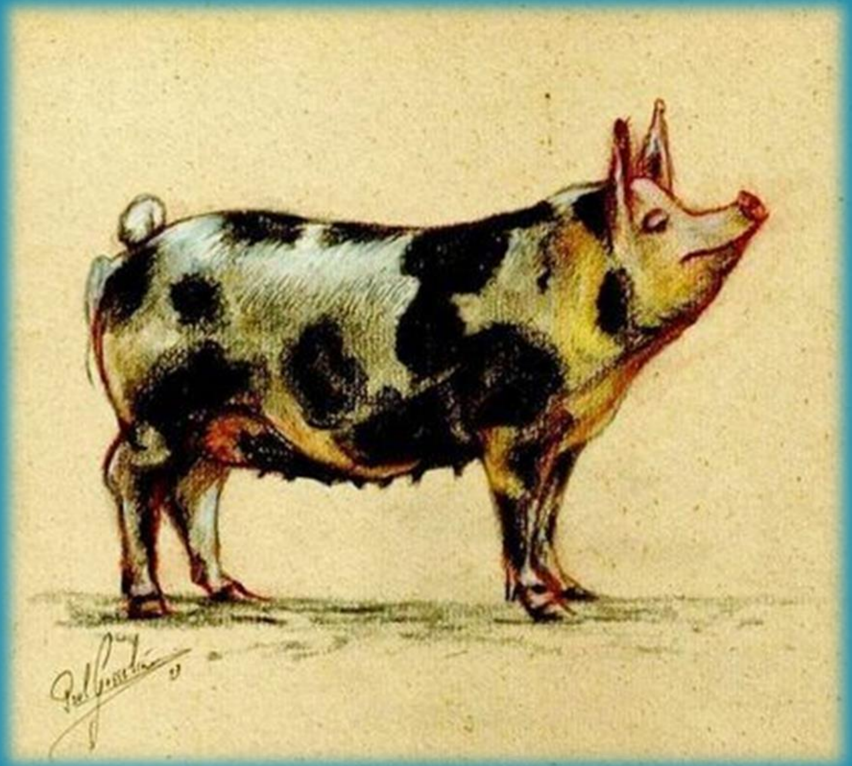
Will the real Piétrain please stand up?



Genomics reveal Piétrain pig diversity and substructure

Wim Gorssen, N. Buys, R. Meyermans, J. Depuydt and S. Janssens

August 28, 2019

Livestock Genetics, Department of Biosystems, KU Leuven, 3001 Leuven, Belgium

A little history on the pig of Piétrain

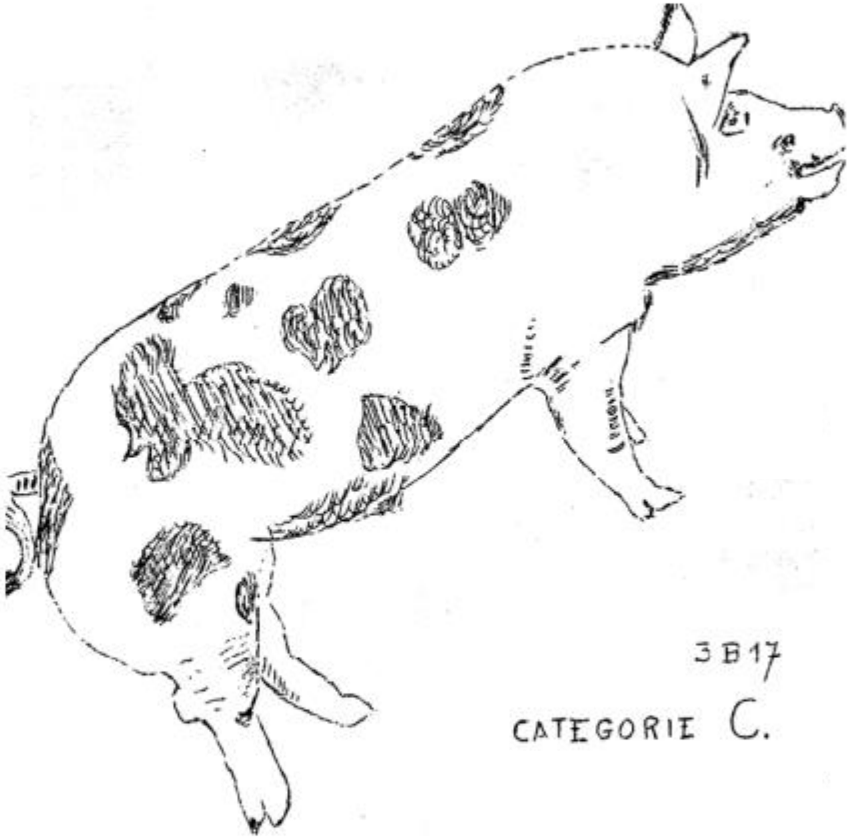
Originates around the small village of Piétrain, Belgium (1920-1950)

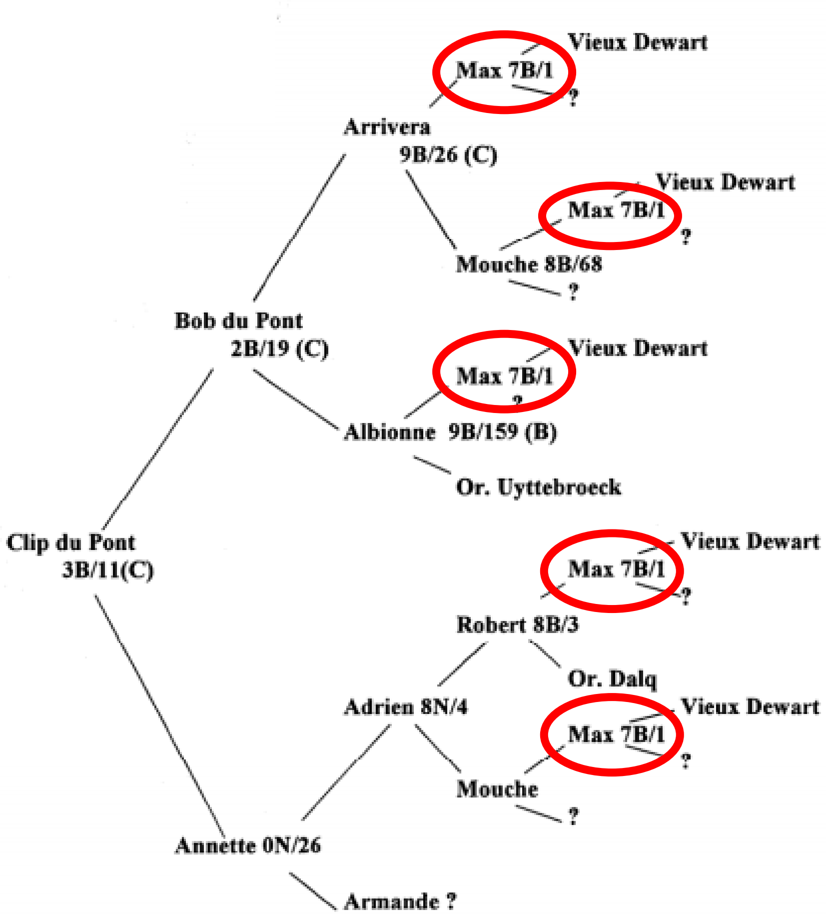
- Unknown origin
  - Berkshire, Bayeux, Large White, wild boars,…?
- Initially very small active population size
- Pigbook started in 1950

After WWII

- Demand for lean meat rises
- Piétrain conquers Belgium… and the world as terminal sire line

Severe inbreeding during breed formation





Research questions

1. Piétrain diversity
    - What is the inbreeding level and effective population size within the Piétrain breed?
    - Has separation in subpopulations from the 196^{2}0’s on resulted into significant breed substructure?

2. Can we detect Piétrain-specific signatures of selection?

Material and methods

1632 Piétrain genotypes and 34551 SNPs after quality control and merging in PLINK 1.9

- Effective population size (N_E) calculated via SNeP
- F_{st} analysis via Hierfstat
- Runs of homozygosity (ROH) analysis and PCA using PLINK

Breed	Subpopulation	N	Array	SNPs
Piétrain	Belgium	577	GGP HD	68516
	France	172	Illumina	61565
	Germany	843	Illumina	62163
	The Netherlands*	20	Illumina	61772
	USA*	20	Illumina	61772

*Yang et al. (2017). *Genetics selection evolution*, *49*(1), 71.

Inbreeding estimation based on runs of homozygosity (ROH) analysis using PLINK*

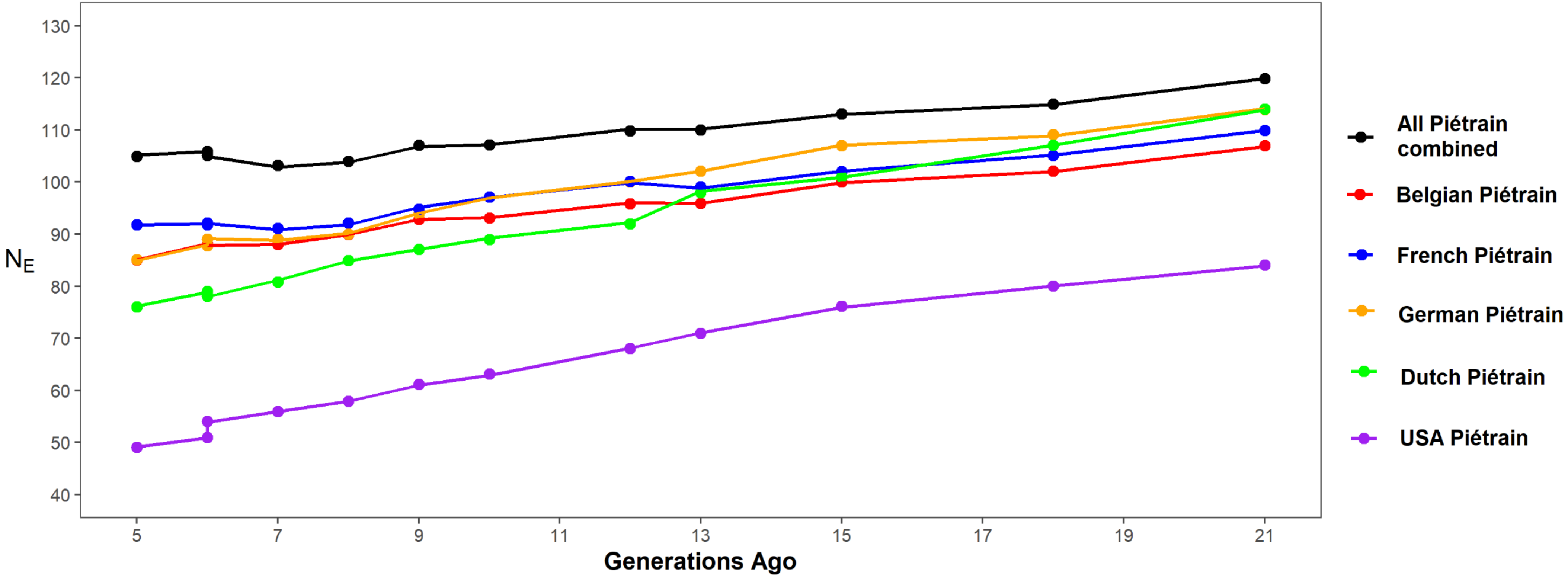
Breed	**Subpopulation**	**F_{ROH} (%)**	**$F_{ROH>4Mb}$ (%)**
Piétrain	Belgium	22.0	19.3
	France	20.5	19.8
	Germany	18.0	16.7
	The Netherlands	21.3	18.7
	USA	23.0	20.9

Crosses between Belgium and German Piétrain (n=8) on average had 15.7% $F_{ROH>4MB}$

➢ Indication that identical homozygous fragments still appear between subpopulations!

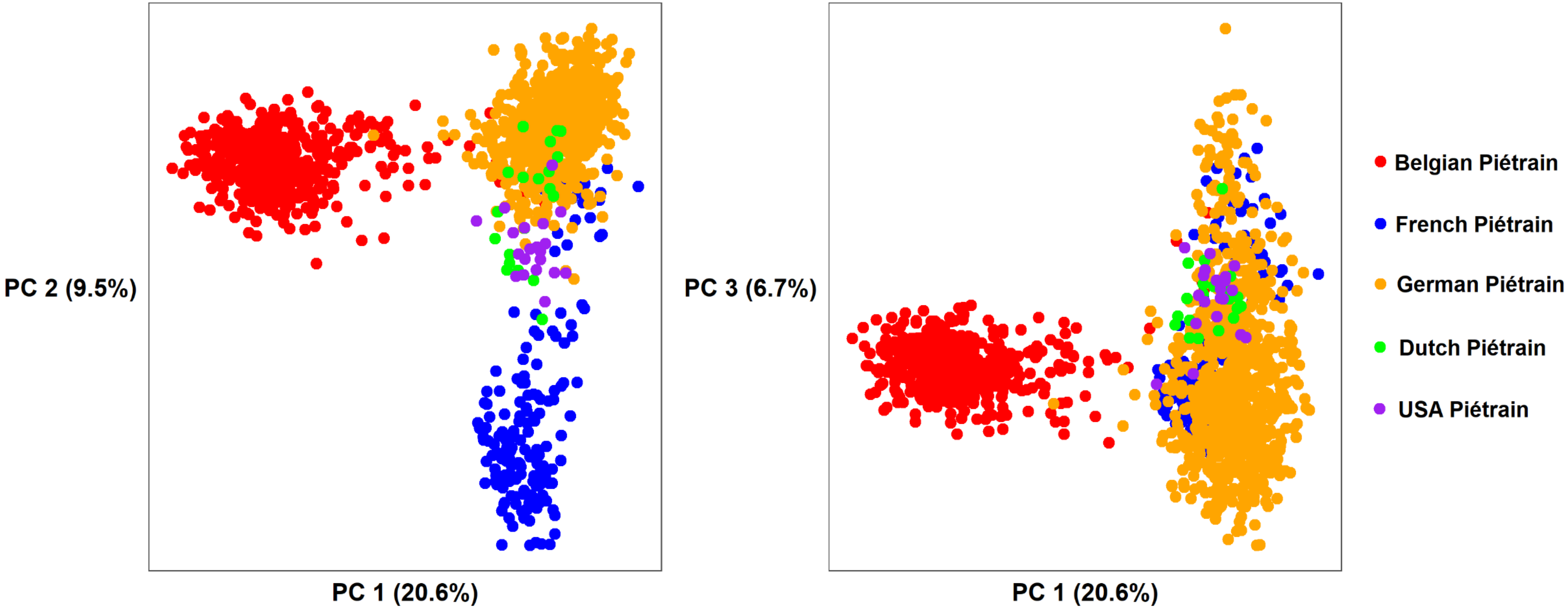
*following Meyermans & Gorssen et al. (See session 54)

Effective population sizes (N_E) indicate sufficient genetic diversity in Piétrain breed



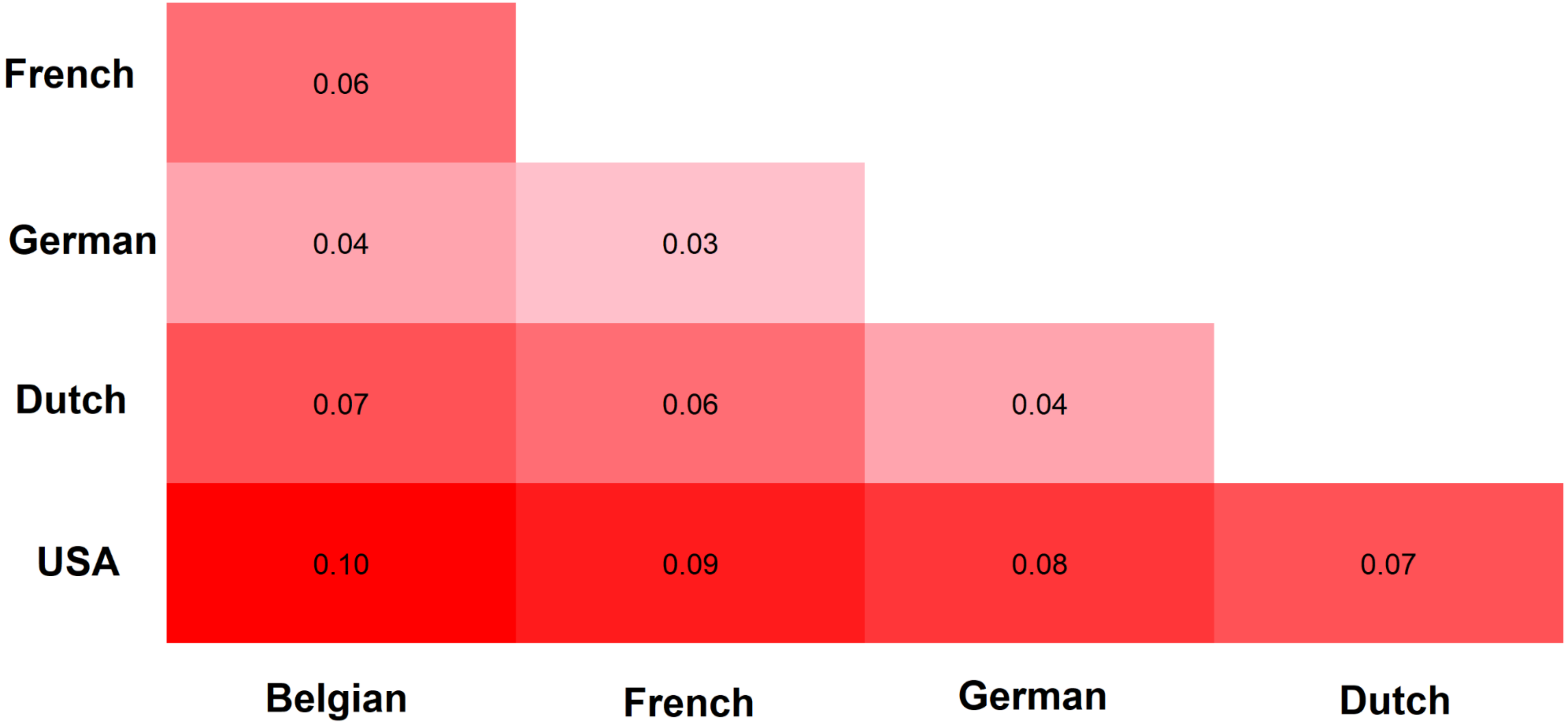
Gorssen et al., submitted

PCA shows three subclusters within the Piétrain breed



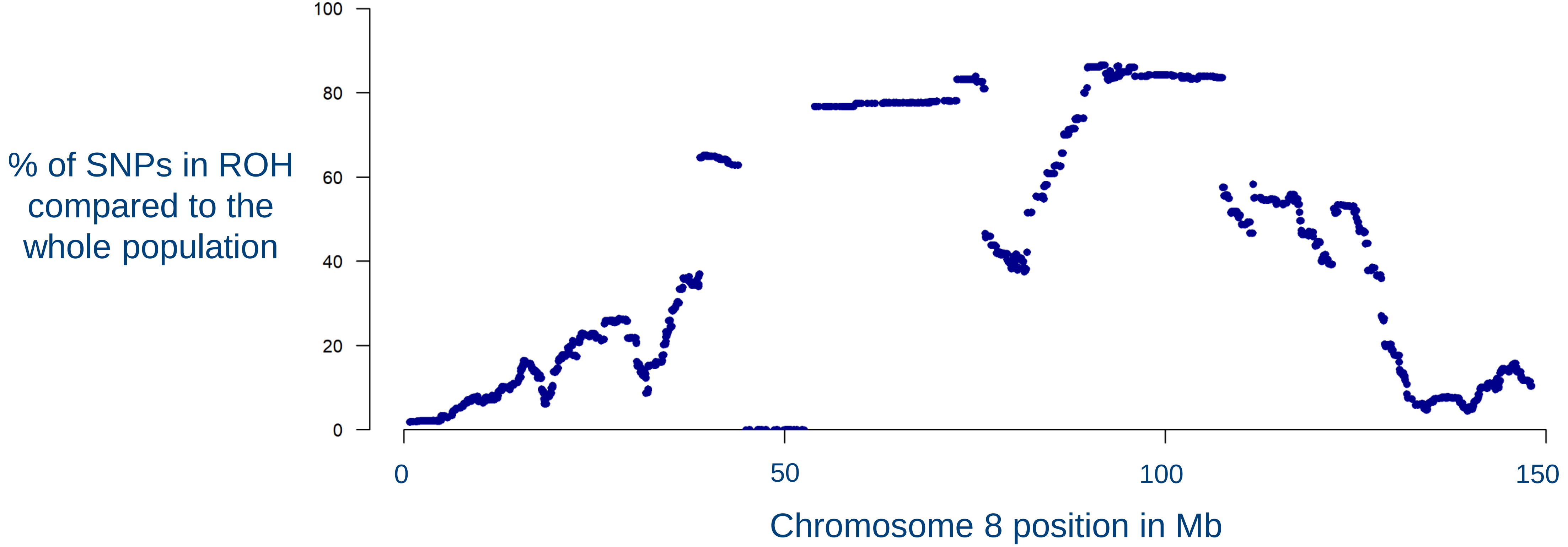
Gorssen et al., submitted

F_{st} heatmap shows USA Piétrains have genetically most diverged from European relatives



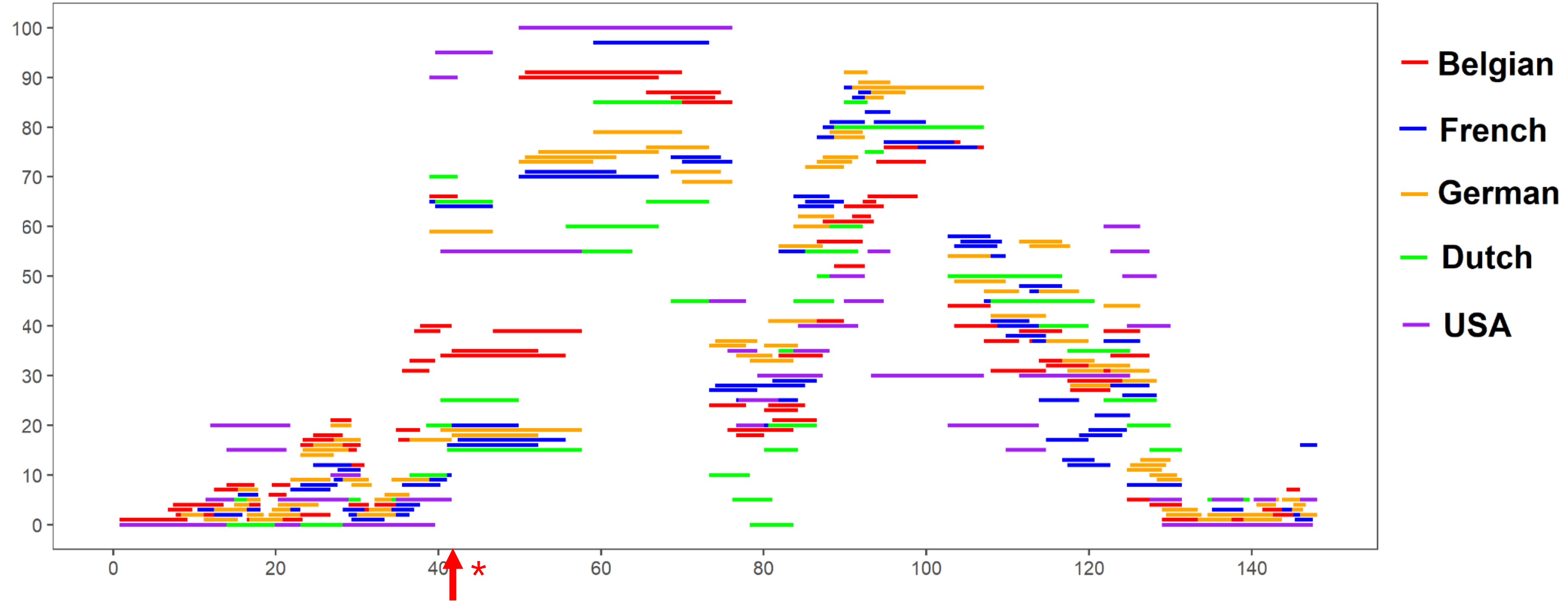
	Belgian	French	German	Dutch
French	0.06			
German	0.04	0.03		
Dutch	0.07	0.06	0.04	
USA	0.10	0.09	0.08	0.07

ROH island analysis on all Piétrains (N=1632) reveals highly inbred regions on SSC8



The majority of Piétrains across all subpopulations showed identical homozygous genotypes on SSC8 from 50-70 Mb and from 90-105 Mb

% Identical 50 SNP homozygous genotype



Chromosome 8 position in Mb

Gorssen et al., submitted; *: location of KIT gene

Take-home messages

The Piétrain shows (weak) genomic differentiation in subpopulations

- Will this differentiation further increase over time?

N_E reasonably high in separate subpopulations and overall at 100

ROH-islands on SSC8 are nearly fixed in all Piétrain animals

- Founder effects during breed formation?
- Coat color selection?
- …

Acknowledgements

Vlaamse Piétrain Fokkerij

- Jürgen Depuydt and Chris Dhondt

INRA, GenPhySE

- Llibertat Tusell Palomero and Catherine Larzul

University of Hohenheim

- Jörn Bennewitz and Patrick Stratz

VPF
PIÉTRAINFOKKERIJ IN EEN NIEUWE DIMENSIE

Flanders
is agriculture and fisheries

Thank you for your attention

wim.gorssen@kuleuven.be

Quality control settings: individual

	Number of genotyped animals					
Country/population	Total	Call rate (<0.91)	Relatedness (>0.95)	Outlying heterozygosity	Removed	Retained
BEL	620	42	1	0	51	577
Flemish	584	41	0	0	41	543
German	13	0	0	0	0	13
Walloon	2	0	0	0	0	2
Flemish*German	10	1	1	0	2	8
Flemish*Austrian	11	0	0	0	0	11
GER	938	7	89	0	95	843
Baden-Württemburg	665	6	88	0	94	571
Nordrhein-Westfalen	196	1	0	0	1	195
Schleswig-Holstein	77	0	0	0	0	77
FRE	173	1	0	0	1	172
Line P1	128	1	0	0	1	127
Line P2	22	0	0	0	0	22
Line P3	23	0	0	0	0	23
NL	20	0	0	0	0	20
USA	20	0	0	0	0	20

Quality control settings: SNP

	BEL	GER	FRE	NL	USA
Total number of SNPs before QC	68516	62163	64232	61772	61772
Mapping SNP-chimp illv2	37264	61177	61565	60835	60835
Call frequency (call rate) (<0.95)	5971	2033	3183	2480	2468
No chromosomal coordinates	2866	7771	7849	7759	7759
SNPs on sex chromosomes	3319	1422	1458	1397	1397
Total SNPs removed during QC	12180	11226	12490	11636	11624
Total SNPs remained after QC	56360	49951	49075	49199	49211
Common autosomal SNPs after QC	34551				

ROH analysis settings

	ROH length category (in Mb)				
	1-2Mb	2-4Mb	4-8Mb	8-16Mb	>16Mb
Allowed heterozygous SNPs (--homozyg-window-het and --homozyg-het)	0	0	0	0	1
Allowed missing SNPs (--homozyg-window-mis)	0	0	1	2	4

- *--homozyg-gap: 1000kb*
- *--homozyg-kb: 1000kb*
- *--homozyg-density 1SNP/150kb*
- *--homozyg-window-snp l*
- *--homozyg-snp l*
  - Where *l* is calculated as

$$l = \frac{log_e \frac{\alpha}{n_s n_i}}{log_e (1 - het)}$$

- Genome coverage was well above 95%

N_E estimation

- SNeP 1.1 software (Barbato et al., 2015)
- Same set of SNPs (34551) was used to compare populations.
- Parameters
  - sample size correction
  - no maf-pruning
  - Sved & Feldman (1973) recombination rate modifier
  - bin distance distribution of 3
  - 30 bins
  - maximum 10 Mb between evaluated SNPs

N_E estimation

Fst analysis

- Hierfstat package in R
- Same set of SNPs (34551) was used to compare populations.
- All pig populations (n>100) from Yang et al. (2017) included

Piétrain distribution according to FAO DAD-IS