Overview of the genetic variability of genes FUT1, FUT2 and FUT3 in various pig populations

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

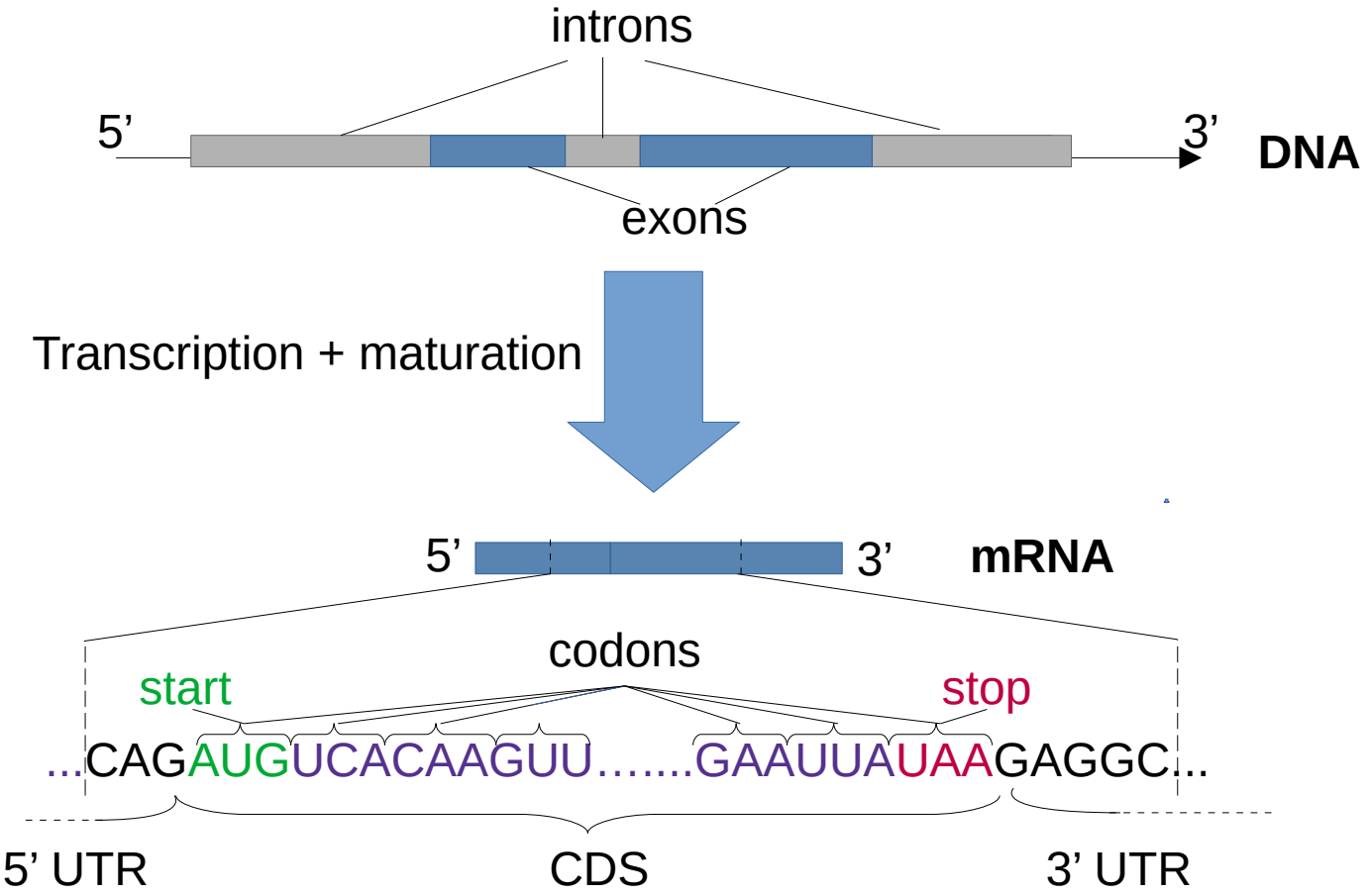
INRAE

Introduction

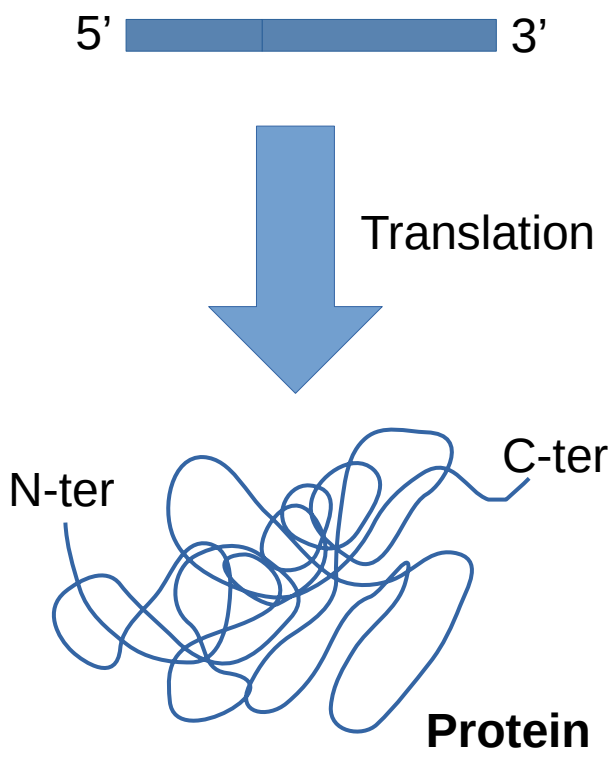
- Milk oligosaccharide structures are conditioned by glycosyltransferases, including **fucosyltransferases** (FUT)
- FUT gene family :
  - In humans : **11** genes (FUT1 → FUT11)
  - In pigs : **9** genes (FUT1 → FUT4, FUT7 → FUT11)
- Our focus : genes **FUT1, FUT2 and FUT3**
  - Involved in formation of Lewis blood group antigens
  - If the enzyme is inactive : decreased diversity in milk oligasaccharides
- Our **goal** : study the variability of the FUT1, FUT2 and FUT3 genes in various pig populations
  - Are there alleles **specific** to some breeds?
  - Are there alleles that produce **non-functional** enzymes?

Fundamentals

- In eukaryotes, protein-coding genes can have introns and exons
  - **Introns** are removed during the mRNA maturation process
  - **Exons** remain in the final mature mRNA and carry the coding sequence (CDS)



Fundamentals



- Function of protein determined by its shape and AA sequence

What is a variant ?

- When a genome is sequenced, it can be aligned on a **reference genome**
- If the sequence is different, it is called a **variant**
- Different types of variants :

Reference	CAGTATGCTA
SNP	CAGT**T**TGCTA
Deletion	CATT**-**TGCTA
Insertion	CATT**A**GTGCTA

- Variants can impact the protein produced :
  - SNP in coding sequence (CDS) :
    - New codon codes for the same AA (synonymous)
    - New codon codes for a different AA (missense)
    - New codon is STOP (nonsense)
  - Insertion/deletion (indel) in CDS :
    - Reading frame shifts (frameshift)

What is a variant ?

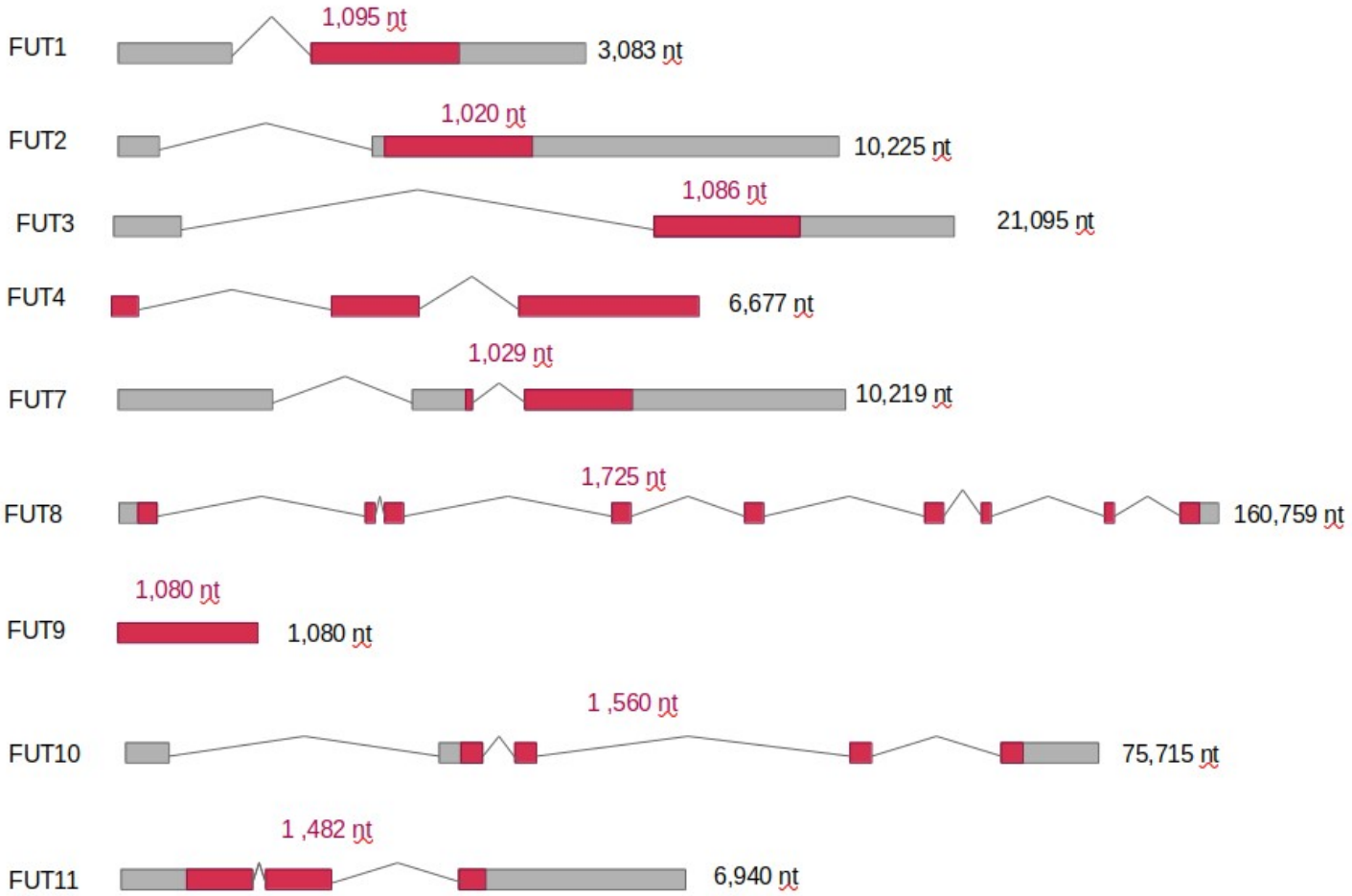
- Variants can impact the protein produced :
  - SNP in coding sequence (CDS) :
    - New codon codes for the same AA (**synonymous**)
    - New codon codes for a different AA (**missense**)
    - New codon is STOP (**nonsense**)
  - Insertion/deletion (**indel**) in CDS :
    - Reading frame shifts (**frameshift**)

Reference	DNA	ATGCATGTGTATAGGCCA
	AA	M H V Y R P
Insertion	DNA	ATGCATGTGTAATAGGCCA
	AA	M H V STOP

  - Variants outside of CDS (UTR, introns, down/upstream) :
    - Promoters and regulatory regions can be affected

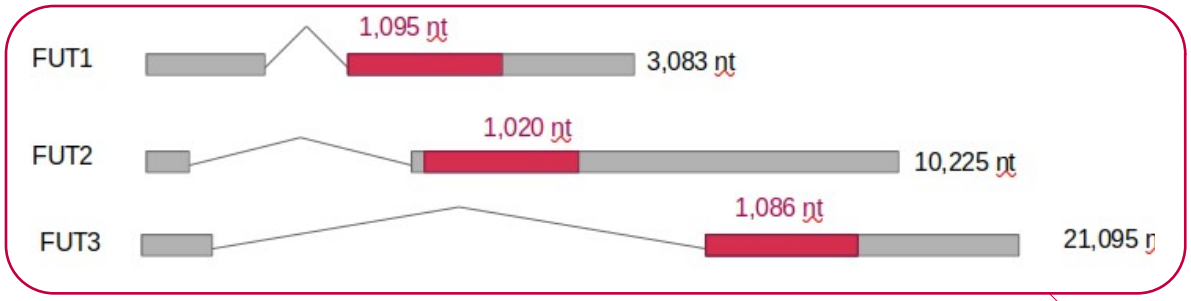
FUT genes in pigs

- Great variability in terms of gene structure (length, number of exons/introns, etc...)



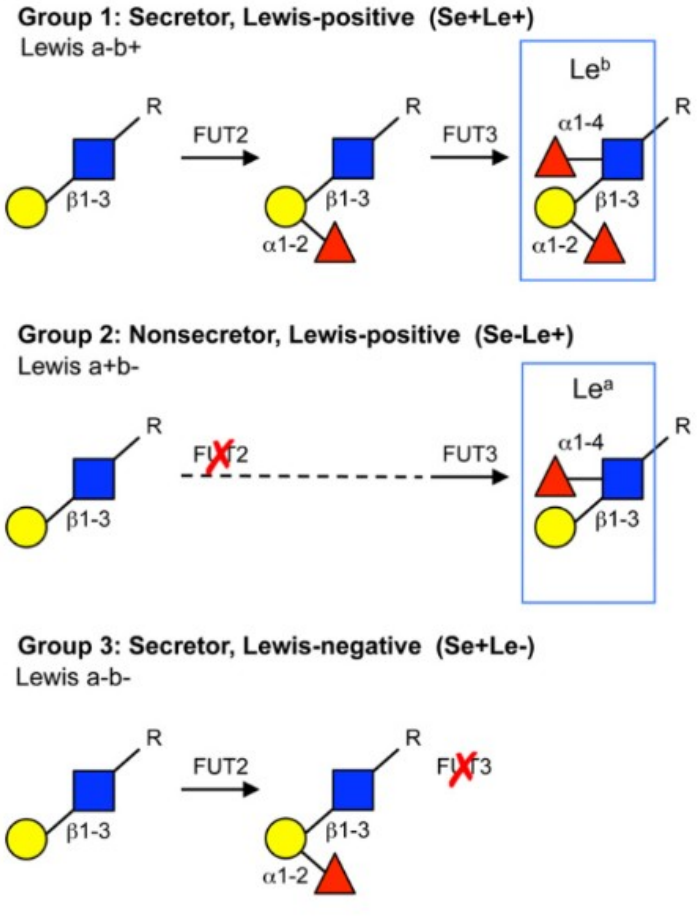
FUT genes in pigs

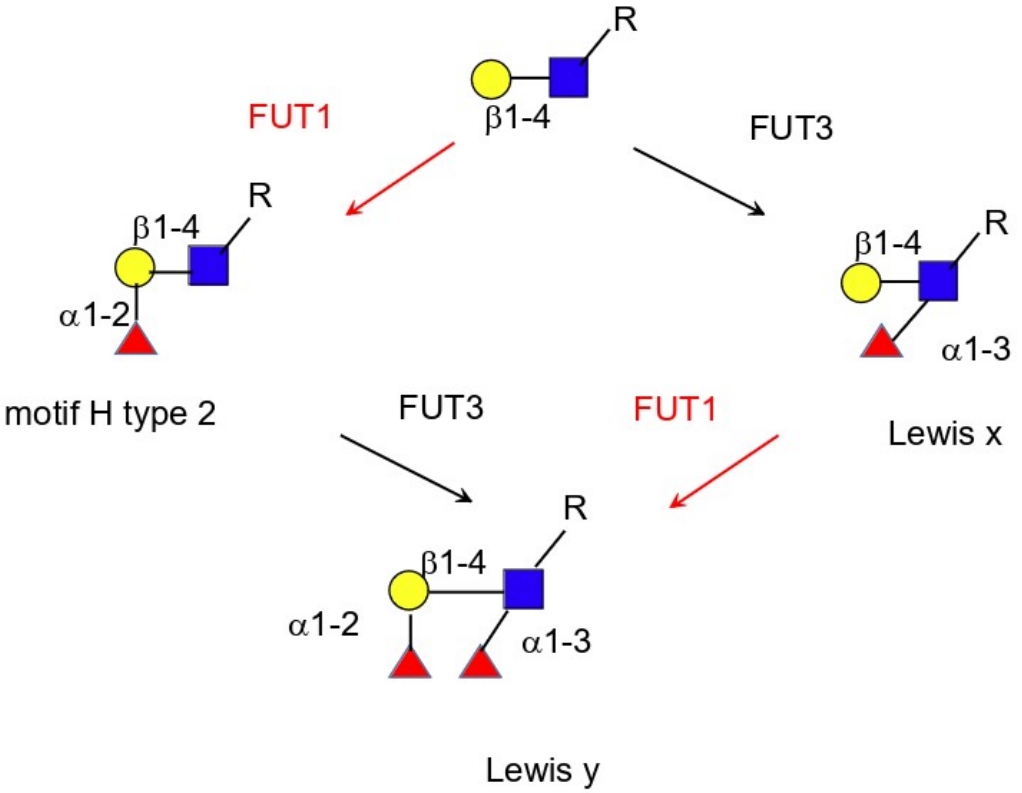
- Great variability in terms of gene structure (length, number of exons/introns, etc...)



For humans :
Lewis group a/b : FUT2, FUT3
Lewis group x/y : FUT1, FUT3

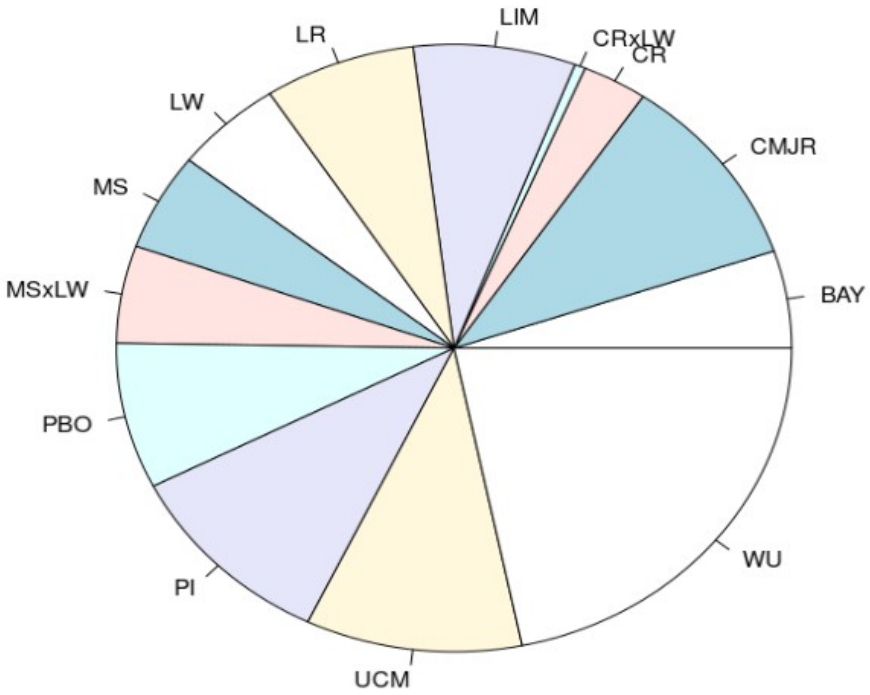
Lewis groups and HMOs





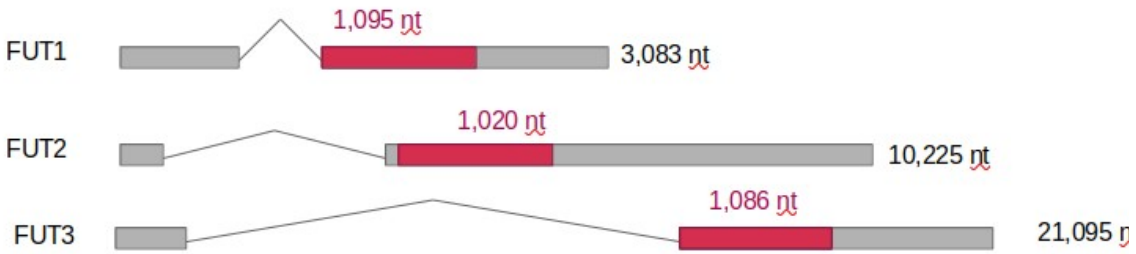
Our goal

- Examine variants and haplotype (unique variant combination) diversity of FUT1, FUT2 and FUT3 across multiple pig populations
  - Use available data : 194 animals
  - 2 unknown populations : WU and UCM



BAY	10
CMJR	20
CR	6
CRxLW	1
LIM	15
LR	14
LW	10
MS	10
MSxLW	10
PBO	15
PI	20
UCM	20
WU	42

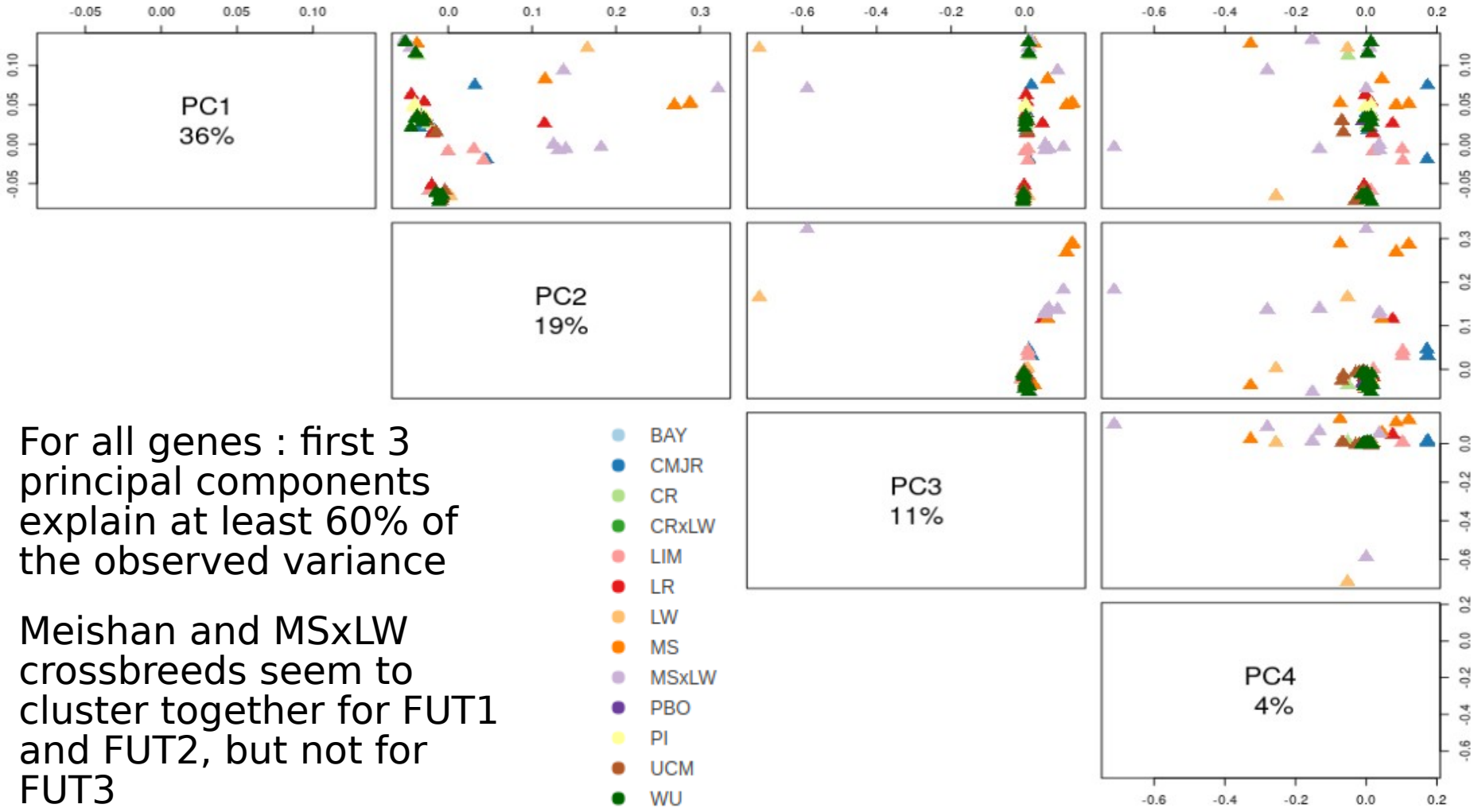
FUT1, FUT2 and FUT3 variants



- Annotation with VEP
- Filters:
  - Quality >= 30
  - Biallelic only

Variant consequence	FUT1	FUT2	FUT3
Frameshift variant	2	0	2
Missense variant	7	4	2
Synonymous variant	3	3	4
5' UTR variant	1	0	0
3' UTR variant	6	45	2
Intron variant	3	91	277
Upstream gene variant	9	36	19
Downstream gene variant	12	7	9
Splice donor 5th base variant	0	0	1
Stop gained	0	1	0
TOTAL	**43**	**187**	**316**

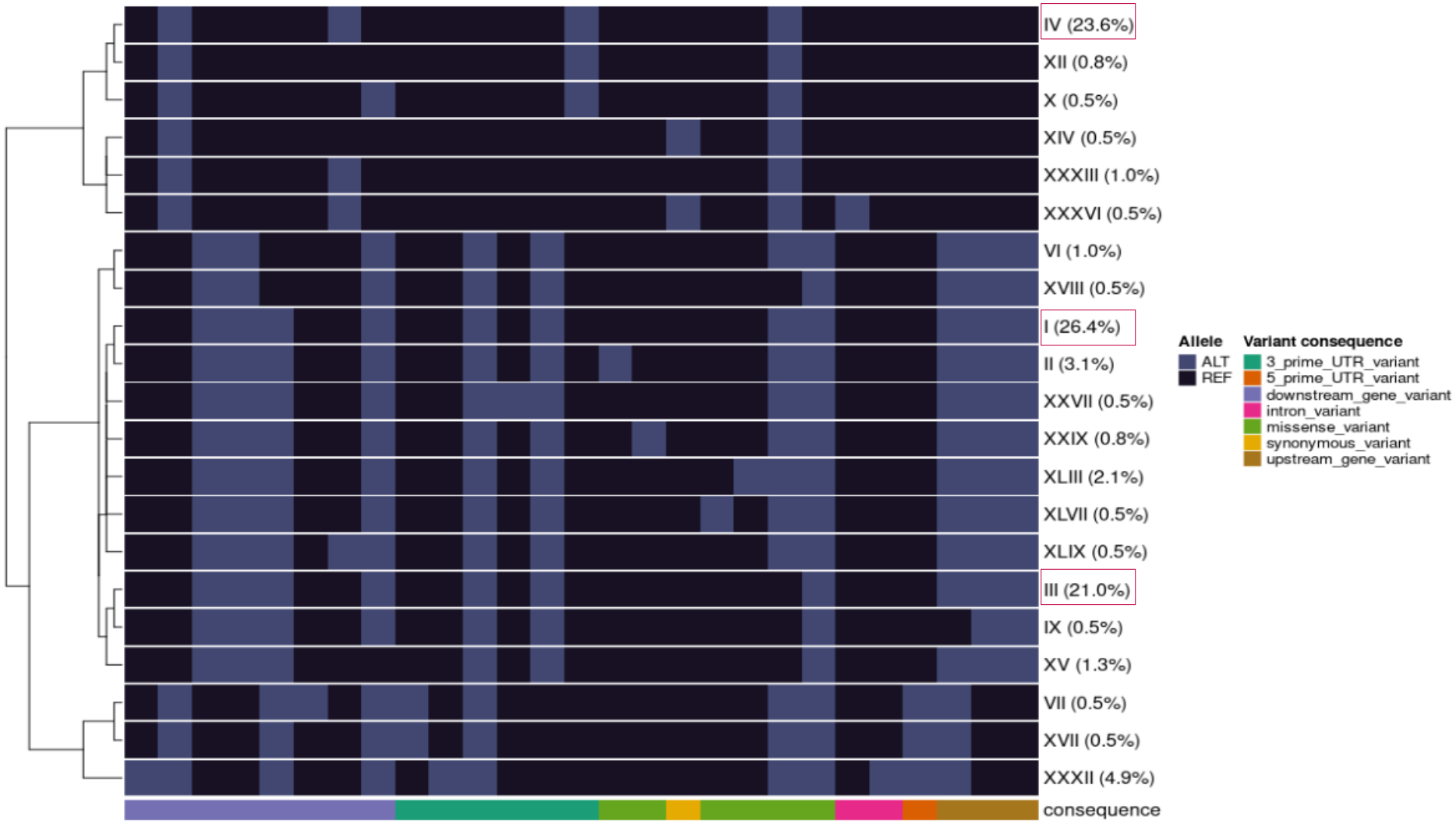
PCA - FUT1



- For all genes : first 3 principal components explain at least 60% of the observed variance
- Meishan and MSxLW crossbreeds seem to cluster together for FUT1 and FUT2, but not for FUT3

Majority haplotypes - FUT1

- 55 haplotypes in total, 25 observed at least twice

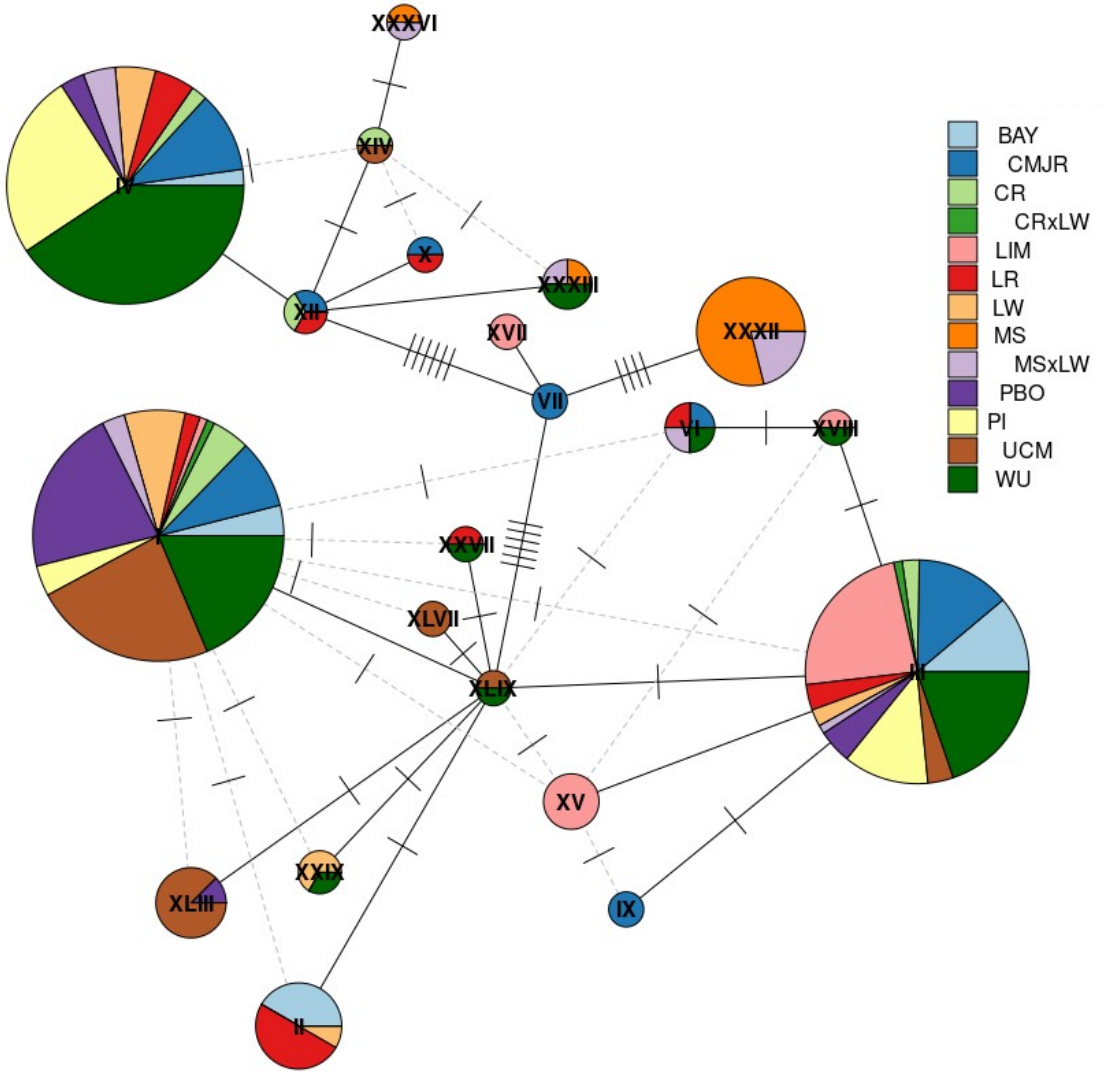


Majority haplotypes in overall population

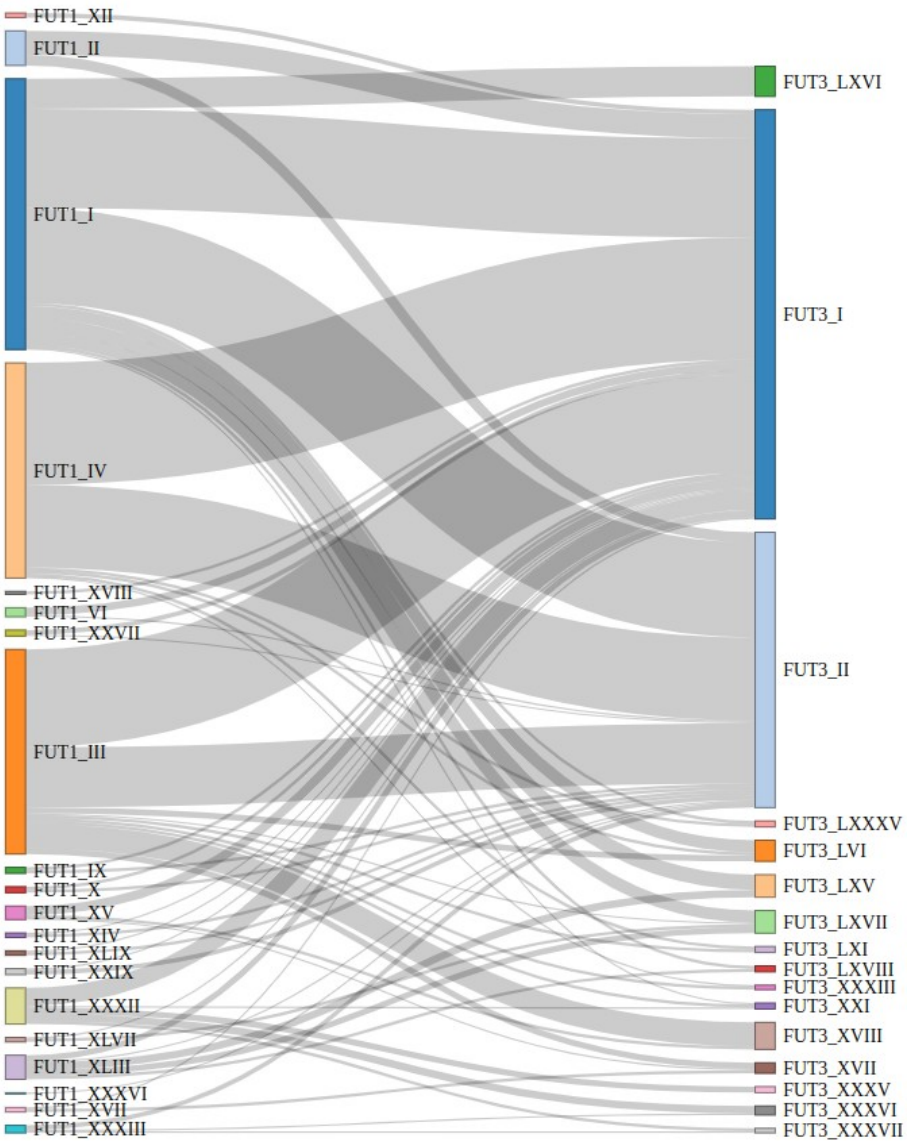
- FUT1:
  - I (26,4%)
  - III (21%)
  - IV (23,6%)
  - Total : 71%
- FUT2:
  - IV (31,1%)
  - V (17,4%)
  - II (7,3%)
  - Total : 55,8%
- FUT3
  - I (38,6%)
  - II (25,4%)
  - Total : 64%

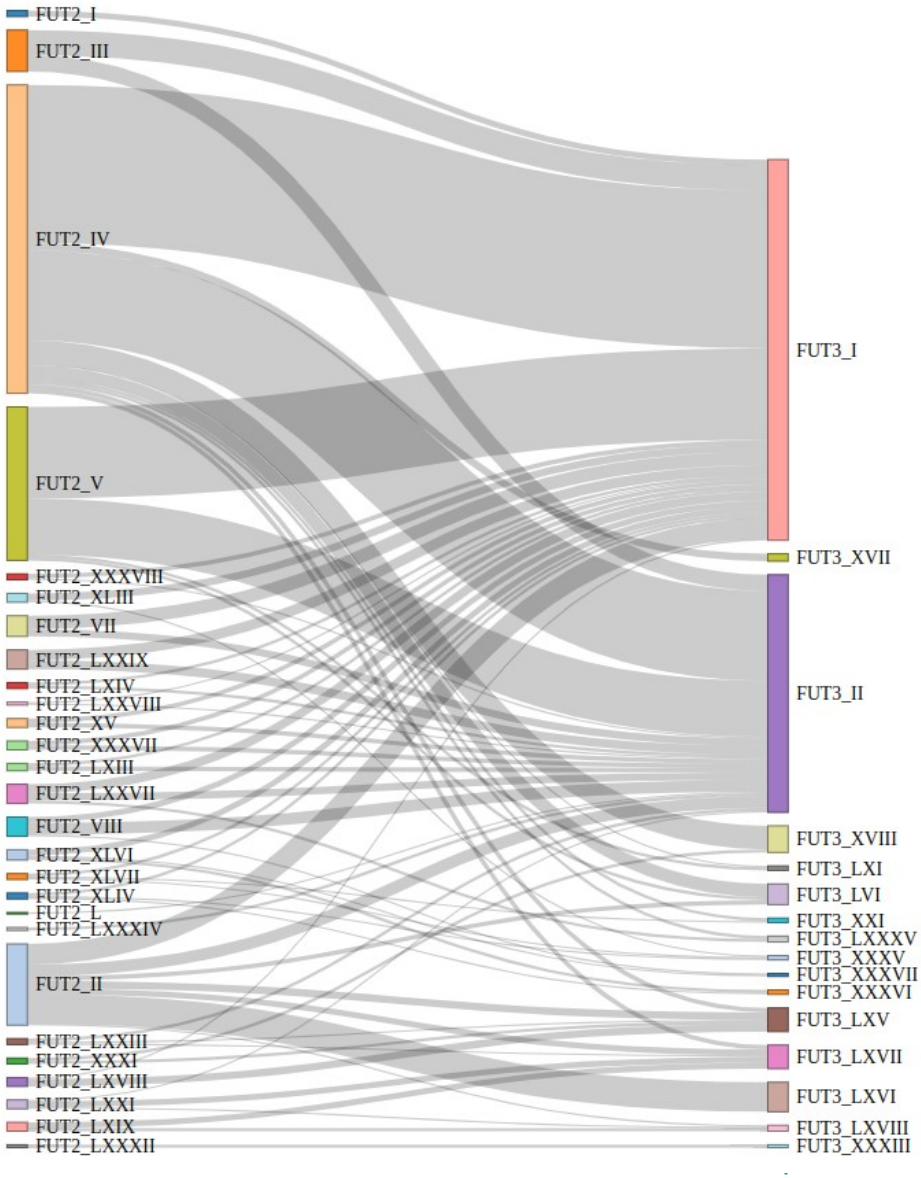
Haplotype distribution across populations - FUT1

Haplotype distribution across populations - FUT1



Haplotype combinations





Conclusions

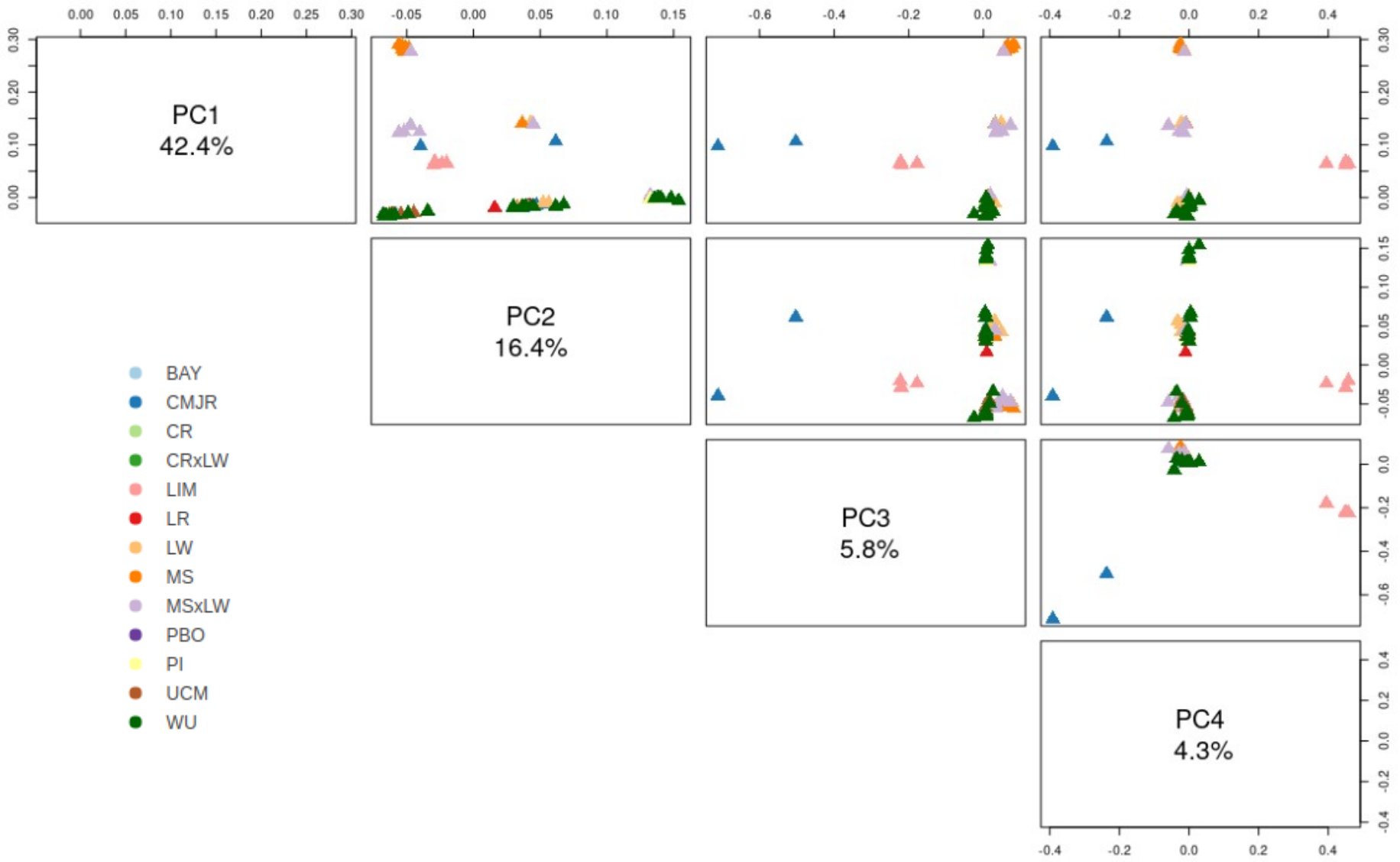
- No strong evidence of inactive FUT1, FUT2 or FUT3
- FUT3 highly conserved
- Meishan population distinct for FUT1 and FUT2, less for FUT3
- Work reproducible and compatible with other genes
- **Perspective** :
  - Apply to other FUT genes
  - Compare variants found with human variants

Thank you !

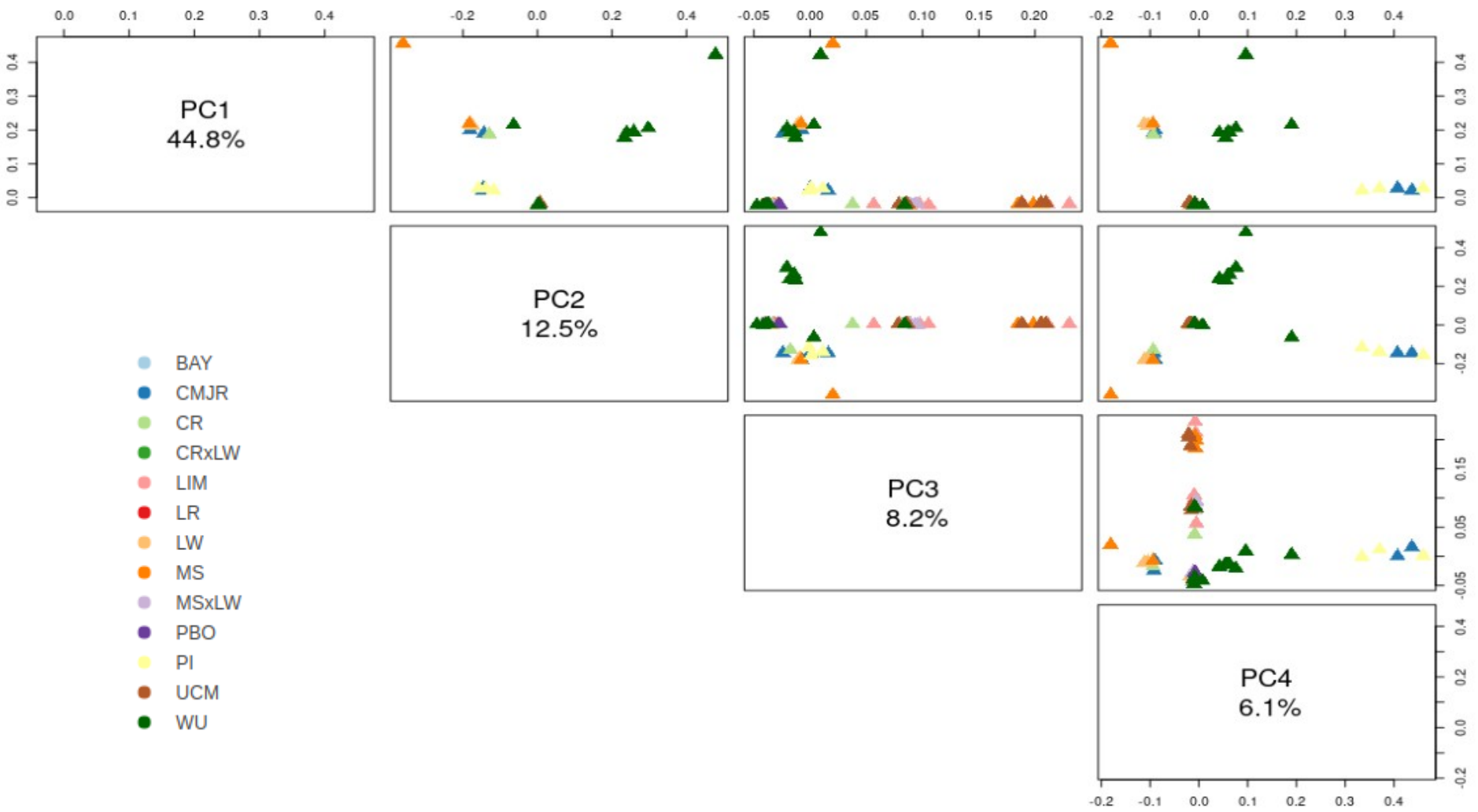
RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

INRAE

PCA - FUT2

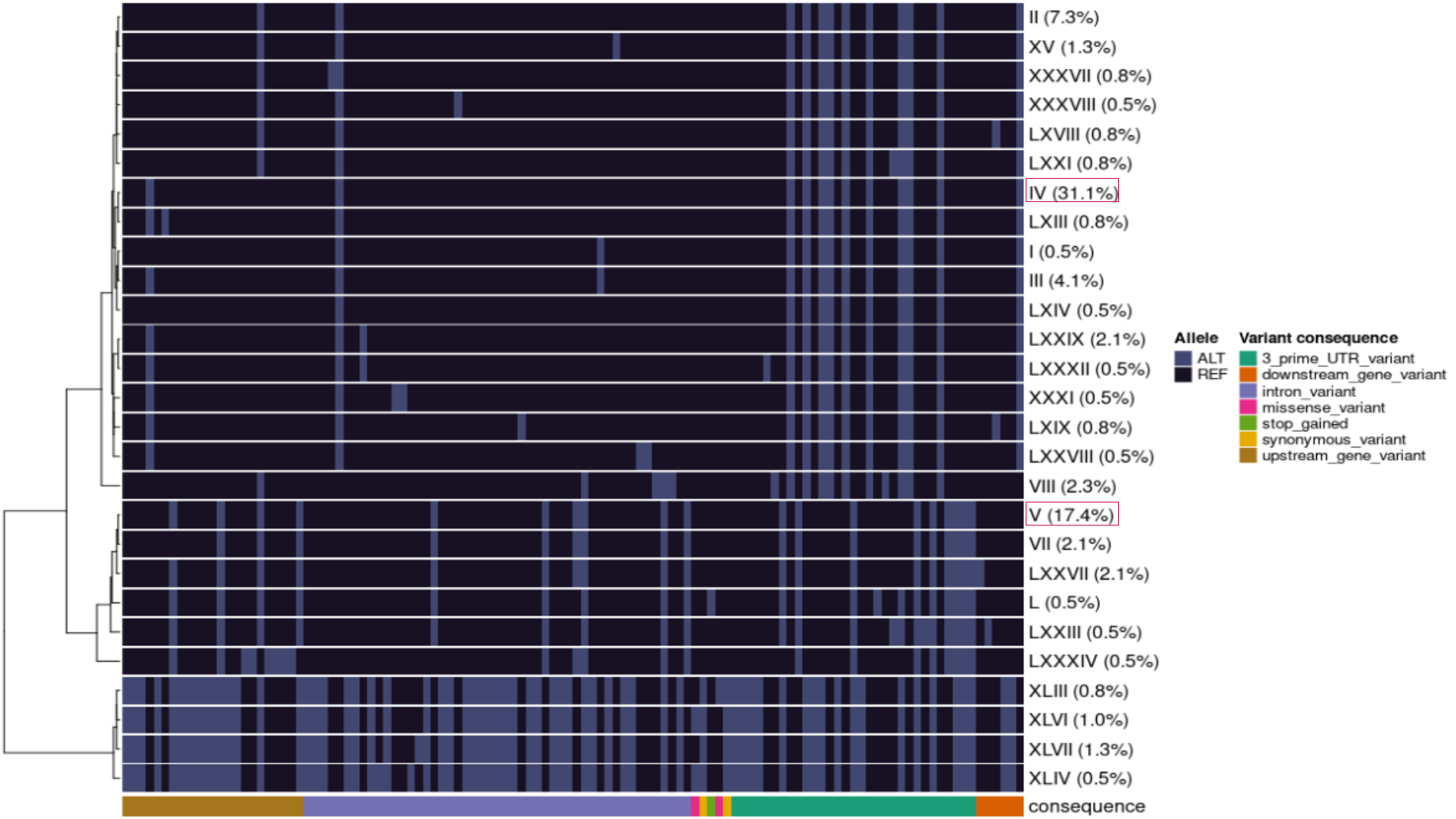

PC1
42.4%
PC2
16.4%
PC3
5.8%
PC4
4.3%
BAY
CMJR
CR
CRxLW
LIM
LR
LW
MS
MSxLW
PBO
PI
UCM
WU

PCA - FUT3


PC1
44.8%
PC2
12.5%
PC3
8.2%
PC4
6.1%
BAY
CMJR
CR
CRxLW
LIM
LR
LW
MS
MSxLW
PBO
PI
UCM
WU

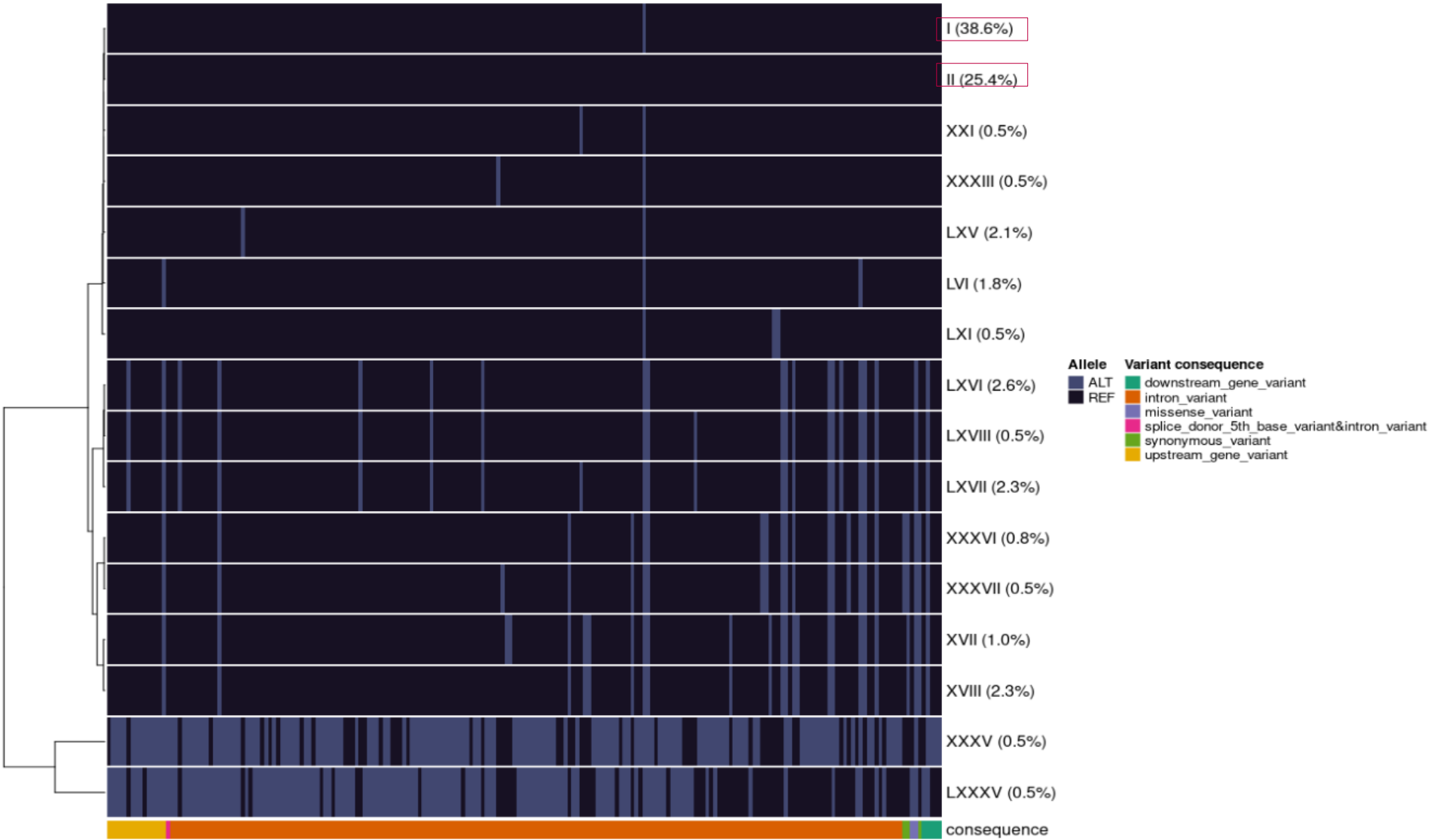
Majority haplotypes - FUT2

- 97 haplotypes in total, 27 observed at least twice

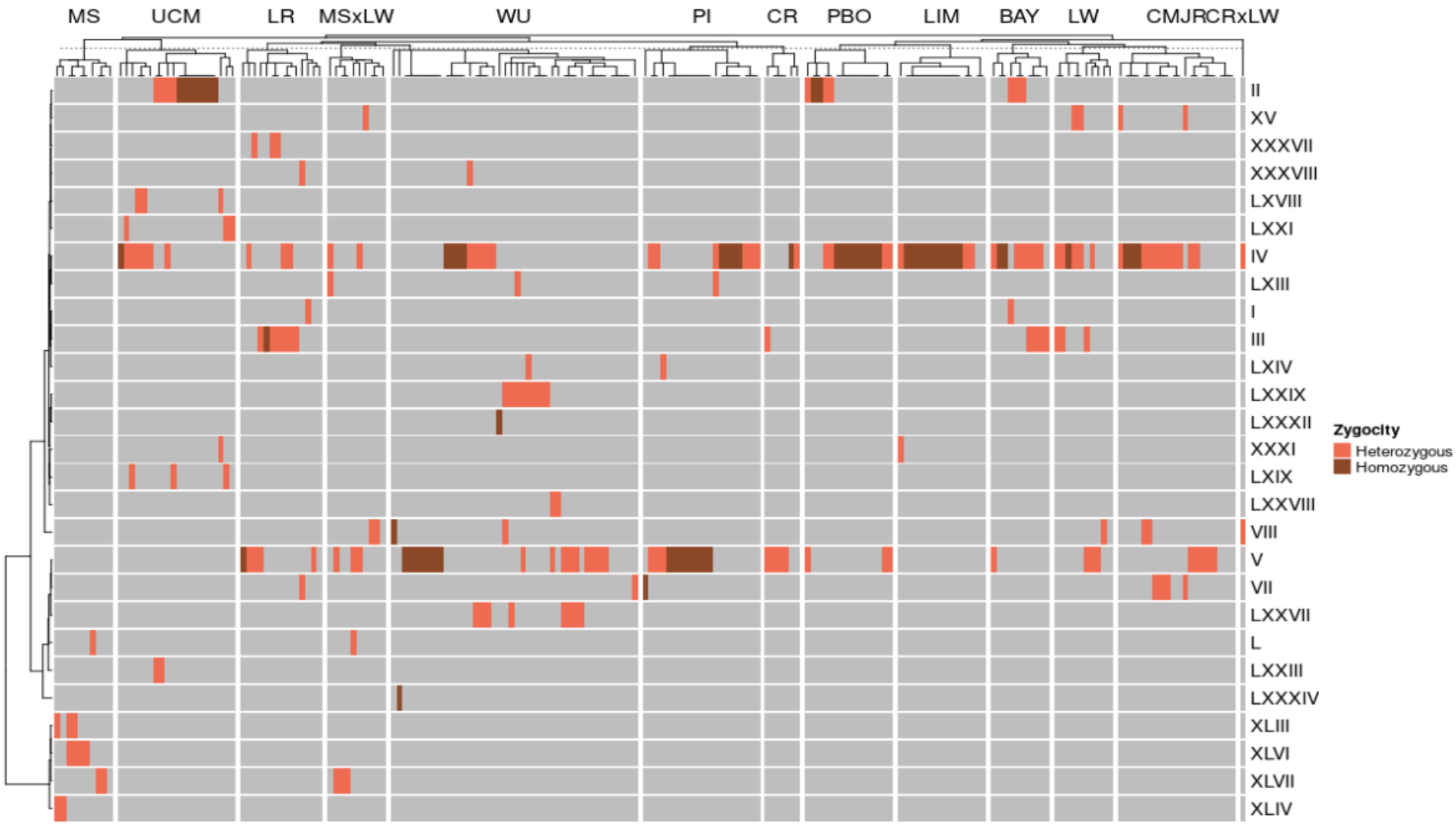


Majority haplotypes - FUT3

- 91 haplotypes in total, 16 observed at least twice

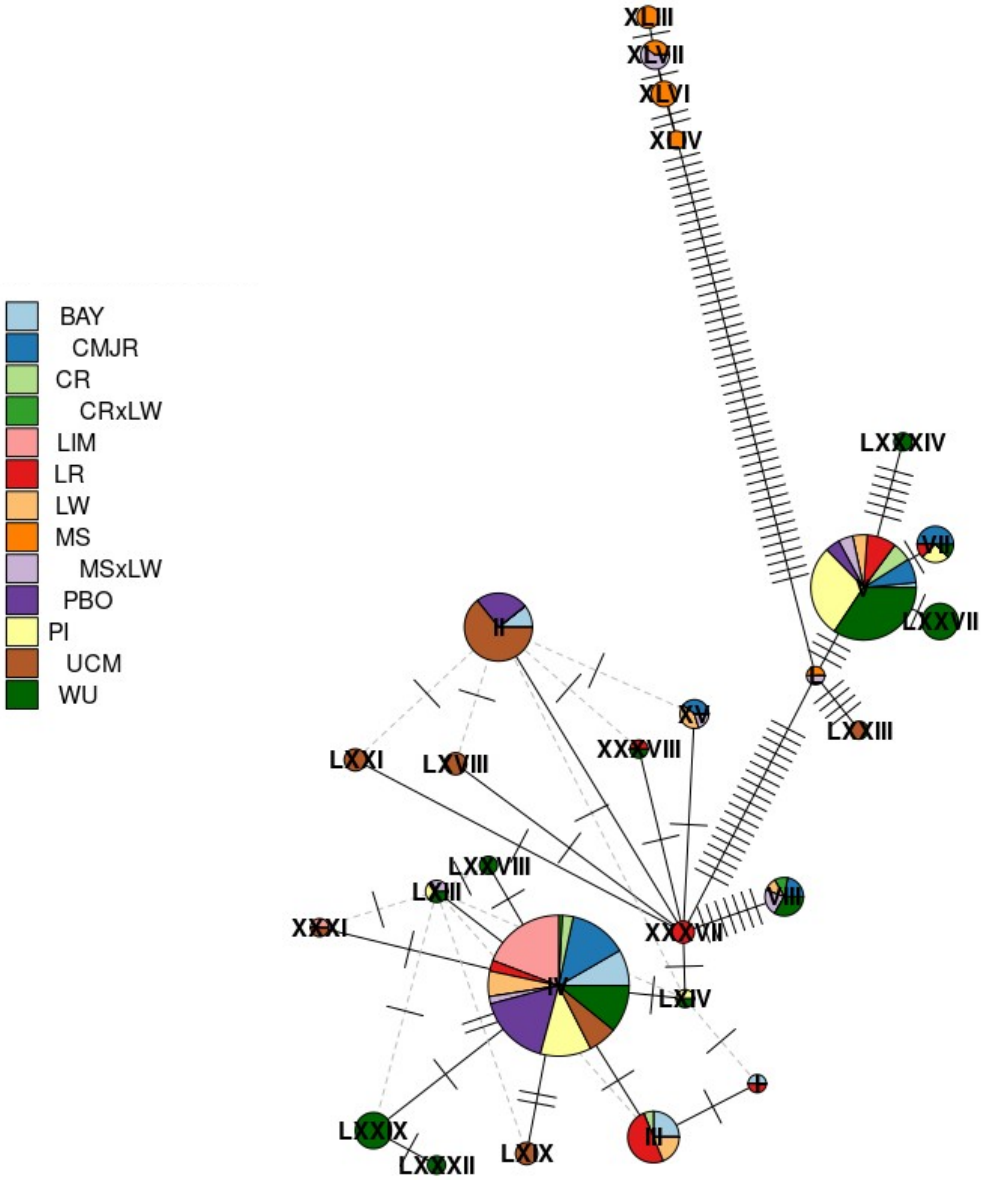
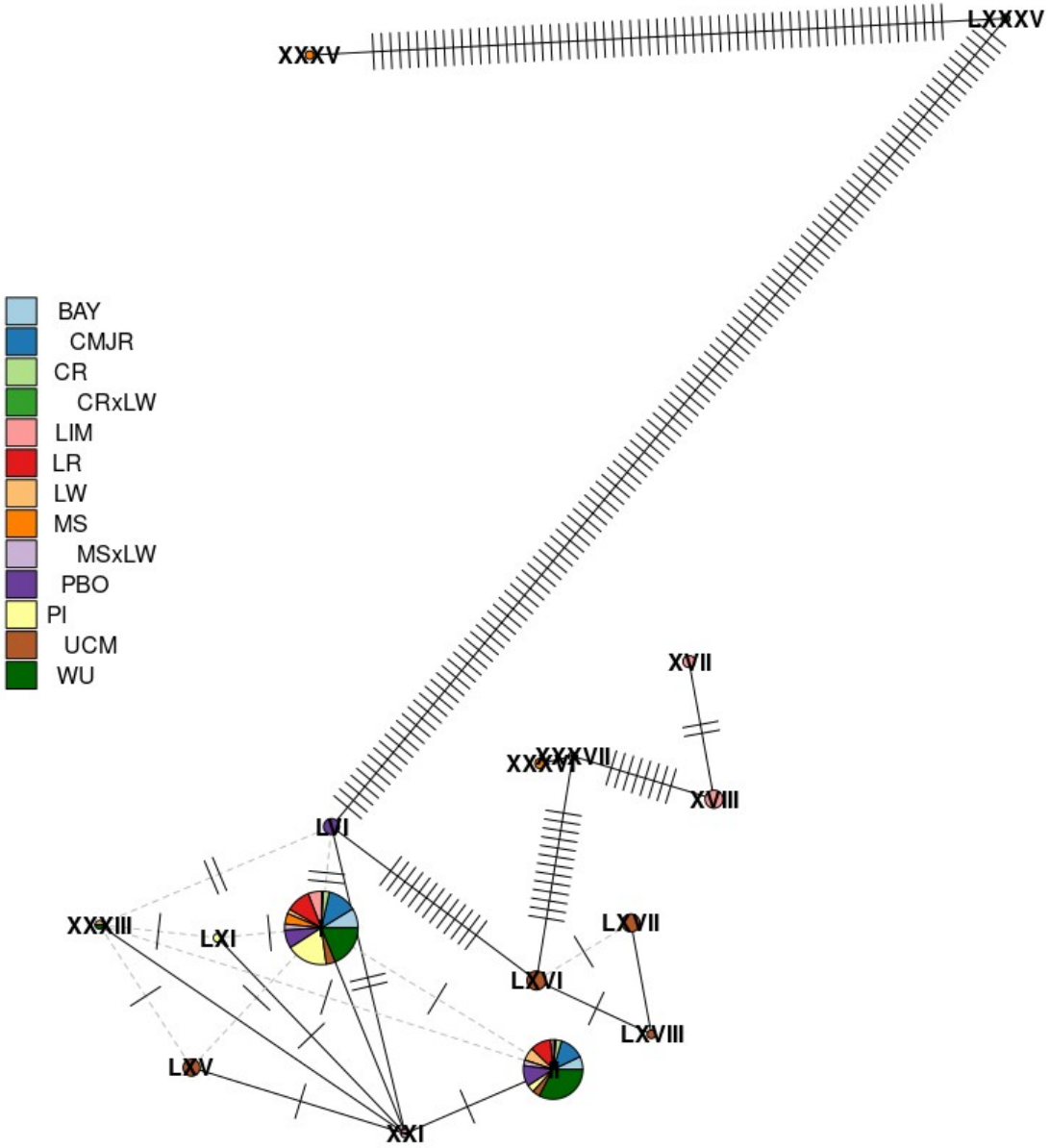


Haplotype distribution across populations - FUT2



Haplotype distribution across populations - FUT3

Haplotype distribution across populations - FUT2



Haplotype distribution across populations - FUT3