# P8

# Integrative Bioinformatic Tools for GRN Reconstruction

GönomiX

## State of the art

Gene regulatory networks (GRNs) are essential for understanding gene expression regulation. Bioinformatics methods, including machine learning (e.g., SpaCeNet, Ensemble-GNN) and multi-omics integration (e.g., ATAC-seq, RNA-seq), have enhanced GRN reconstruction. Predicting transcription factor binding sites and leveraging databases (e.g., Reactome, ENCODE) further improve accuracy.

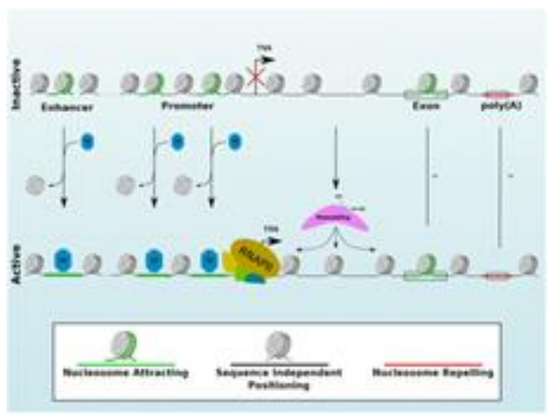


**Figure 1: Regulatory Element prediction from DNA Structure**
*Combing Structure information can make prediction much more accurate (Sahrhage et al. 2024).*

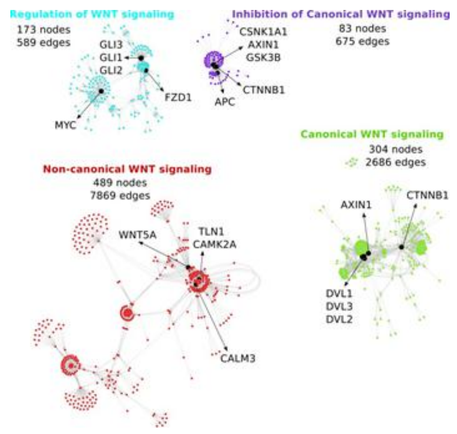


**Figure 2: WNT Signaling in breast cancer**
Prior knowledge from literature, other information is still crucial (Bayerlova et al 2015)

## Primary Questions

- How to reconstruct GRNs from *Omics*-data of multiple model-species?

- Methods and GRN information are mainly available for model species (*Mus musculus, Drosophila, Arabidopsis*)
- Transfer to non-model species is essential and facilitated by databases like OrthoDB and ENCODE

## Objectives

- Develop and apply bioinformatics methods for GRN reconstructing (open chromatin, transcription factor binding sites, *multi-omics* data, for the WNT signaling pathway)
- Reconstruction of GRNs using expression profile correlations, time-series, and perturbation data to capture dynamic regulatory interactions
- Refinement of GRN models through cross-species comparisons - identify conserved and divergent regulatory mechanisms
- Application and extension of GRN tools to GönomiX data

## Workplan

### A) Developing and Validating Bioinformatics Methods for GRN-Reconstruction and Cross Species Analysis of WNT Signaling Modules

- Developing and testing bioinformatic tools using publicly available datasets
- Enhancing WNT modules with bioinformatic predictions of gene regulatory elements (BERT modules) transferring to different species
- Constructing ML models to predict gene expression based on DNA sequence and ATAC-seq data
- Validating models with data from the consortium (e.g. *Tribolium*, Bucher lab)

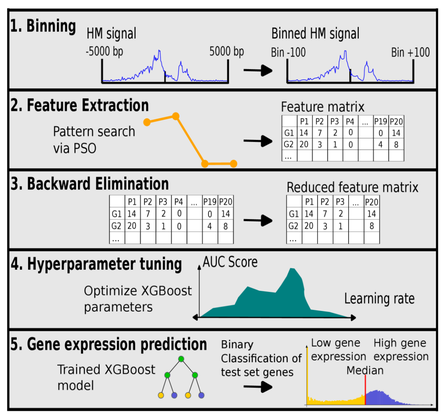


**Figure 3:** Example workflow for predicting gene expression data, from promoter structure Information (Paul et al. submitted)

B) Network reconstruction based on challenged networks
Use of probabilistic GRN reconstruction models for Graphical Gaussian Models

- GRN reconstruction will use probabilistic methods (e.g., Schrod et al., 2024a) to build GGMs, addressing co-expression limits in identifying interaction direction.
- RNAi knock-down data (collaboration Bucher lab) will capture dynamic responses, enhancing insights into regulatory interactions.

### C) Refining and Applying Cross-Species Network Reconstruction Methods using GönomiX Data

- Analyze multi-species RNA-seq, ATAC-seq, and single-cell data to improve models.
- Use evolutionary insights and experimental validation to refine models iteratively.

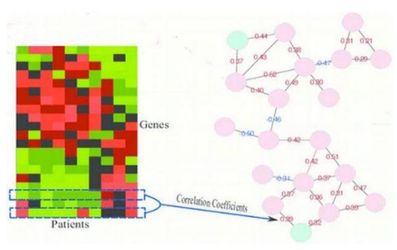


**Figure 4:** Principle of reconstructing GRNs based on correlation.

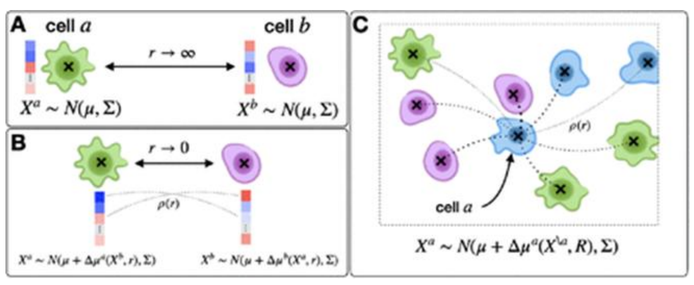


**Figure 5:** Example Workflow for GRN reconstruction using GGMs (Schrod et al, 2024)

D) Model development and Tools

- Developing an Integrated Cross-Species Wnt GRN Model
- Develop advanced bioinformatics tools for GRN reconstruction using RNA-seq, ATAC-seq, and single-cell sequencing data

## Synergy and collaborations

- **Collaborative Project 1: Wnt signaling in anterior development**
  Comparision of direct WNT target genes and GRNs across species through coordinated RNA-seq and ATAC-seq time course experiments
- **Collaborative Project 2: Reconstructing evolving GRNs**
  Analyzing diverse omics data generated across all particular projects (RNA-seq, ATAC-seq, chromatin structure, single-cell RNA-seq). Reconstructing evolving gene regulatory networks by focusing on coordinated data gathering the anterior gene regulatory network in various species. Moreover:
- **Collaborative Project 3: Novel bioinformatics and genetic tools**
  Development of tools for bioinformatic analyses across clades; Enhancment of cross-species comparision methods (in Collaboration with Project Partner 9)

## Technical innovation

- Development of sophisticated bioinformatics tools for GRN reconstruction
- Integration of predictions of regulatory elements and facilitate cross-species GRN comparisons
- enhance study of WNT signaling pathways

## Specific qualification

- Bioinformatics; Machine Learning; High-Throughput Omics Data Analysis; Systems Biology; Programming of Tools and Workflows.

**Tim Beißbarth, Medical Center University**

## References

1. Sahrhage M, Paul NB, **Beißbarth T**, Haubrock M. The importance of DNA sequence for nucleosome positioning in transcriptional regulation. *Life Sci Alliance, 2024, 7(8):e202302380.*
2. Bayerlová M, …, **Beißbarth T**, Bleckmann A. Newly Constructed Network Models of Different WNT Signaling Cascades Applied to Breast Cancer Expression Data. *PloS One, 2015, 10:e0144014.*
3. Paul NB, …, **Beißbarth T**, Haubrock M. Prediction of gene expression using histone modification patterns extracted by Particle Swarm Optimization. *Bioinformatics, 2025, accepted.*
4. Schrod S, …, **Beißbarth T**, …, Altenbuchinger M. Spatial Cellular Networks from omics data with SpaCeNet. *Genome Res, 2024, 34(9):1371-1383*.
5. Schrod S, … **Beißbarth T**, Hauschild AC, Altenbuchinger M. CODEX: COunterfactual Deep learning for the in silico EXploration of cancer cell line perturbations. *Bioinformatics, 2024, 40(Suppl 1):i91-i99.*
6. Sitte M, …, **Beißbarth T**. Reconstruction of Different Modes of WNT Dependent Protein Networks from Time Series Protein Quantification. *Stud Health Technol Inform, 2019, 267:175-180.*

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN IN PUBLICA COMMODA SEIT 1737

MAX PLANCK INSTITUTE FOR MULTIDISCIPLINARY SCIENCES

UNIVERSITÄTSMEDIZIN GÖTTINGEN UMG Göttingen Campus