

THETA

*Statistical Genetics Group*  
*Institute of Animal Genetics*

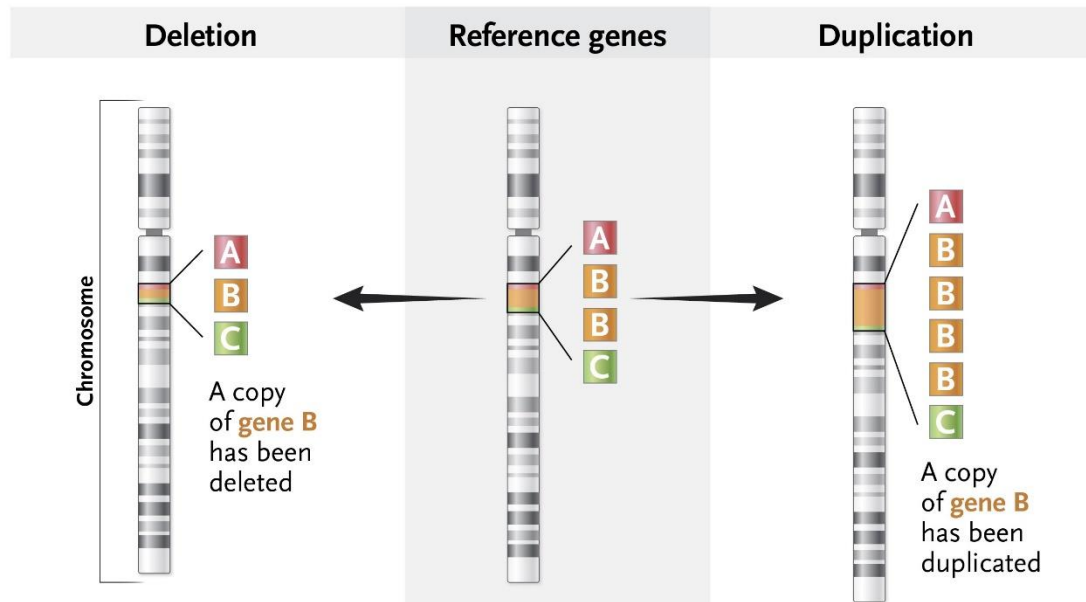


# Impact of copy number polymorphisms on gene expression

M. Mielczarek, M. Frąszczak, J.Szyda

# Background & Objectives

- Genome-wide CNVs detection
- **CNV impact on gene expression**

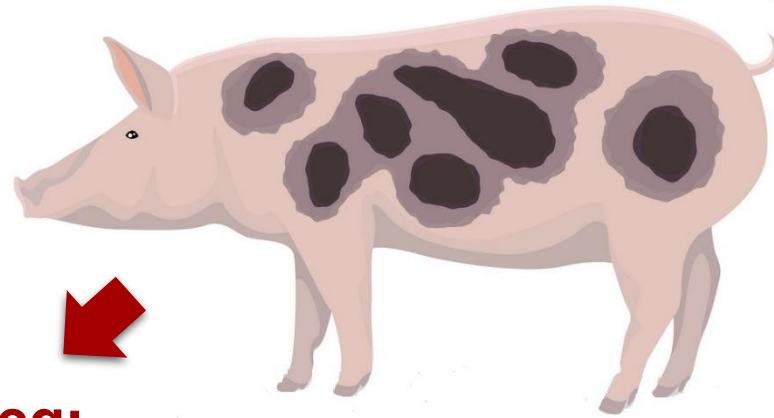


source: <https://illustrated-glossary.nejm.org>

# Dataset

**Accession** → NCBI  
BioProject PRJNA354435

**one** Duroc x Pietrain  
crossbreed pig ♀



## **DNA-seq:**

Illumina HiSeq 2500  
150×2 PE  
average coverage: 15

## **RNA-seq:**

Illumina HiSeq 2500  
150×2 PE

fat:	2 x 93 725 452
liver:	2 x 78 943 463
muscle:	2 x 85 905 149

# Methods

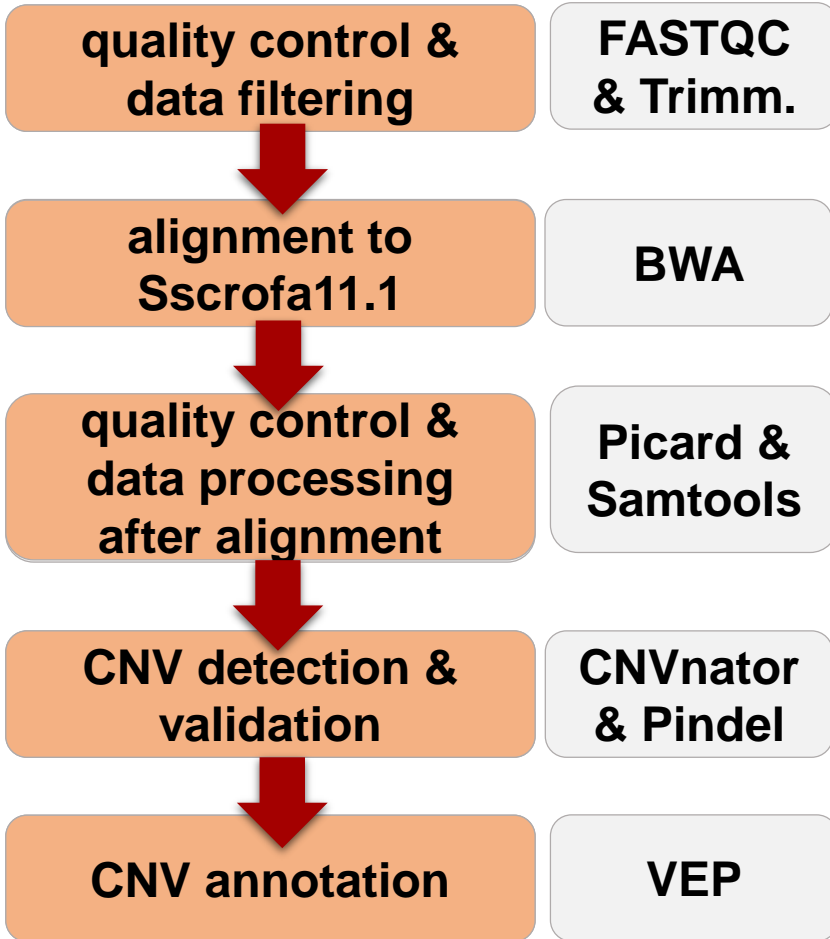
1. CNVs detection (DNA)

2. Transcript quantification (RNA)

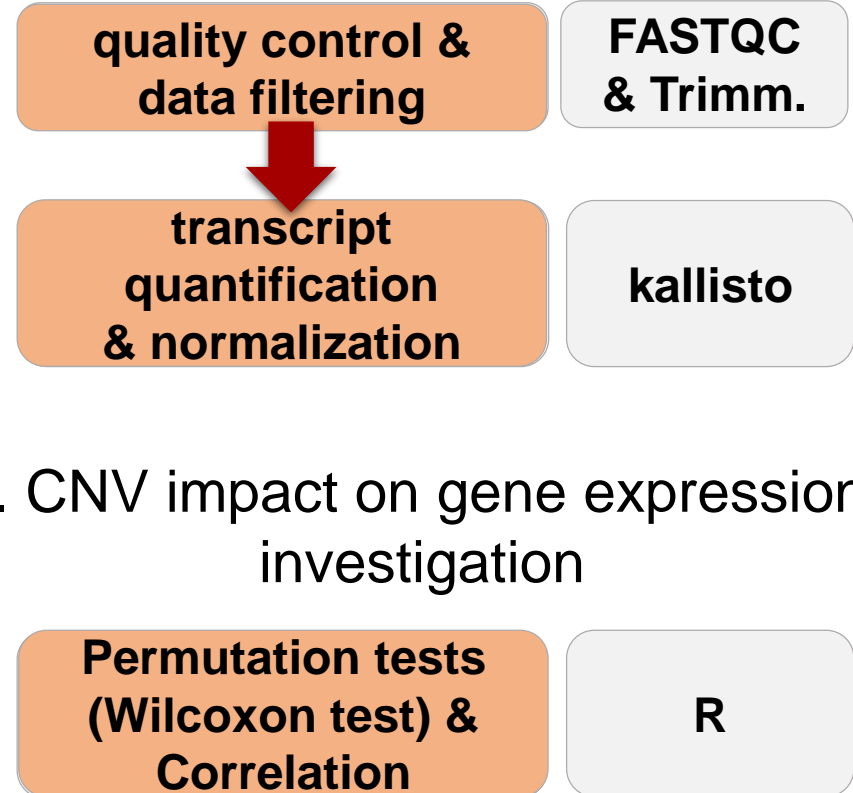
3. CNV impact on gene expression investigation

# Methods

## 1. CNVs detection (DNA)



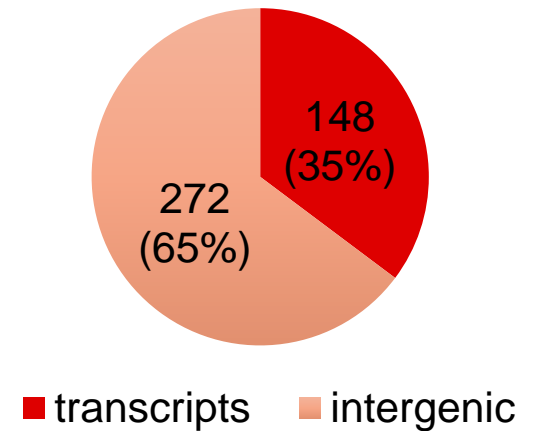
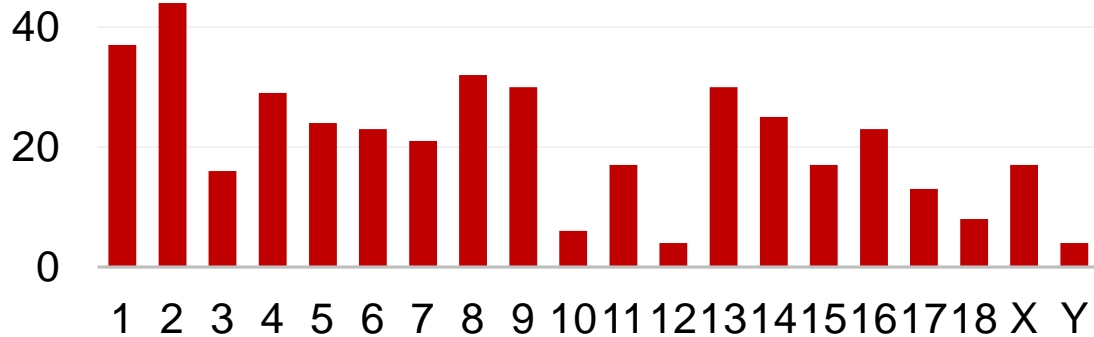
## 2. Transcript quantification (RNA)



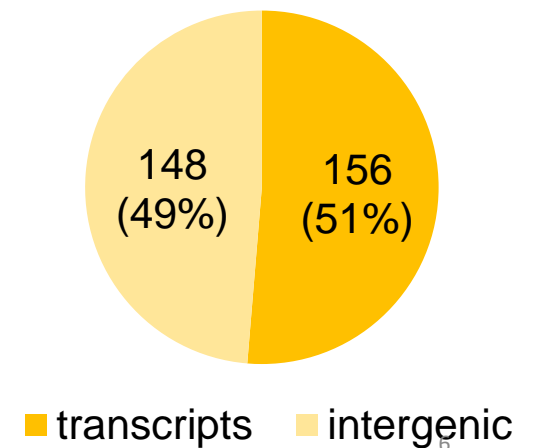
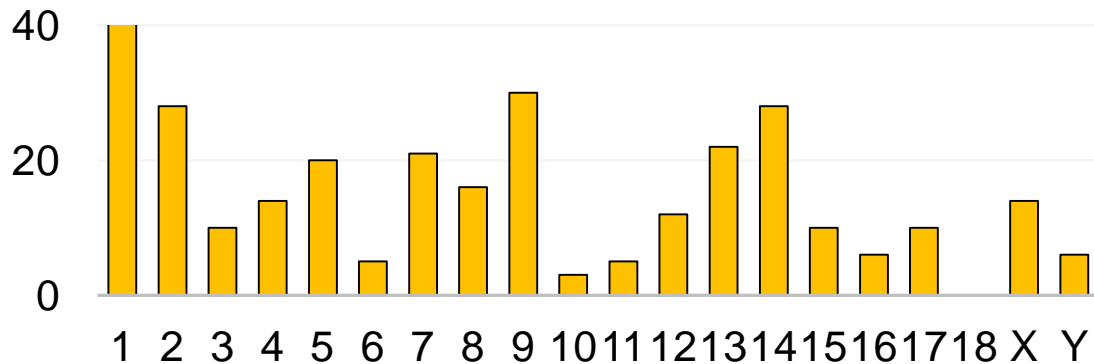
# Results

CNVs characterization → **the number of CNVs**

# deletions = 420

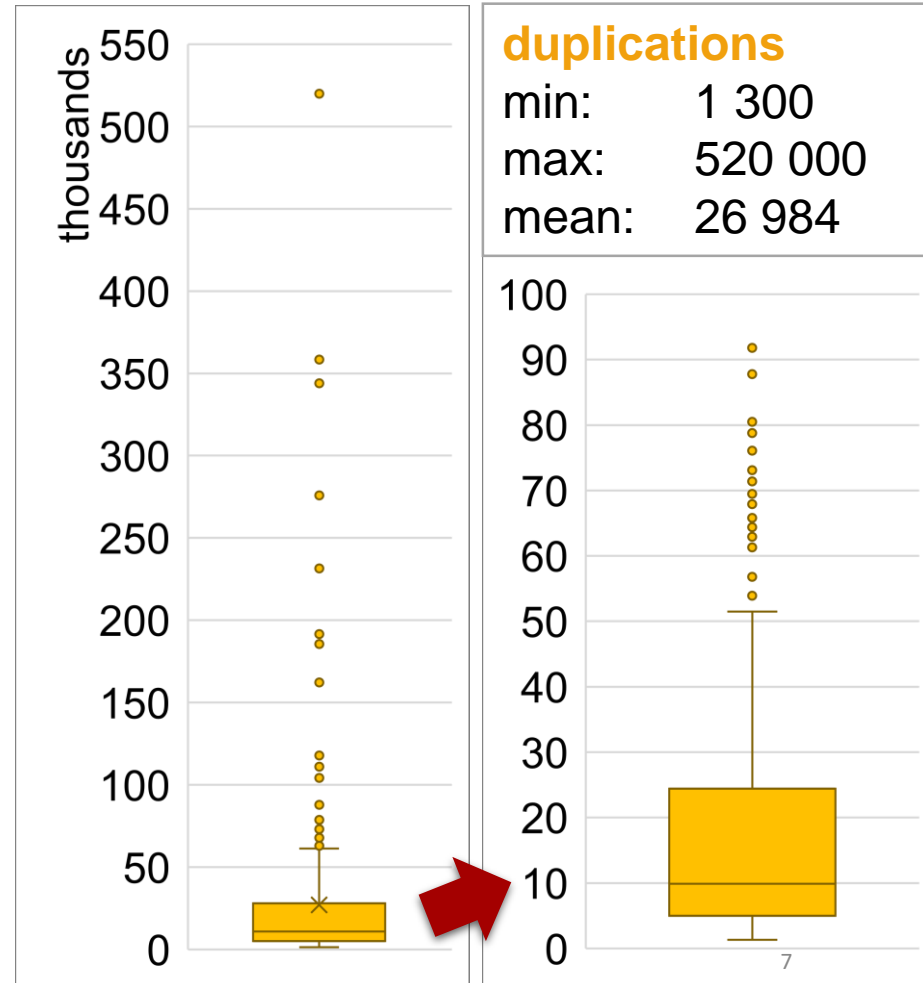
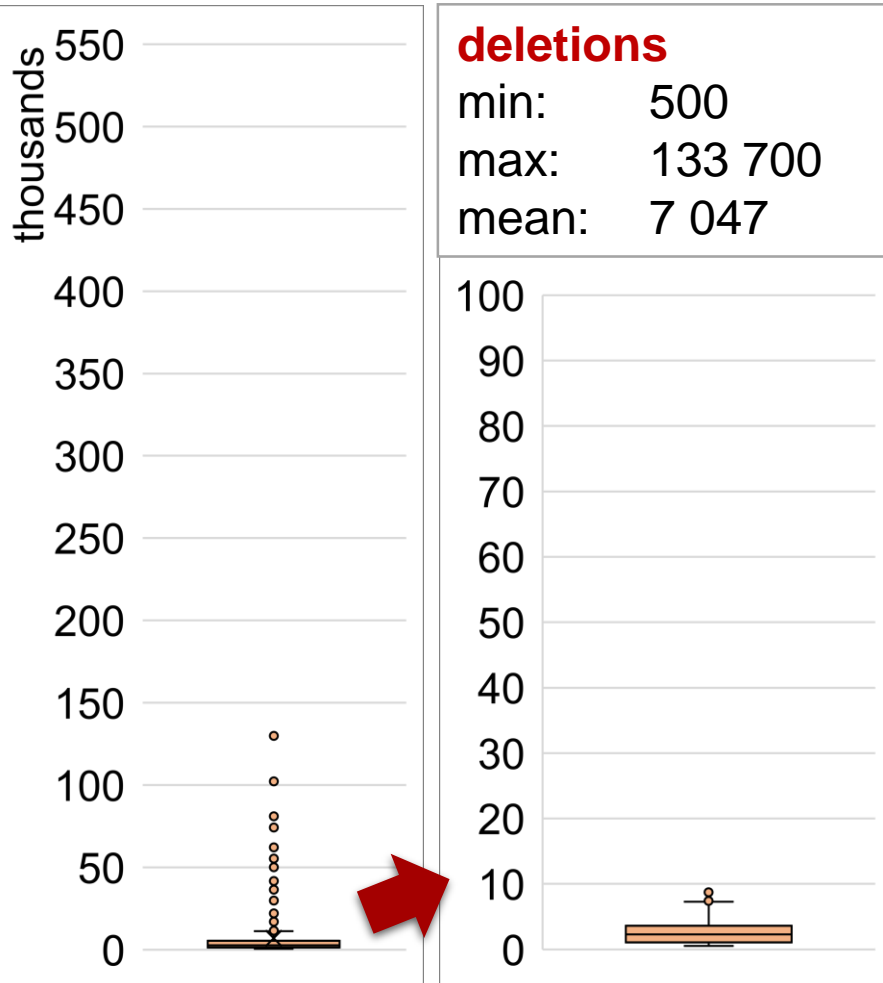


# duplications = 304



# Results

CNVs characterization → **CNVs length (bp)**



# Results

CNVs characterization → **CNVs functional annotation**

Deletions

GO	KEGG	QTL
-	-	-

Duplications

GO (+/-)	KEGG (+)	QTL
<ul style="list-style-type: none"><li>● sensory perception of smell (+)</li><li>● G-protein coupled receptor (+)</li><li>● cellular metabolic processes (-)</li><li>● regulation of metabolic processes (-)</li></ul>	<ul style="list-style-type: none"><li>● pathways of lipid metabolism class</li><li>● olfactory transduction</li><li>● inflammatory mediator regulation</li><li>● carcinogenesis</li><li>● serotonergic synapse</li></ul>	<ul style="list-style-type: none"><li>● production traits and the meat quality</li><li>● skeletal system and bone structure</li><li>● immune system</li><li>● ...and others</li></ul>



# Results

CNVs impact → **CNVs in transcripts**

## Permutation Test

$H_0: E_{\text{CNV}} \geq E_{\text{random}}$

$H_1: E_{\text{CNV}} < E_{\text{random}}$

$E_{\text{CNV}}$  - expression of a transcript with CNV

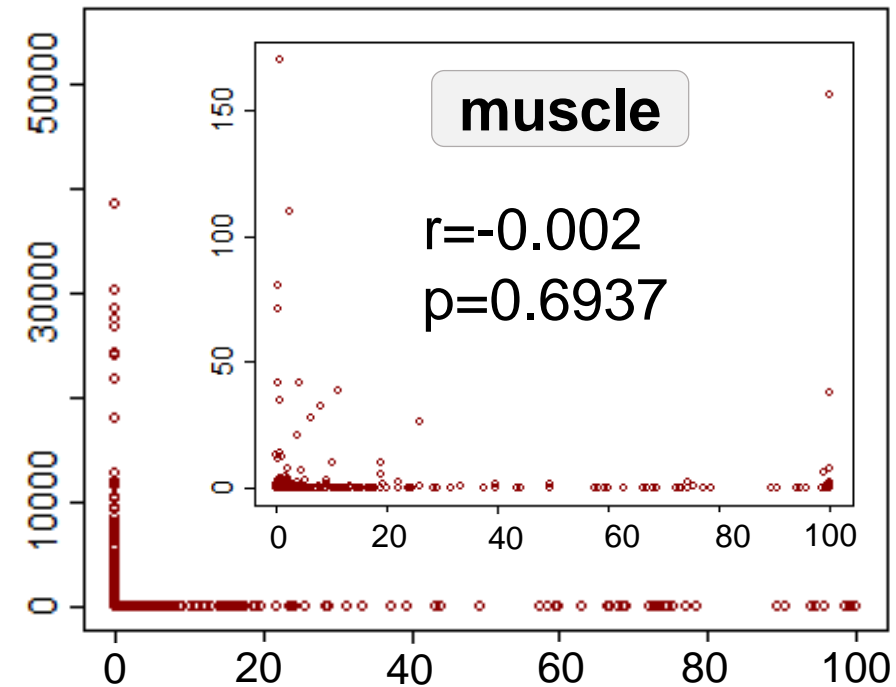
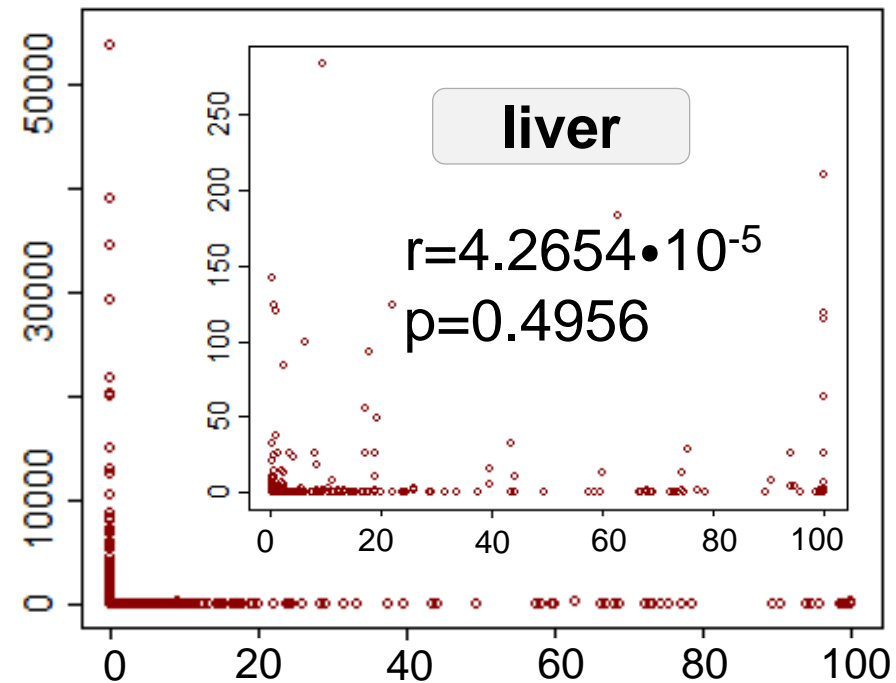
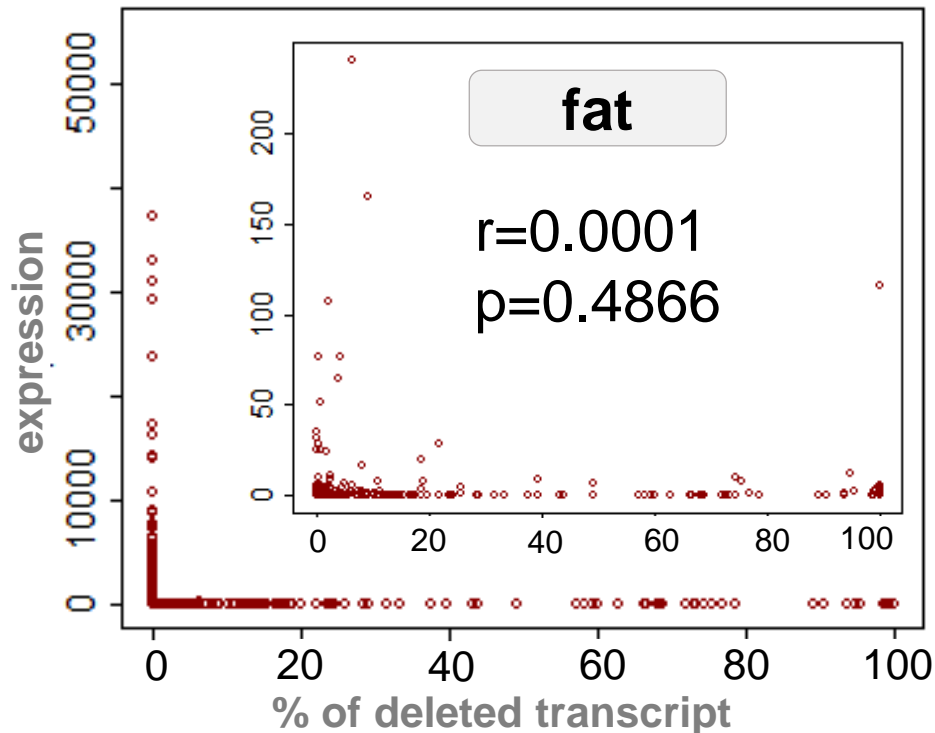
$E_{\text{random}}$  - expression of a random transcript

	Deletions	Duplications
fat	p=0.0008	p<0.0001
muscle	p=0.0010	p<0.0001
liver	p=0.1418	p=0.5908

# Results

CNVs impact →

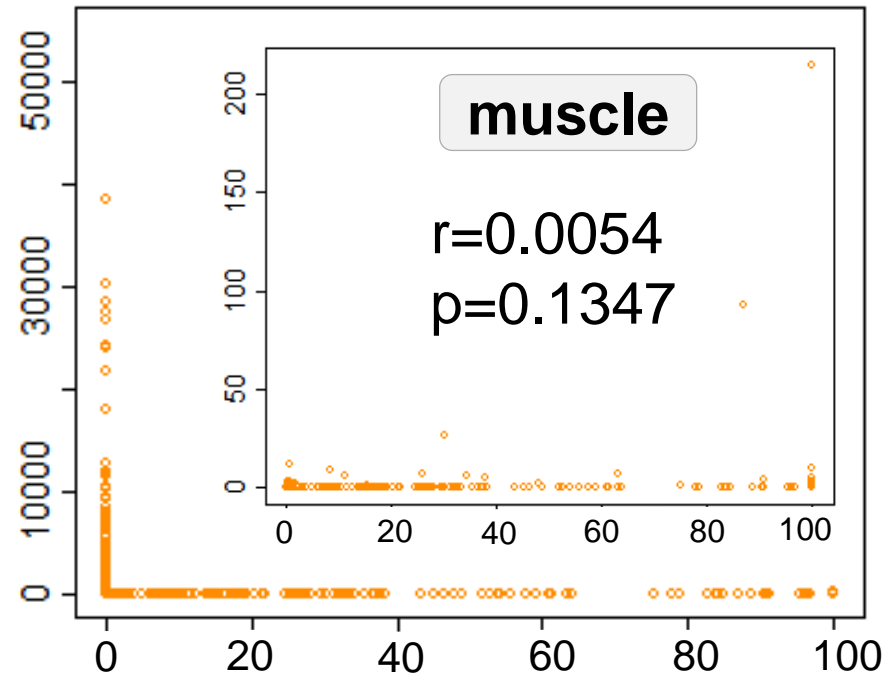
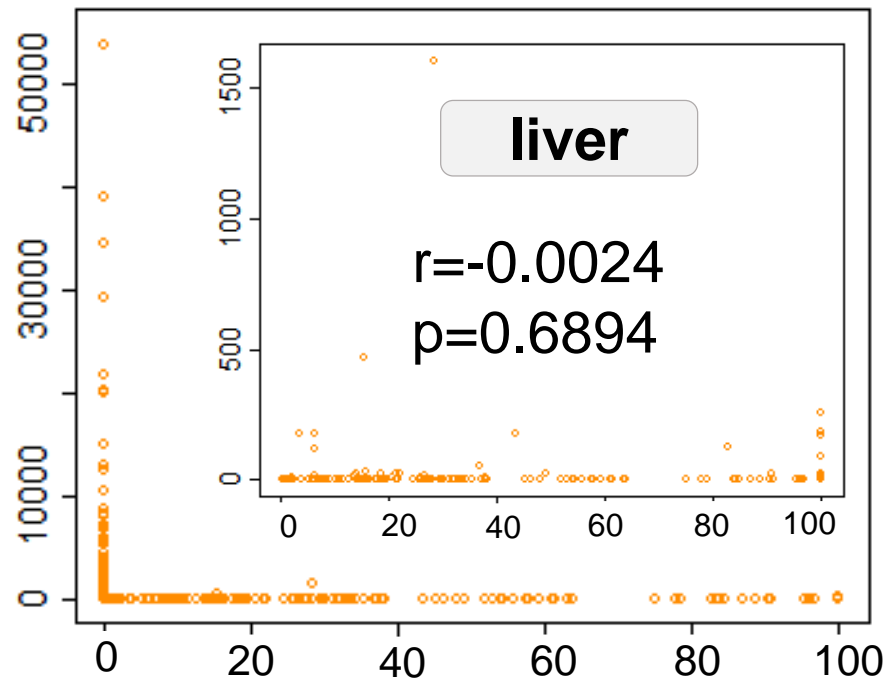
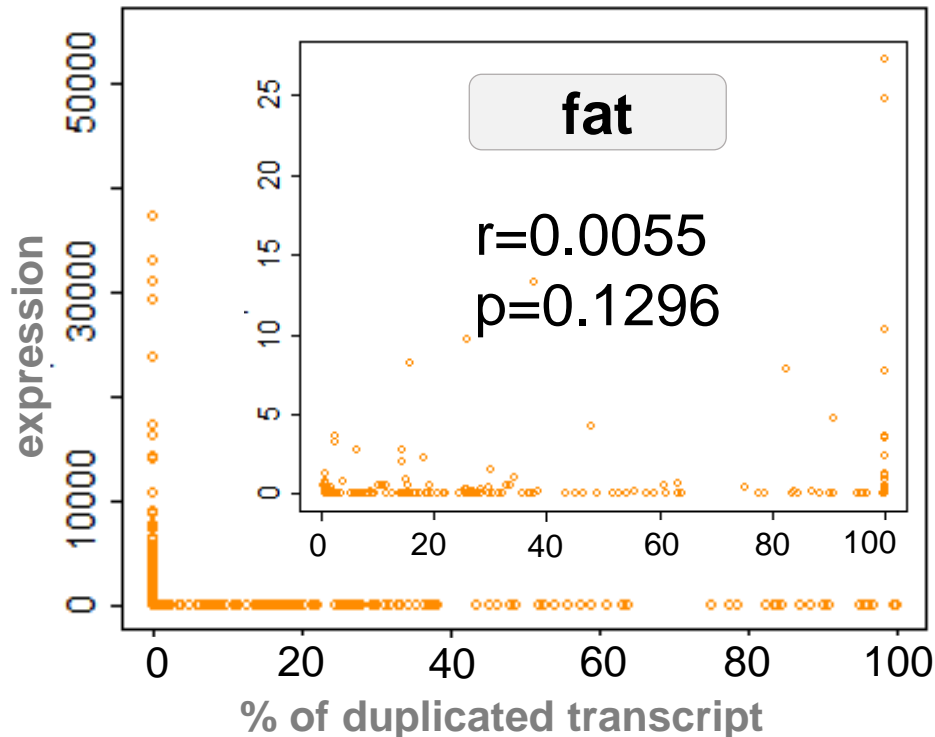
**size of a deleted transcript**



# Results

CNVs impact →

**size of duplicated transcript**



# Summary & conclusions

- Genome: more **deletions** than **duplications**  
**deletions** shorter than **duplications**
- Transcriptome: less **deletions** than **duplications**,  
**deletions** located in „neutral” regions  
„important” regions **duplicated**
- **Deletions** and **duplications** → lower expression in fat and muscle
- No correlation between the size of **deletions/duplications** and expression level

## Future work

- DNA-seq and RNA-seq of 20 male Polish Large White pigs
- Matched by age, sex, breed and environmental factors (diet, housing, slaughter conditions)
- Additional CNVs validation by PCR
- Statistical analysis to provide population-wide inferences

# Acknowledgements

Leading National Research  
Centre in Poland



Poznan Supercomputing  
and Networking Center



Polish National  
Science Centre



**...thank YOU!**