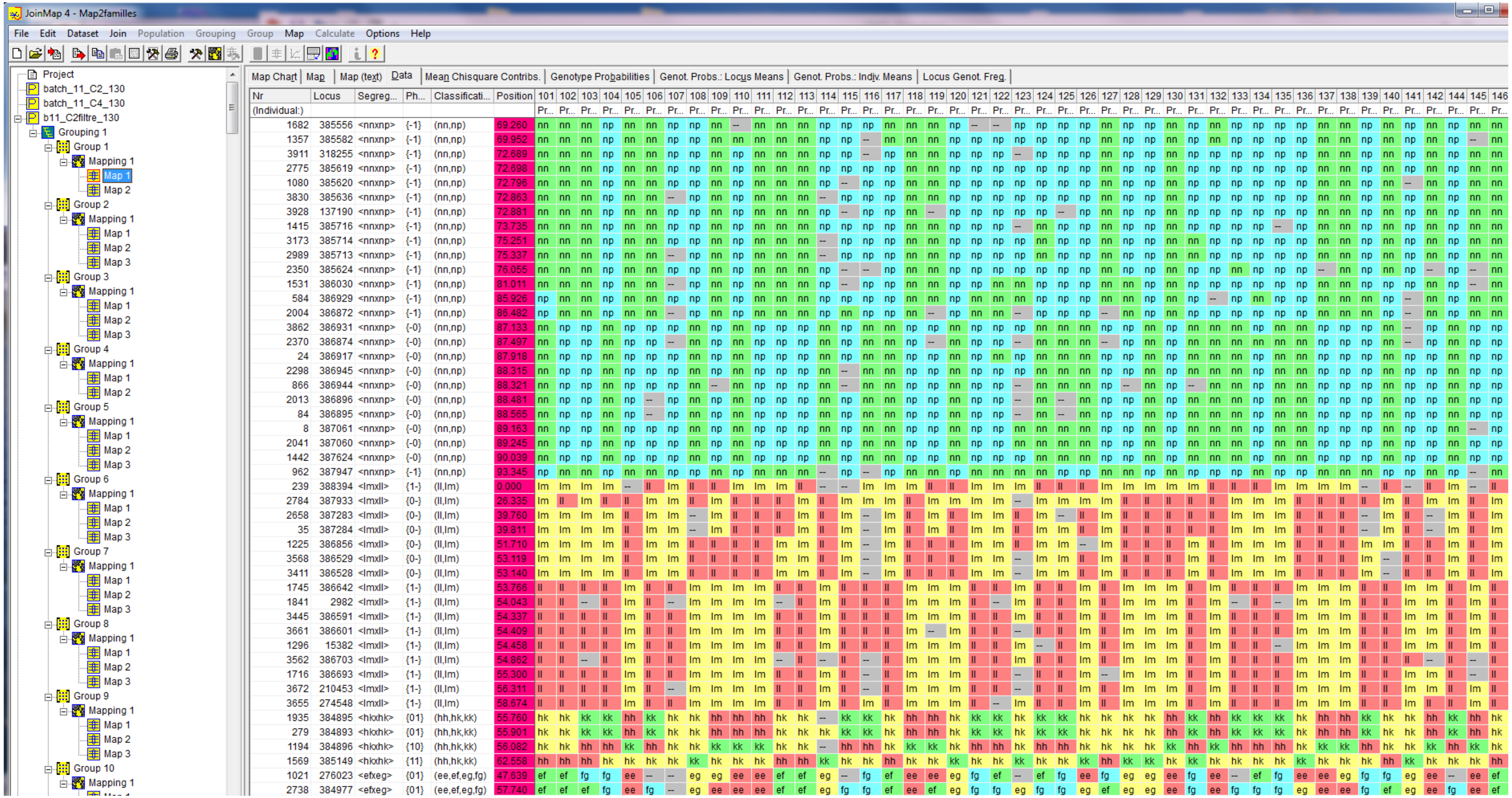
→ 3 Metacentics

→32 Acrocentrics

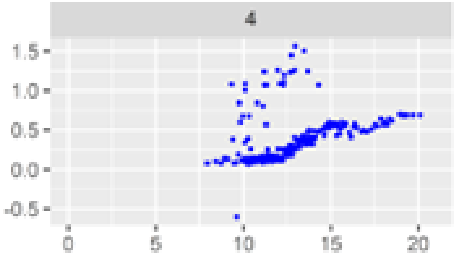
→5 undeterminde

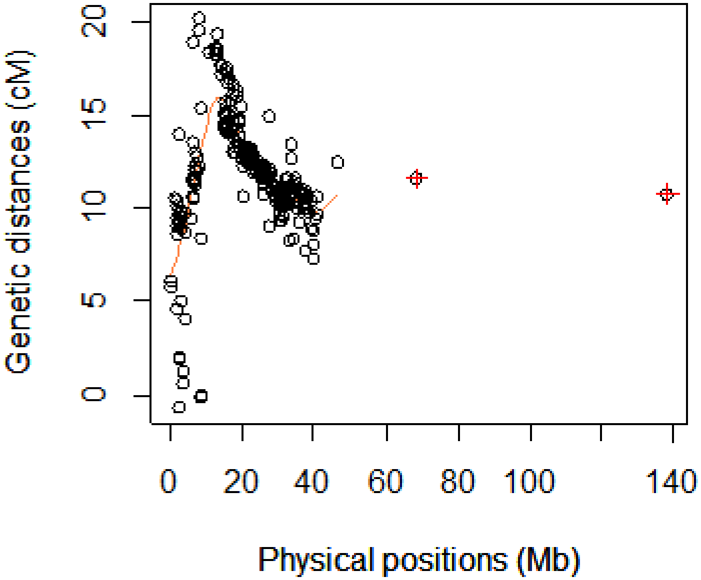
NF = 86 -96
Expected NF (karyotype studies) = 96-104

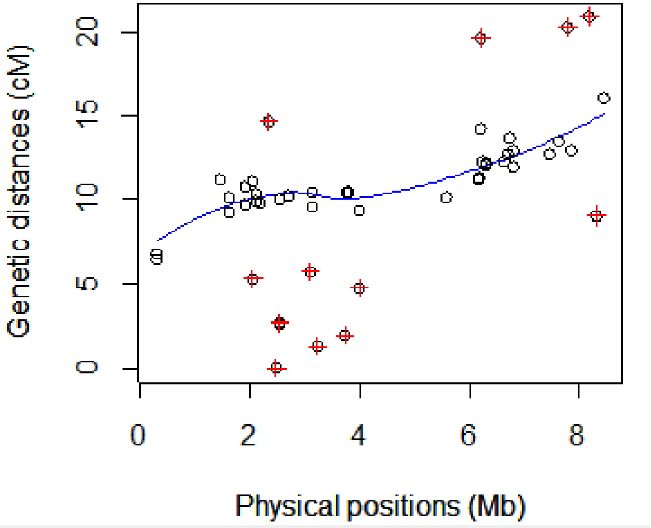
Joinmap 4.0

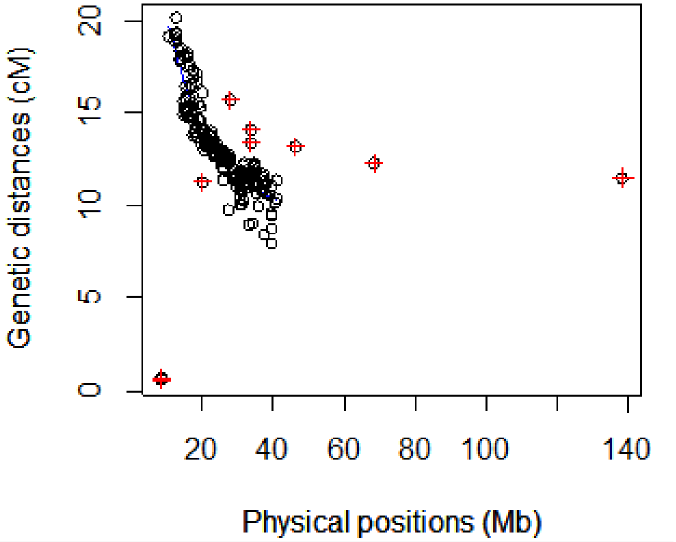


Nr	Locus	Segreg...	Ph...	Classificati...	Position	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	144	145	146
(Individual:)						Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...	Pr...
1682	385556	<nnxnp>	{-1}	(nn,np)	69.260	nn	nn	nn	np	nn	nn	np	np	nn	--	nn	nn	nn	np	np	np	nn	nn	nn	np	--	--	np	np	np	np	nn	np	np	nn	np	nn	np	np	np	np	nn	nn	np	nn	np	nn	np	nn	nn
1357	385582	<nnxnp>	{-1}	(nn,np)	69.952	nn	nn	nn	np	nn	nn	np	np	nn	nn	nn	nn	nn	np	np	--	nn	nn	nn	np	np	np	np	np	np	np	nn	np	np	nn	np	nn	np	np	np	np	nn	nn	np	nn	np	nn	np	--	nn
3911	318255	<nnxnp>	{-1}	(nn,np)	72.689	nn	nn	nn	np	nn	nn	np	np	nn	np	nn	nn	nn	np	np	--	np	nn	nn	np	np	np	--	np	np	np	nn	np	np	nn	np	np	np	np	np	np	nn	nn	np	nn	np	nn	np	nn	nn
2775	385619	<nnxnp>	{-1}	(nn,np)	72.698	nn	nn	nn	np	nn	nn	np	np	nn	np	nn	nn	nn	np	np	np	np	nn	nn	np	np	np	np	np	np	np	nn	np	np	nn	np	np	np	np	np	np	nn	nn	np	nn	np	nn	np	nn	nn
1080	385620	<nnxnp>	{-1}	(nn,np)	72.796	nn	nn	nn	np	nn	nn	np	np	nn	np	nn	nn	nn	np	--	np	np	nn	nn	np	np	np	np	np	np	np	nn	np	np	nn	np	np	np	np	np	np	nn	nn	np	nn	--	nn	np	nn	nn
3830	385636	<nnxnp>	{-1}	(nn,np)	72.863	nn	nn	nn	np	nn	nn	--	np	nn	np	nn	nn	nn	--	np	np	np	nn	nn	np	np	np	np	np	np	np	nn	np	np	nn	np	np	np	np	np	np	nn	nn	np	nn	np	nn	np	nn	nn
3928	137190	<nnxnp>	{-1}	(nn,np)	72.881	nn	nn	nn	np	nn	nn	np	np	nn	np	nn	nn	nn	np	--	np	np	nn	--	np	np	np	np	np	--	np	nn	np	np	nn	np	np	np	np	np	np	nn	nn	np	nn	np	nn	np	nn	nn
1415	385716	<nnxnp>	{-1}	(nn,np)	73.735	nn	nn	nn	np	nn	nn	np	np	nn	np	nn	nn	nn	np	np	np	np	nn	nn	np	np	np	--	nn	np	np	nn	np	np	nn	np	np	np	np	--	np	nn	nn	np	nn	np	nn	np	nn	nn
3173	385714	<nnxnp>	{-1}	(nn,np)	75.251	nn	nn	nn	np	nn	nn	np	np	nn	np	nn	nn	nn	--	np	np	np	nn	nn	np	np	np	np	nn	np	np	nn	np	np	nn	nn	np	np	np	np	np	nn	nn	np	nn	np	nn	np	nn	nn
2989	385713	<nnxnp>	{-1}	(nn,np)	75.337	nn	nn	nn	np	nn	nn	--	np	nn	np	nn	nn	nn	--	np	np	np	nn	nn	np	np	np	np	nn	np	np	nn	np	np	nn	nn	np	np	np	np	np	nn	nn	np	nn	np	nn	np	nn	nn
2350	385624	<nnxnp>	{-1}	(nn,np)	76.055	nn	nn	nn	np	nn	nn	np	np	nn	np	nn	nn	nn	np	--	--	np	nn	nn	np	np	np	np	np	np	np	nn	np	np	nn	np	np	nn	np	np	np	--	nn	np	nn	np	--	np	--	nn
1531	386030	<nnxnp>	{-1}	(nn,np)	81.011	nn	nn	nn	np	nn	nn	--	np	nn	np	nn	nn	nn	np	--	np	np	nn	nn	np	np	nn	nn	np	np	np	nn	nn	np	nn	np	np	np	np	np	np	nn	nn	np	np	np	nn	np	--	nn
584	386929	<nnxnp>	{-1}	(nn,np)	85.926	np	nn	nn	np	nn	nn	np	np	nn	np	nn	nn	nn	np	np	np	np	nn	nn	np	nn	nn	nn	np	np	np	nn	nn	np	nn	np	--	np	nn	np	np	nn	nn	nn	np	--	nn	np	nn	nn
2004	386872	<nnxnp>	{-1}	(nn,np)	86.482	np	nn	nn	np	nn	nn	--	np	nn	np	nn	nn	nn	np	nn	np	np	nn	--	np	nn	nn	--	np	np	np	--	nn	np	nn	np	np	np	np	np	np	nn	nn	nn	np	--	nn	np	nn	nn
3862	386931	<nnxnp>	{-0}	(nn,np)	87.133	nn	np	np	nn	np	np	np	nn	np	nn	np	np	np	nn	np	nn	nn	np	np	nn	np	np	np	nn	nn	nn	np	np	nn	np	nn	nn	nn	np	nn	nn	np	np	np	nn	--	np	nn	np	np
2370	386874	<nnxnp>	{-0}	(nn,np)	87.497	nn	np	np	nn	np	np	--	nn	np	nn	np	np	np	nn	np	nn	nn	np	--	nn	np	np	--	nn	nn	nn	--	np	nn	np	nn	nn	nn	nn	nn	nn	np	np	np	nn	--	np	nn	np	np
24	386917	<nnxnp>	{-0}	(nn,np)	87.918	nn	np	np	nn	np	np	np	nn	np	nn	np	np	np	nn	np	nn	nn	np	np	nn	np	nn	np	nn	nn	nn	np	np	nn	np	nn	nn	nn	np	nn	nn	np	np	np	nn	np	np	nn	np	np
2298	386945	<nnxnp>	{-0}	(nn,np)	88.315	nn	np	np	nn	np	np	np	nn	np	nn	np	np	np	nn	--	nn	nn	np	np	nn	np	np	np	nn	nn	nn	np	np	nn	np	nn	nn	nn	np	nn	nn	np	np	np	nn	np	np	nn	np	np
866	386944	<nnxnp>	{-0}	(nn,np)	88.321	nn	np	np	nn	np	np	np	nn	--	nn	np	np	np	nn	--	nn	nn	np	np	nn	np	np	--	nn	nn	nn	np	--	nn	np	--	nn	nn	np	nn	nn	np	np	np	nn	np	np	nn	np	np
2013	386896	<nnxnp>	{-0}	(nn,np)	88.481	nn	np	np	nn	np	--	np	nn	np	nn	np	np	np	nn	np	nn	nn	np	np	nn	np	np	--	nn	--	nn	np	np	nn	np	nn	nn	nn	np	nn	nn	np	np	np	nn	np	np	nn	np	np
84	386895	<nnxnp>	{-0}	(nn,np)	88.565	nn	np	np	nn	np	--	np	nn	np	nn	np	np	np	nn	np	nn	nn	np	np	nn	np	np	--	nn	--	nn	np	np	nn	np	nn	nn	nn	np	nn	nn	np	np	np	nn	np	np	nn	np	np
8	387061	<nnxnp>	{-0}	(nn,np)	89.163	nn	np	np	nn	np	np	np	nn	np	nn	np	np	np	nn	np	nn	nn	np	np	nn	np	np	nn	nn	nn	nn	np	np	nn	np	nn	nn	nn	np	nn	nn	np	np	np	nn	np	np	nn	--	np
2041	387060	<nnxnp>	{-0}	(nn,np)	89.245	nn	np	np	nn	np	np	np	nn	np	nn	np	np	np	nn	np	nn	nn	np	np	nn	np	np	nn	nn	nn	nn	np	np	nn	np	nn	nn	nn	np	nn	nn	np	np	np	nn	np	np	nn	np	np
1442	387624	<nnxnp>	{-0}	(nn,np)	90.039	nn	np	np	nn	np	np	np	nn	np	nn	np	np	np	nn	np	nn	nn	np	np	nn	np	np	np	nn	nn	nn	np	np	nn	np	nn	nn	nn	np	nn	nn	np	np	np	nn	np	np	nn	np	np
962	387947	<nnxnp>	{-1}	(nn,np)	93.345	np	nn	nn	np	nn	nn	np	np	nn	np	nn	nn	nn	--	np	--	np	nn	nn	np	nn	nn	nn	nn	np	np	nn	nn	np	nn	np	np	np	nn	np	np	nn	nn	nn	np	np	nn	np	--	nn
239	388394	<lmxll>	{1-}	(ll,lm)	0.000	lm	lm	lm	lm	--	ll	lm	ll	ll	lm	lm	lm	ll	--	--	lm	lm	lm	ll	ll	lm	lm	lm	ll	ll	ll	lm	lm	lm	lm	lm	ll	ll	ll	lm	lm	lm	lm	--	ll	--	ll	lm	--	ll
2784	387933	<lmxll>	{0-}	(ll,lm)	26.335	lm	ll	lm	ll	ll	lm	lm	ll	lm	ll	ll	ll	lm	ll	lm	lm	lm	ll	lm	lm	lm	lm	--	lm	lm	lm	lm	ll	ll	ll	ll	ll	lm	lm	lm	ll	ll	ll	ll	lm	ll	lm	lm	ll	lm
2658	387283	<lmxll>	{0-}	(ll,lm)	39.760	lm	lm	lm	lm	ll	lm	lm	--	lm	ll	ll	ll	lm	ll	lm	--	lm	ll	lm	ll	lm	lm	ll	lm	--	ll	lm	ll	ll	ll	ll	ll	lm	lm	lm	ll	ll	ll	--	lm	ll	--	lm	ll	lm
35	387284	<lmxll>	{0-}	(ll,lm)	39.811	lm	lm	lm	lm	ll	lm	lm	--	lm	ll	ll	ll	lm	ll	lm	--	lm	ll	lm	ll	lm	lm	ll	lm	lm	ll	lm	ll	ll	ll	ll	ll	lm	lm	lm	ll	ll	ll	--	lm	ll	--	lm	ll	lm
1225	386856	<lmxll>	{0-}	(ll,lm)	51.710	lm	lm	lm	lm	ll	lm	lm	ll	ll	ll	ll	lm	lm	ll	lm	--	lm	ll	ll	ll	lm	lm	ll	lm	lm	--	lm	ll	ll	ll	lm	ll	lm	lm	lm	ll	ll	ll	lm	lm	ll	ll	lm	ll	lm
3568	386529	<lmxll>	{0-}	(ll,lm)	53.119	lm	lm	lm	lm	ll	lm	lm	ll	ll	ll	ll	lm	lm	ll	lm	--	lm	ll	ll	ll	lm	lm	--	lm	lm	ll	lm	ll	ll	ll	lm	ll	lm	lm	lm	ll	ll	ll	lm	--	ll	ll	lm	ll	lm
3411	386528	<lmxll>	{0-}	(ll,lm)	53.140	lm	lm	lm	lm	ll	lm	lm	ll	ll	ll	ll	lm	lm	ll	lm	--	lm	ll	ll	ll	lm	lm	--	lm	lm	ll	lm	ll	ll	ll	lm	ll	lm	lm	lm	ll	ll	ll	lm	--	ll	ll	lm	ll	lm
1745	386642	<lmxll>	{1-}	(ll,lm)	53.766	ll	ll	ll	ll	lm	ll	ll	lm	lm	lm	lm	ll	ll	lm	ll	ll	ll	lm	lm	lm	ll	ll	lm	ll	ll	lm	ll	lm	lm	lm	ll	lm	ll	ll	ll	lm	lm	lm	ll	ll	lm	lm	ll	lm	ll
1841	2982	<lmxll>	{1-}	(ll,lm)	54.043	ll	ll	--	ll	lm	ll	--	lm	lm	lm	lm	--	ll	lm	ll	ll	ll	lm	lm	lm	ll	--	lm	ll	ll	lm	ll	lm	lm	lm	ll	lm	--	ll	--	lm	lm	lm	ll	ll	lm	lm	ll	lm	ll
3445	386591	<lmxll>	{1-}	(ll,lm)	54.337	ll	ll	ll	ll	lm	ll	ll	lm	lm	lm	lm	ll	ll	lm	ll	ll	ll	lm	lm	lm	ll	ll	lm	ll	ll	lm	ll	lm	lm	lm	ll	lm	ll	ll	ll	lm	lm	lm	ll	ll	lm	lm	ll	lm	ll
3661	386601	<lmxll>	{1-}	(ll,lm)	54.409	ll	ll	ll	ll	lm	ll	ll	lm	lm	lm	lm	ll	ll	lm	ll	ll	ll	lm	--	lm	ll	ll	--	ll	ll	lm	ll	lm	lm	lm	ll	lm	ll	ll	ll	lm	lm	lm	ll	ll	lm	lm	ll	lm	ll
1296	15382	<lmxll>	{1-}	(ll,lm)	54.458	ll	ll	ll	ll	lm	ll	ll	lm	lm	lm	lm	ll	ll	lm	ll	ll	ll	lm	lm	lm	ll	ll	lm	--	ll	lm	ll	lm	lm	lm	ll	lm	ll	ll	--	lm	lm	lm	ll	ll	lm	lm	ll	lm	ll
3562	386703	<lmxll>	{1-}	(ll,lm)	54.862	ll	ll	--	ll	lm	ll	ll	lm	lm	lm	lm	--	ll	--	ll	--	ll	lm	lm	lm	ll	ll	lm	ll	ll	lm	ll	lm	lm	lm	ll	lm	ll	ll	ll	lm	lm	lm	ll	ll	ll	--	ll	--	ll
1716	386693	<lmxll>	{1-}	(ll,lm)	55.300	ll	ll	ll	ll	lm	ll	ll	lm	lm	lm	lm	ll	ll	lm	ll	--	ll	lm	lm	lm	ll	ll	--	ll	ll	lm	--	lm	lm	lm	ll	lm	ll	ll	ll	lm	lm	lm	ll	ll	lm	lm	ll	--	ll
3672	210453	<lmxll>	{1-}	(ll,lm)	56.311	ll	ll	ll	ll	lm	ll	--	lm	lm	lm	lm	ll	ll	lm	ll	--	ll	lm	lm	lm	ll	ll	--	ll	ll	lm	ll	lm	lm	lm	ll	lm	ll	ll	ll	lm	lm	lm	ll	ll	lm	lm	ll	lm	ll
3655	274548	<lmxll>	{1-}	(ll,lm)	58.674	ll	ll	ll	ll	lm	ll	ll	lm	lm	lm	lm	ll	ll	lm	ll	lm	ll	lm	lm	lm	ll	--	lm	ll	ll	lm	ll	lm	lm	lm	ll	lm	ll	ll	ll	lm	lm	lm	ll	ll	lm	ll	ll	lm	ll
1935	384895	<hkxhk>	{01}	(hh,hk,kk)	55.760	hk	hk	kk	kk	hh	kk	hk	hk	hh	hh	hh	hk	hk	--	kk	kk	hk	hh	hh	hk	kk	kk	hk	kk	kk	hk	hk	hk	hk	hh	kk	hh	kk	kk	kk	hk	hh	hh	kk	hk	hk	hh	kk	hh	hk
279	384893	<hkxhk>	{01}	(hh,hk,kk)	55.901	hk	hk	kk	kk	hh	kk	hk	hk	hh	hh	hh	hk	hk	hk	kk	kk	hk	hh	hh	hk	kk	kk	hk	kk	kk	hk	hk	hk	hk	hh	kk	hh	kk	kk	kk	hk	hh	hh	kk	hk	hk	hh	kk	hh	hk
1194	384896	<hkxhk>	{10}	(hh,hk,kk)	56.082	hk	hk	hh	hh	kk	hh	hk	hk	kk	kk	kk	hk	hk	--	hh	hh	hk	kk	kk	hk	hh	hh	hk	hh	hh	hk	hk	hk	hk	kk	hh	kk	hh	hh	hh	hk	kk	kk	hh	hk	hk	kk	hh	kk	hk
1569	385149	<hkxhk>	{11}	(hh,hk,kk)	62.558	hh	hh	hh	hk	hk	hk	hk	kk	hk	hk	hk	hh	hh	kk	hk	hk	hh	hk	hk	kk	hk	hk	kk	hk	hk	kk	hh	kk	kk	hk	kk	hk	hk	hk	hk	kk	hk	hk	hk	hh	kk	hk	hk	hk	hh
1021	276023	<efxeg>	{01}	(ee,ef,eg,fg)	47.639	ef	ef	fg	fg	ee	--	--	eg	eg	ee	ee	ef	ef	eg	--	fg	ef	ee	ee	eg	fg	ef	--	ef	fg	ee	fg	eg	eg	ee	fg	ee	--	ef	fg	ee	ee	eg	fg	fg	eg	ee	--	ee	ef
2738	384977	<efxeg>	{01}	(ee,ef,eg,fg)	57.740	ef	ef	ef	fg	ee	fg	--	eg	ee	ee	ee	ef	ef	eg	fg	fg	ef	ee	ef	eg	fg	fg	eg	fg	fg	eg	ef	eg	eg	ee	fg	ee	fg	fg	fg	eg	ee	ee	fg	ef	eg	ee	fg	ee	ef

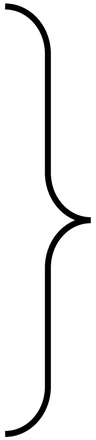

4
1.5
1.0
0.5
0.0
-0.5
0
5
10
15
20


Genetic distances (cM)
20
15
10
5
0
0
20
40
60
80
100
140
Physical positions (Mb)

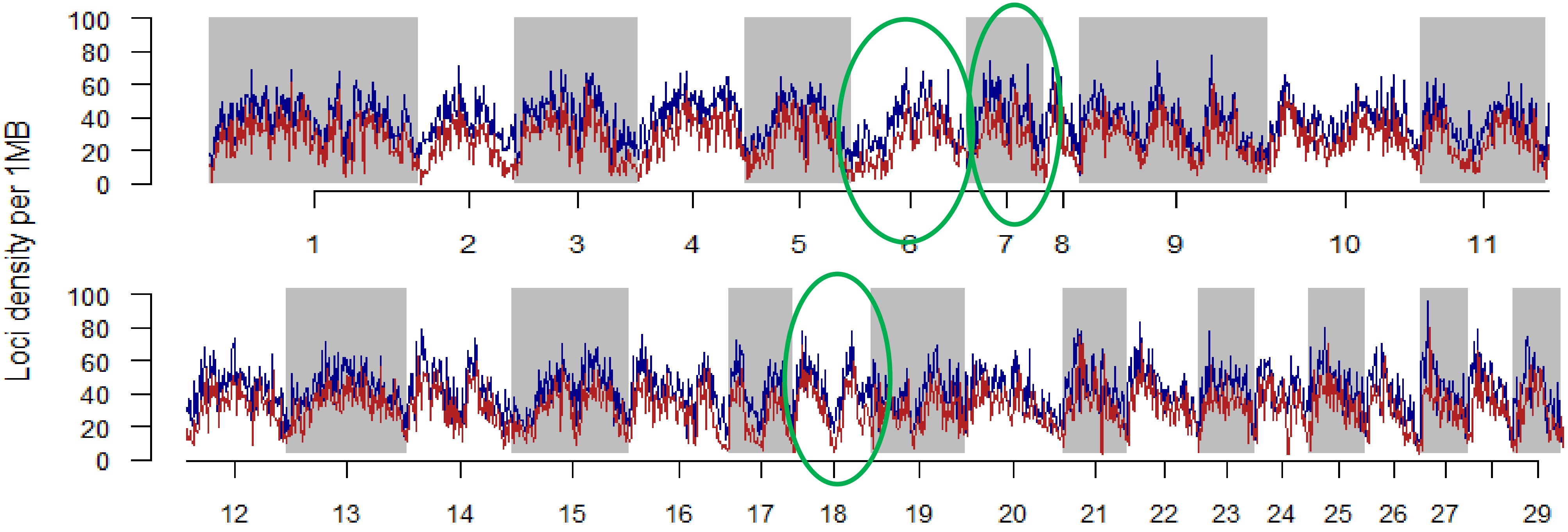

Genetic distances (cM)
20
15
10
5
0
0
2
4
6
8
Physical positions (Mb)


Genetic distances (cM)
20
15
10
5
0
20
40
60
80
100
140
Physical positions (Mb)

- → 3 Metacentics
- →32 Acrocentrics
- →5 undeterminde



NF = 86 -96
Expected NF (karyotype studies) = 96-104



Part 3

Work to do…

→ Detecting hybrid individuals and evaluating the amount of admixture for all individuals samples (N = 184)

→ Assess the introgression patterns of individual loci along the genome and identify outliers SNPs

→ Use the local recombination landscape to determine linkage disequilibrium between loci

→ Assess the introgression at haplotype level then identify introgressed haplotype into wild individuals

→ Use the introgressed haplotype size as a proxy of the timing of the introgression

→ Determine the introgression rate along the genome

→ Find signatures of adaptive introgression or barriers to gene flow