



FROM ANCESTRY TO ORIENTATION: CRISPR ARRAY TRANSCRIPTIONAL DIRECTION PREDICTION FROM AN EVOLUTIONARY PERSPECTIVE

Axel Fehrenbach*, Alexander Mitrofanov¹, Omer Alkhnbashi¹, Rolf Backofen¹, Franz Baumdicker*

*University of Tübingen, Cluster of Excellence Controlling Microbes to Fight Infections & Cluster of Excellence Machine Learning for Science

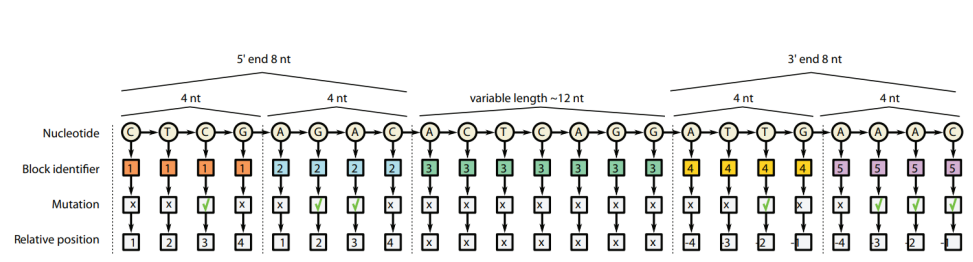
¹ University of Freiburg, Department of Computer Science

Summary:

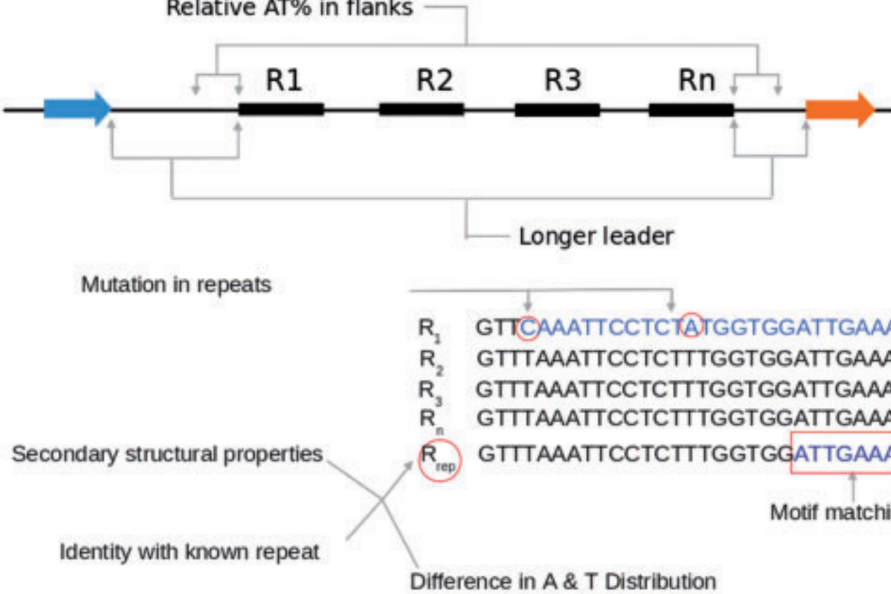
- CRISPR arrays provide a history of acquired spacers.
- We can estimate the CRISPR array orientation from multiple spacer arrays based on an evolutionary model.
- Direction prediction differs between CRISPR types.

So far direction prediction is repeat based, but...

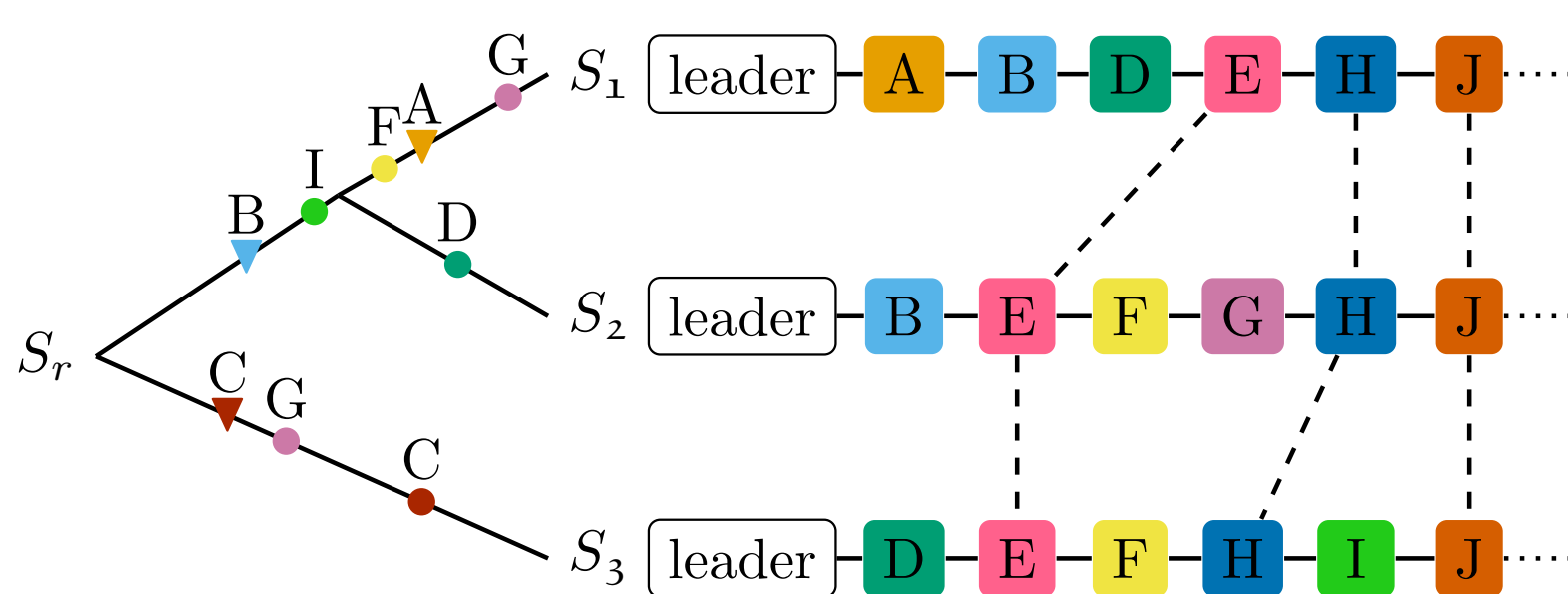
CRISPRstrand [1]



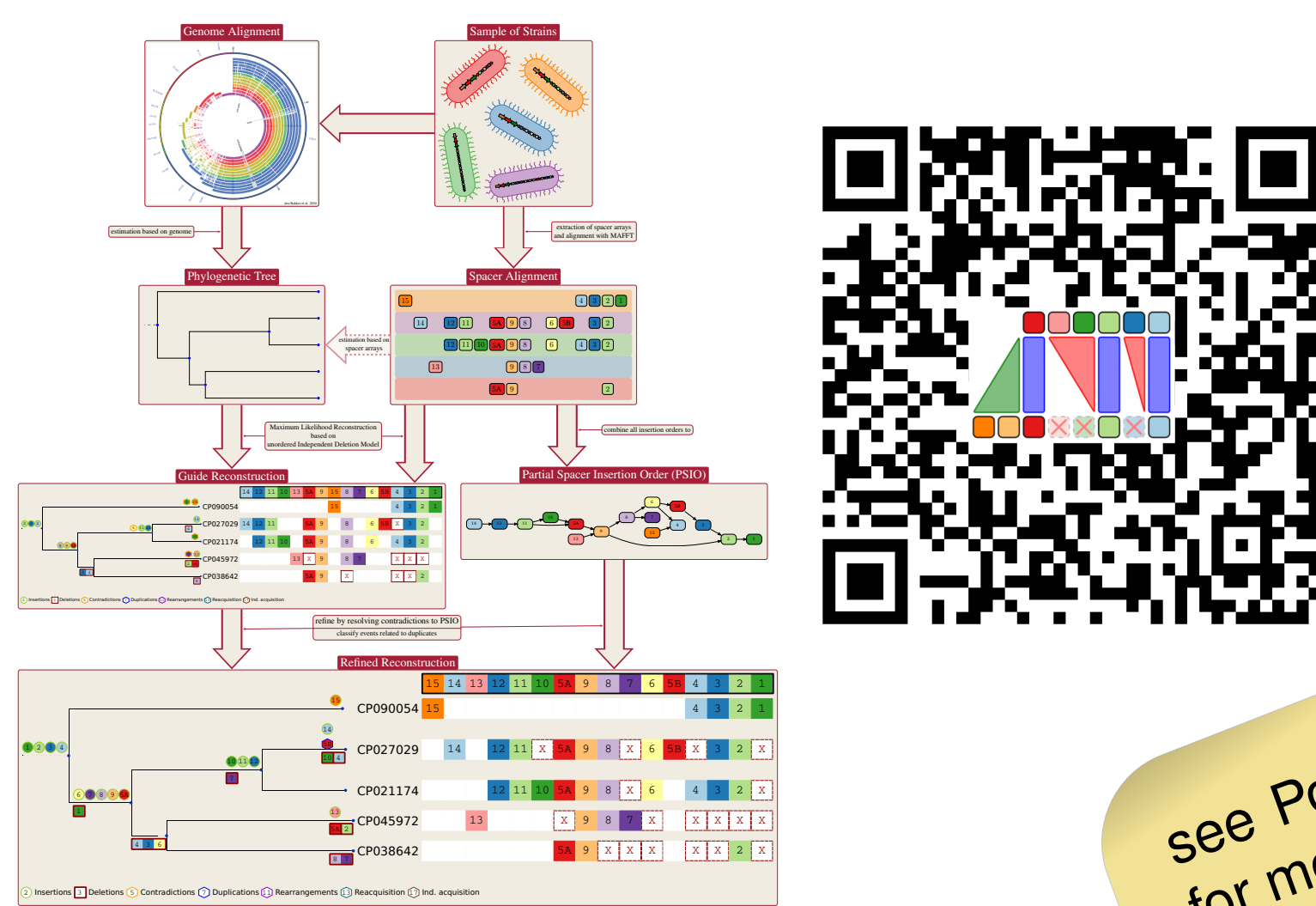
CRISPRdirection [2, 3]



... we can also use evolutionary patterns of CRISPR arrays



Reconstruct ancestral spacer acquisitions & deletions



Results and Conclusions

- We can predict the orientation of CRISPR arrays based on the reconstructed ancestral events from multiple spacer arrays.
- CRISPR-evo-inator is independent of repeat based orientation tools like *CRISPRDirection* [3] and *CRISPRstrand* [1].
- Increases the number of arrays with known orientation.
- Type I-F directions are correct in CRISPRCasdb.
- Type I-E, I-B, and II-A directions can be misleading.

Data

- CRISPR arrays extracted from *CRISPRdb* [4]
- Clustered by species and Cas type into 337 groups
- 5192 arrays (2327 are unique).

Contact

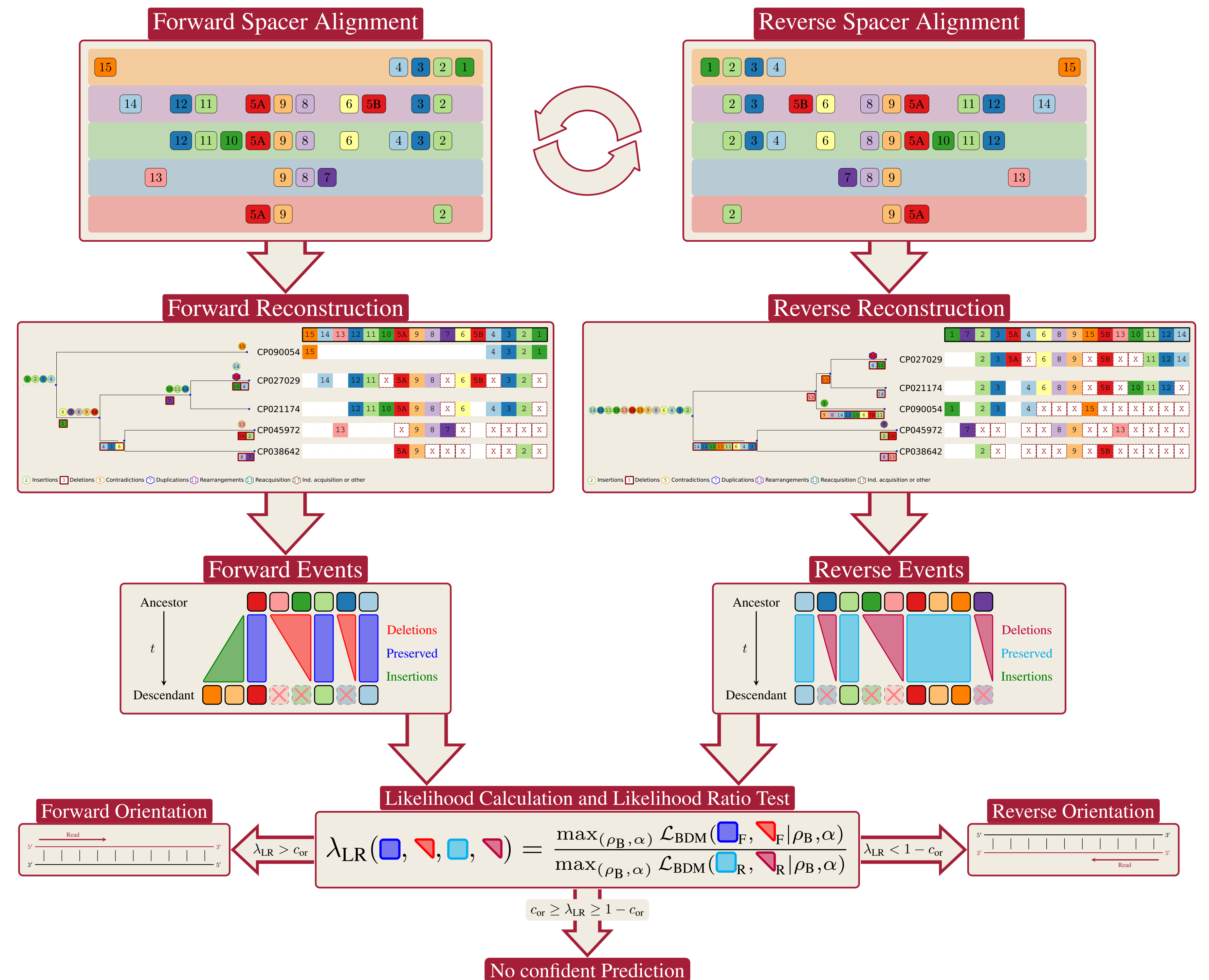
franz.baumdicker@uni-tuebingen.de
axel.fehrenbach@uni-tuebingen.de
www.baumdickergroup.de
fbaumdicker



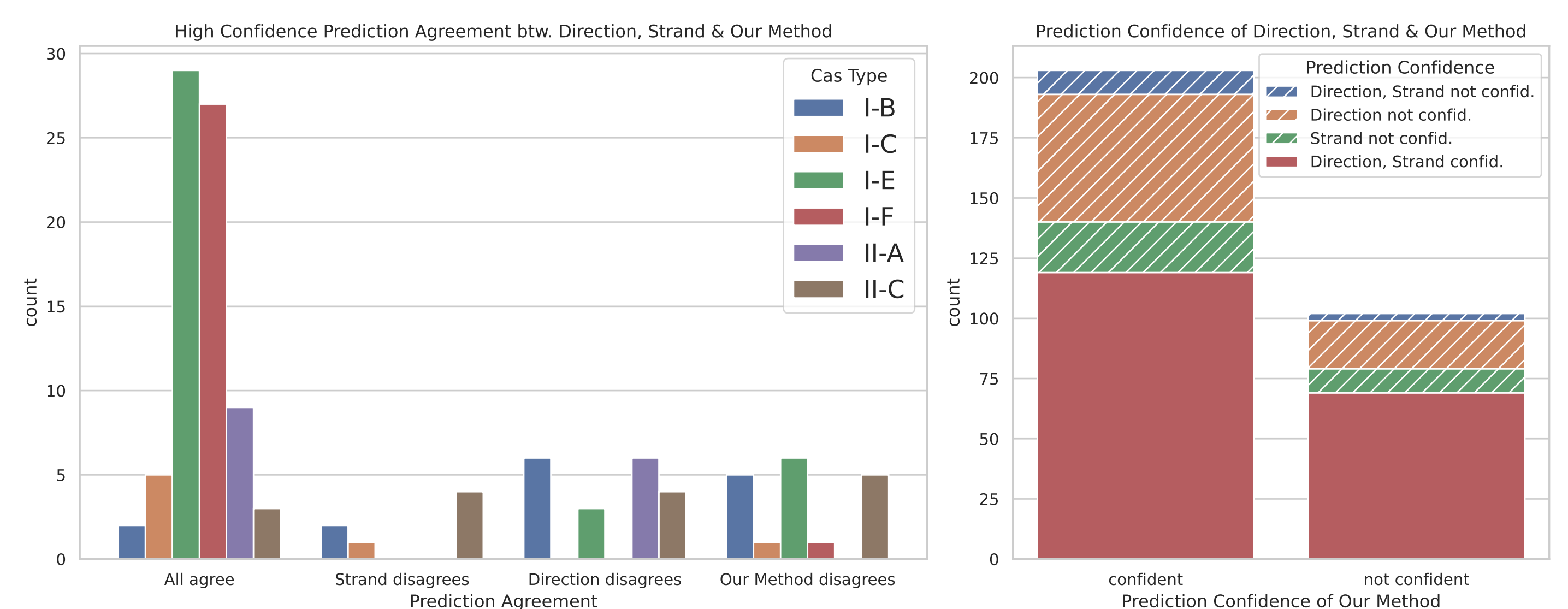
How can we predict Orientation from multiple CRISPR arrays? Behold the ...

CRISPR-evo-inator

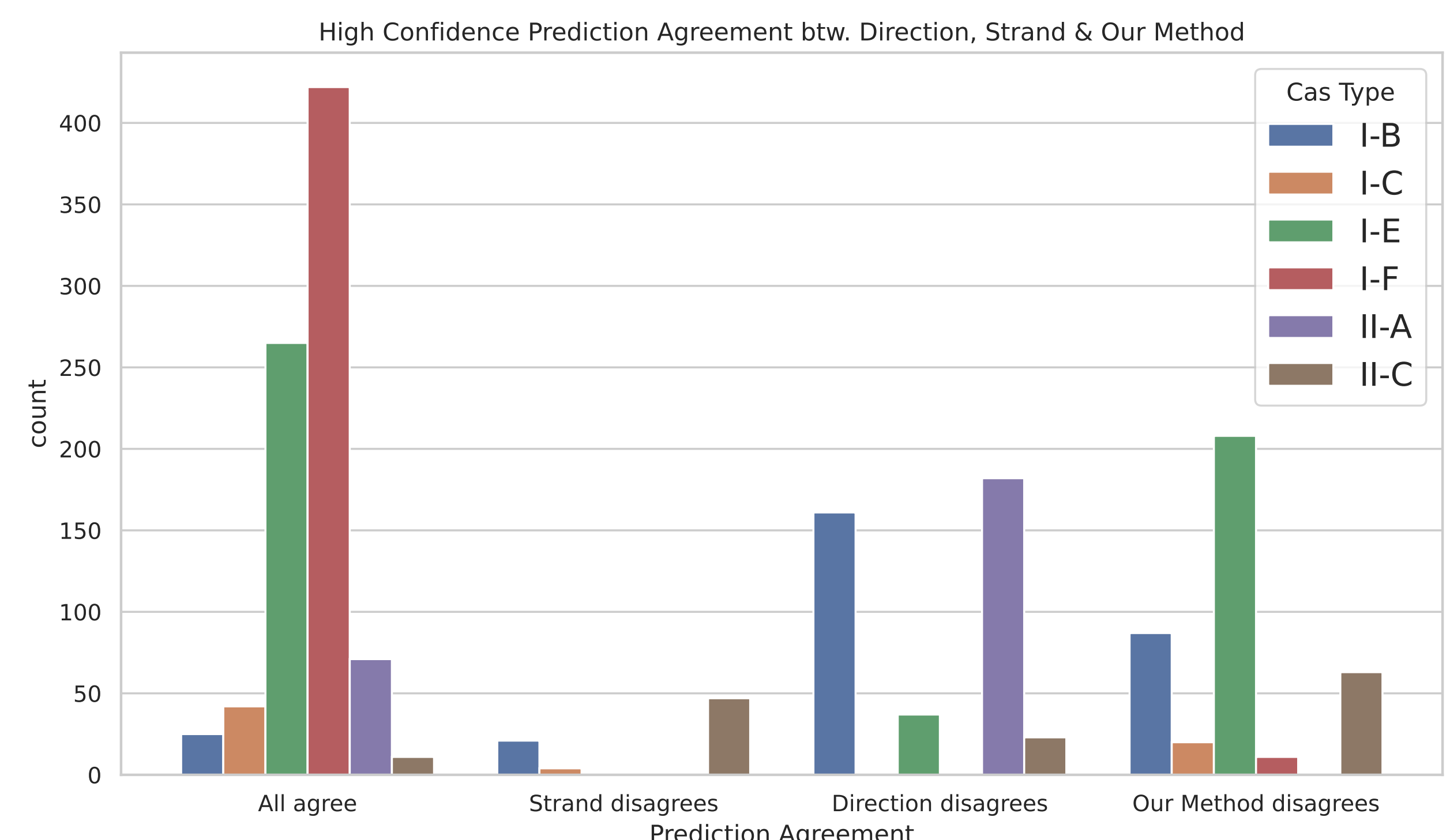
CRISPR evolution based inference of array transcription orientation



Orientation Prediction Results (groupwise)



Orientation Prediction Results (per array)



References

- [1] Omer S Alkhnbashi et al. "CRISPRstrand: predicting repeat orientations to determine the crRNA-encoding strand at CRISPR loci". In: *Bioinformatics* 30.17 (2014).
- [2] Ambarish Biswas, Peter C. Fineran, and Chris M. Brown. "Accurate computational prediction of the transcribed strand of CRISPR non-coding RNAs". In: *Bioinformatics* 30.13 (2014), pp. 1805–1813.
- [3] David Couvin et al. "CRISPRCasFinder, an update of CRISPRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins". In: *Nucleic acids research* 46.W1 (2018).
- [4] Christine Pourcel et al. "CRISPRCasdb a successor of CRISPRdb containing CRISPR arrays and cas genes from complete genome sequences, and tools to download and query lists of repeats and spacers". In: *Nucleic acids research* 48.D1 (2020), pp. D535–D544.