Genomic Insights into Postglacial Expansion Dynamics of the Italian Endemic Amphibian, Bombina pachypus

Sebastiano Fava

DiSVA, *Laboratorio di **Genetica** e **Genomica***   *Tutor: Prof. Emiliano Trucchi*

1. INTRODUCTION AND AIMS

Geographic range expansion impacts **population dynamics and thereby influences biodiversity distribution**

Postglacial range expansion of the Italian Apennine yellow-bellied toad, *Bombina pachypus*

Our goal is to assess **genetic diversity**, **genetic load gradients** along the expansion route, and signs of **natural selection**

10 GB genome

To reduce sequencing costs and not lose genomic information we will focus on 24Mb target regions.

Target genomic regions will be selected to represent four categories with potentially **different functional effects**:
1. **High effect region:** low expression exons
2. **Low effect region:** high expression exons
3. **Regulatory region**
4. **Neutral Intergenic regions**

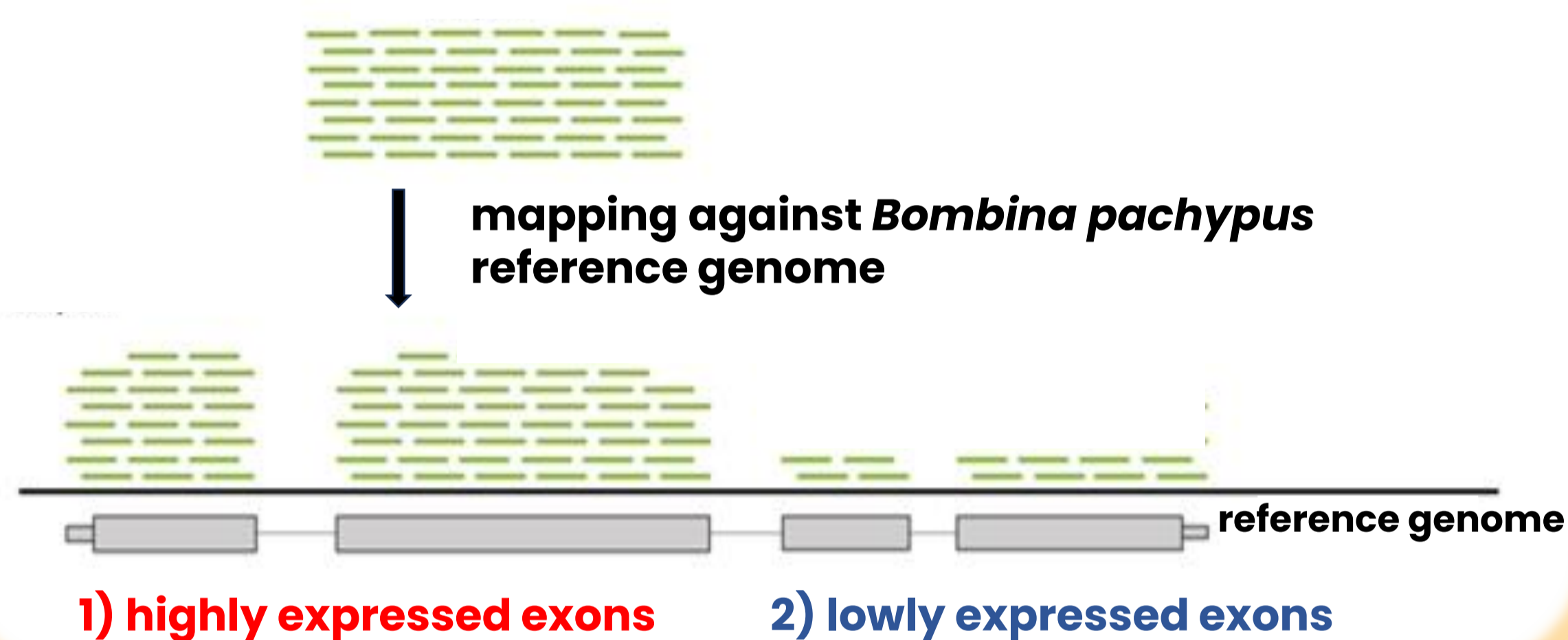

Figure 1: *In yellow: B. pachypus range. Black arrow: B. pachypus post-glacial expansion route from core to front.*

FRONT

CORE

< ENDANGERED >
EN

2. PRELIMINARY MATERIAL AND METHODS

Definition of high- and low-expression exons

mRNA extracted from 3 tissues of 6 individuals and sequenced

mapping against *Bombina pachypus* reference genome

reference genome

1) highly expressed exons   **2) lowly expressed exons**

- **mRNA was extracted** from **16 samples** belonging to **6 individuals** (**7 brain, 7 gonads, 2 bulk**) and **sequenced** using the Illumina platform.

- **The sequenced RNA** was **mapped** to the **reference genome** using STAR software.

- **Reads mapped to each exon** of each previously annotated gene were **counted** using HTSEQ software.

- Read counts per exon were **separated by tissue type, filtered, and normalized to TPM** (transcripts per million).

- **The average expression for each exon in each tissue** was calculated and **divided into five quantiles**.

- Exons in the **first quantile** of average expression were classified as **low expression exons**, while those in the **fifth quantile** were classified as **high expression exons**.

- **The lengths and coordinates of exons** with **low** and **high expression** were then **extracted across all three tissues**

REFERENCES

- Peischl, S., Dupanloup, I., Kirkpatrick, M., & Excoffier, L. (2013). On the accumulation of deleterious mutations during range expansions. *Molecular ecology, 22*(24), 5972-5982.
- Burton, O. J., & Travis, J. M. (2008). The frequency of fitness peak shifts is increased at expanding range margins due to mutation surfing. *Genetics, 179*(2), 941-950.
- Yang, J. R., Liao, B. Y., Zhuang, S. M., & Zhang, J. (2012). Protein misinteraction avoidance causes highly expressed proteins to evolve slowly. *Proceedings of the National Academy of Sciences, 109*(14), E831-E840.
- Wu, Z., Cai, X., Zhang, X., Liu, Y., Tian, G. B., Yang, J. R., & Chen, X. (2022). Expression level is a major modifier of the fitness landscape of a protein coding gene. *Nature Ecology & Evolution, 6*(1), 103-115
- Dobin, A., & Gingeras, T. R. (2015). Mapping RNA-seq reads with STAR. *Current protocols in bioinformatics, 51*(1), 11-14.
- Putri, G. H., Anders, S., Pyl, P. T., Pimanda, J. E., & Zanini, F. (2022). Analysing high-throughput sequencing data in Python with HTSeq 2.0. *Bioinformatics, 38*(10), 2943-2945.
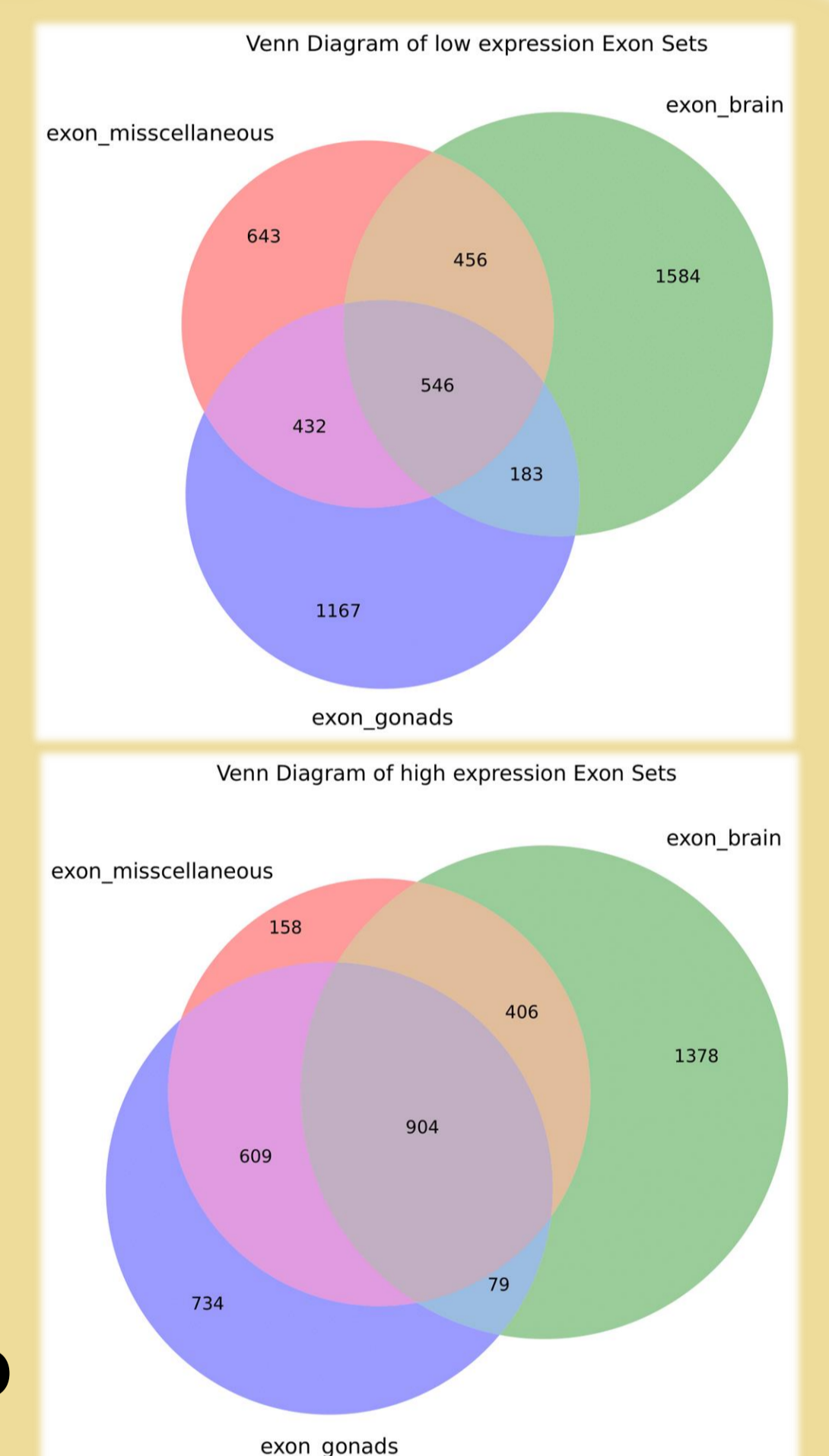
3. PRELIMINARY RESULTS

Total number of **low espression** exon: **5011**

Total length sum of **low expression** exons: **5 Mb**

Total number of **high espression** exon: **4268**

Total length sum of **high expression** exons: **2,8 Mb**

Venn Diagram of low expression Exon Sets
exon_misscellaneous — exon_brain — exon_gonads
643 | 456 | 1584 | 432 | 546 | 183 | 1167

Venn Diagram of high expression Exon Sets
exon_misscellaneous — exon_brain — exon_gonads
158 | 406 | 1378 | 609 | 904 | 79 | 734

4. NEXT STEPS

- **Combination of high- and low-expression exonic regions with regulatory sequences and neutral intergenic regions inferred from ATAC-seq data**.

- **Synthesis of 24 Mb DNA target enrichment probes**, by Agilent , SureSelect company.

- **Capture** and **sequencing of genomic regions in approximately 300 individuals** along the expansion route.

- **Alignment of sequencing data to the reference genome with** BWA mem software.

- **SNP callig** and **annotation of synonymous, missense, or nonsense SNPs** using GATK HaplotypeCaller and SNPEff software.

- **Estimation of genetic diversity** and **load gradient along the expansion route**