

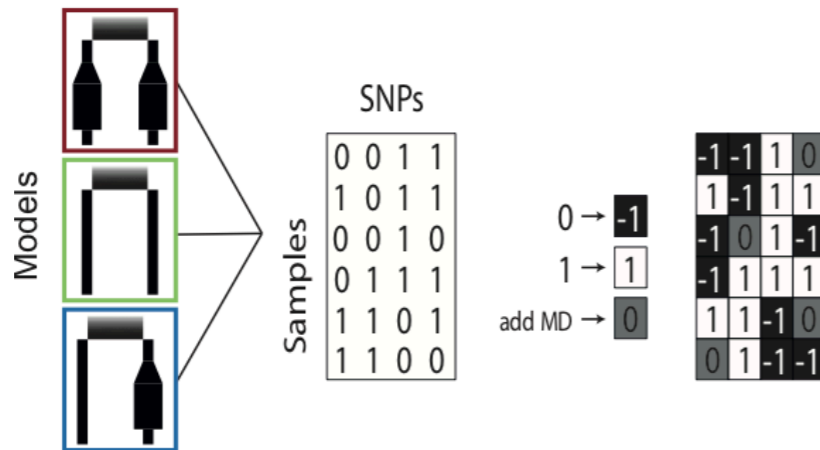
Redes neuronales como alternativa a modelos probabilísticos en inferencia filogenética

Isabel Sanmartín, Real Jardín Botánico

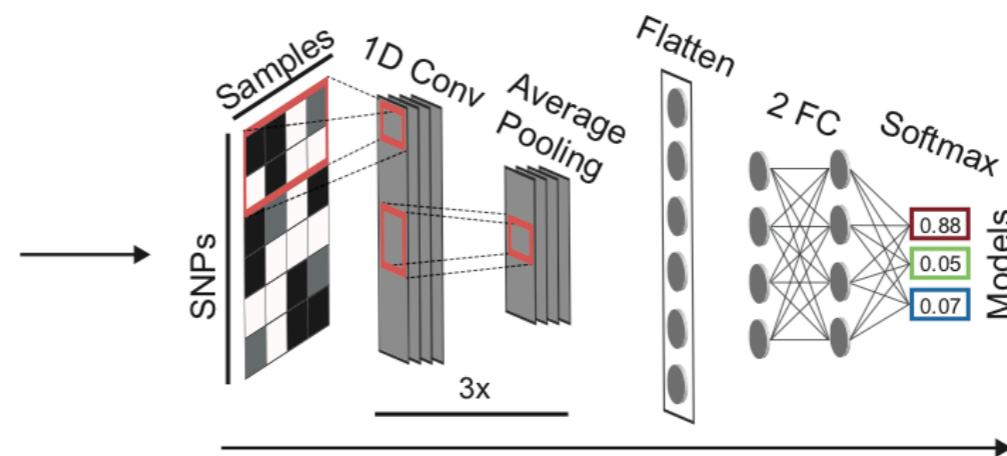
- **Phylogenetic inference** is considered a **non-polynomial time problem**: e.g., the set or property of problems for which no polynomial-time algorithm is known and for which the only NP-algorithms require a number of steps which grow exponentially with the size of the problem (e.g., MCMC inference).
- Many commonly used models of evolutionary inference have issues with **non-identifiability**: different model parameters induce the same probability distributions and are mathematically or “practically” undistinguishable.
- **Deep learning neural networks** have been proposed as an alternative to likelihood-based approaches to tackle these issues.

Deep learning in microevolution

1) Convert simulated SNPs to image



2) Train CNN with images from all models

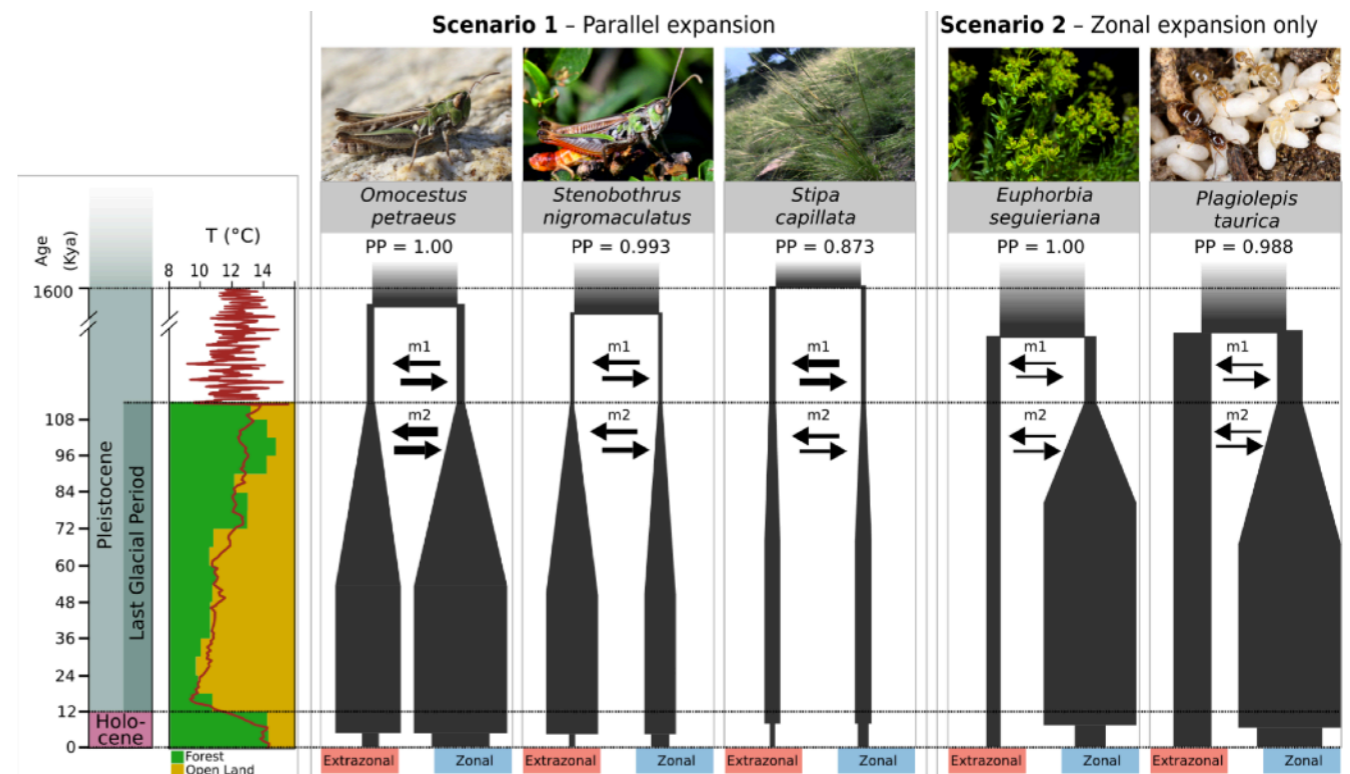


Convolutional Neural Network (CNN)

trained with simulated single-site polymorphism (SNP) genomic data.

Replaces summary statistics (ABC): CNN uses **image pattern recognition** to capture site variance/covariance patterns in the data.

Higher discriminatory power between “non-identifiable” demographic models: populations-size bottleneck vs. positive selective sweep; incomplete lineage sorting vs. secondary contact + migration, etc.

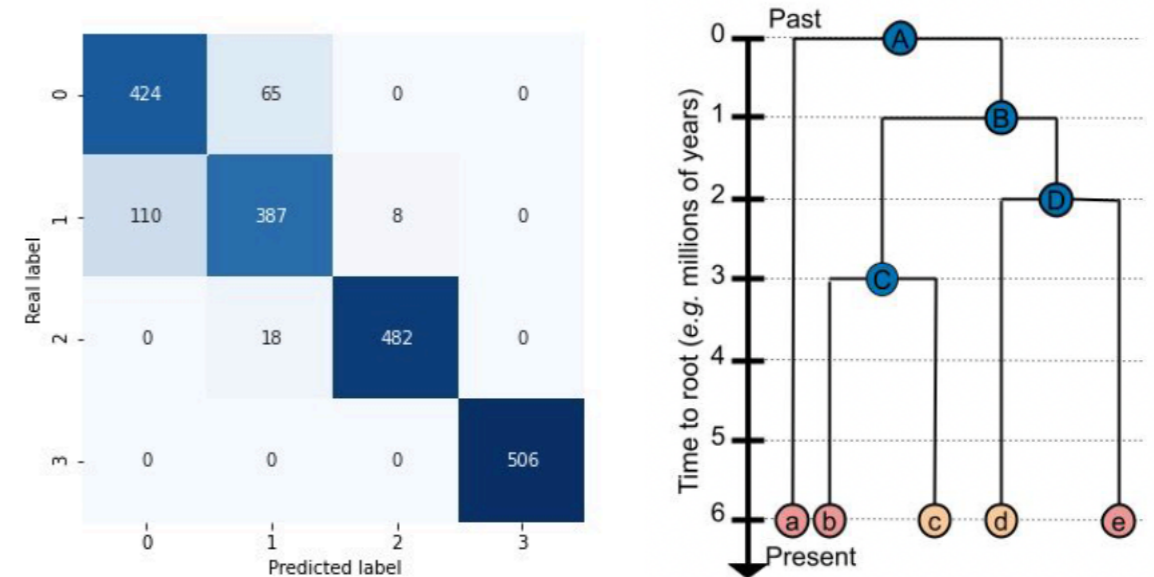
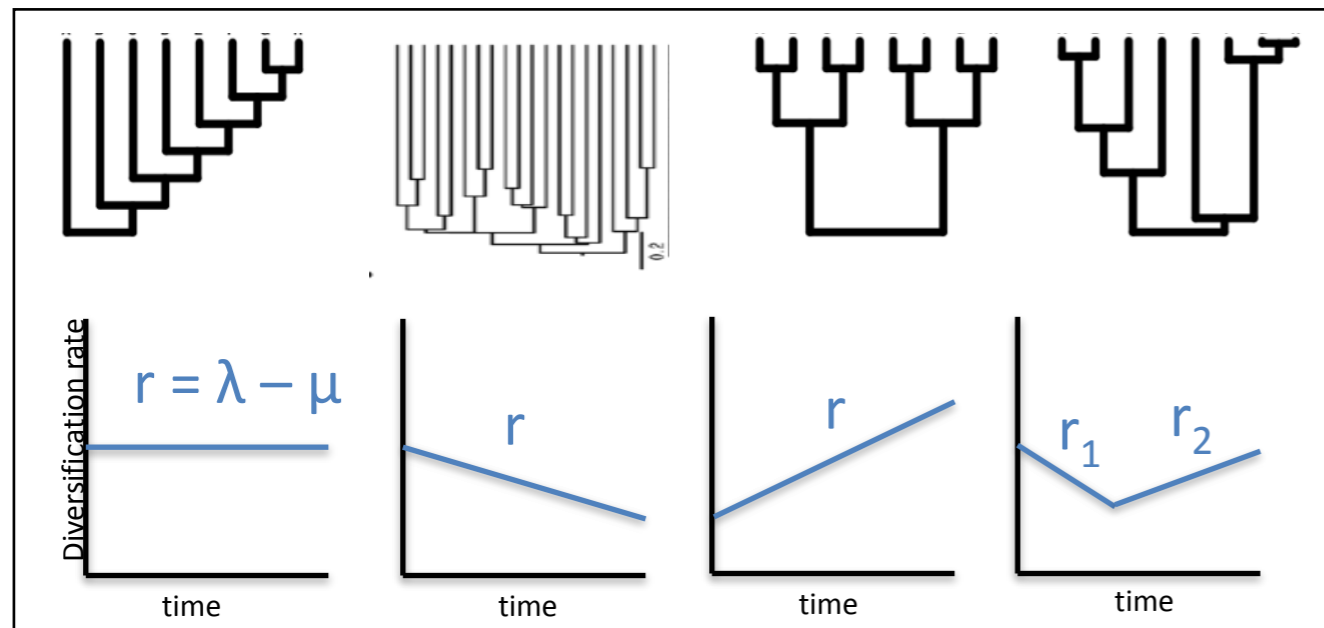


Deep Learning in macroevolution

Convolutional neural networks (CNN) trained with simulated diversification scenarios: time trees.

Instead of image recognition, trees are decomposed into **vector information** (nodes, tips, tree height) to capture topology and branch lengths.

Higher discriminatory power between “non-identifiable” models of diversity generation: evolutionary stasis + accelerated speciation vs. mass extinction; exponentially increasing extinction vs. diversity-dependence, etc.



	precision	recall	f1-score	support
0	0.79	0.87	0.83	489
1	0.82	0.77	0.79	505
2	0.98	0.96	0.97	500
3	1.00	1.00	1.00	506
accuracy			0.90	2000
macro avg	0.90	0.90	0.90	2000
weighted avg	0.90	0.90	0.90	2000