

Integrative OMICS approach to characterize Sf21 and Hi5 cell lines

Jonathan Landry¹, Bence Galik², Attila Gyenesei², Vladimir Benes¹ and GeneCore team¹

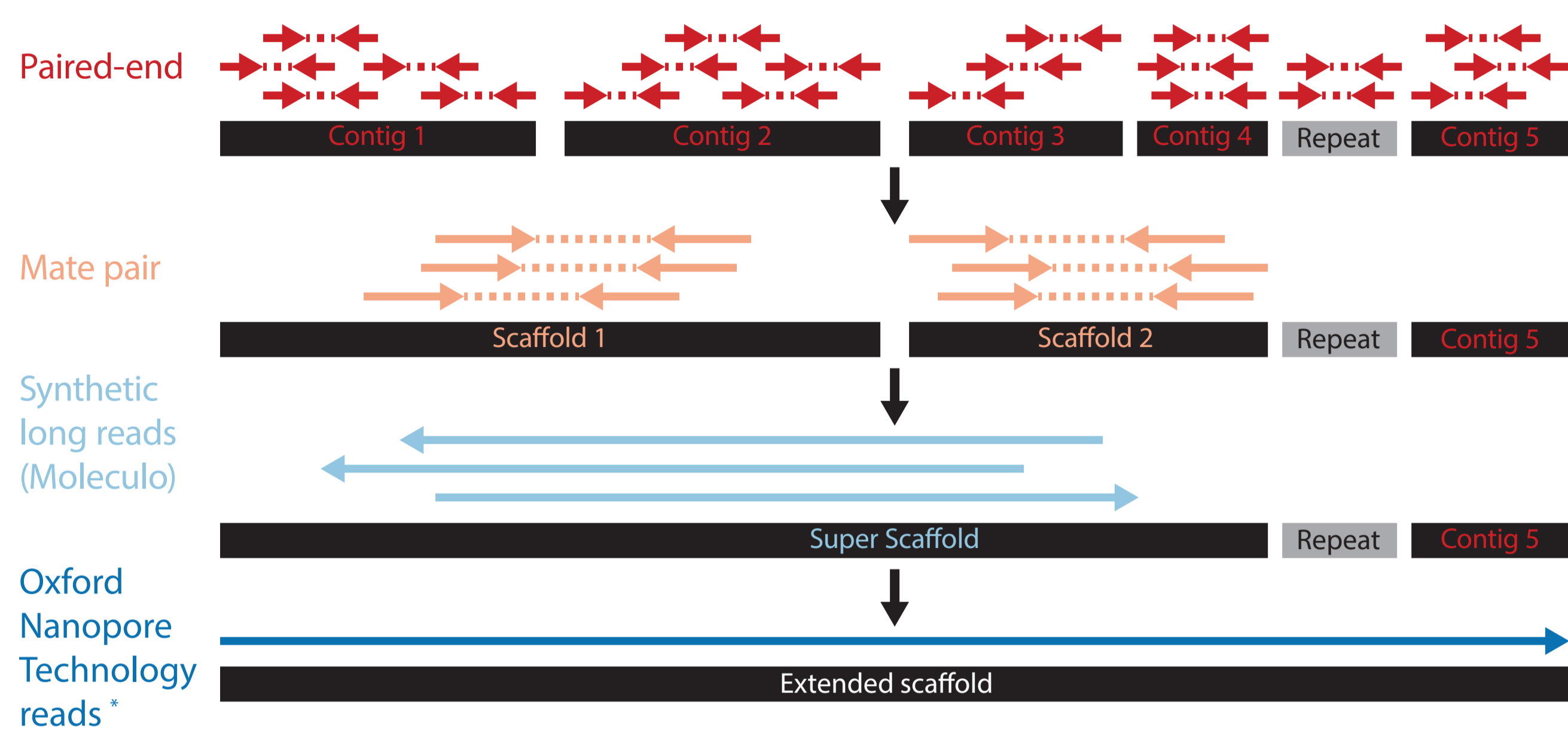
1. European Molecular Biology Laboratory (EMBL), Genomics Core Facility, Heidelberg (DE)
 2. Vienna BioCenter Core Facilities (VBCF), Bioinformatics & Scientific Computing, Vienna (A)
 Correspondence: jonathan.landry@embl.de



Motivation:

Insect derived cell-lines, from *Spodoptera frugiperda* (Sf21) and from *Trichoplusia ni* (Hi5), are widely used systems for recombinant protein expression. Genomic sequences and annotations are still incomplete for Sf21 or absent for Hi5.

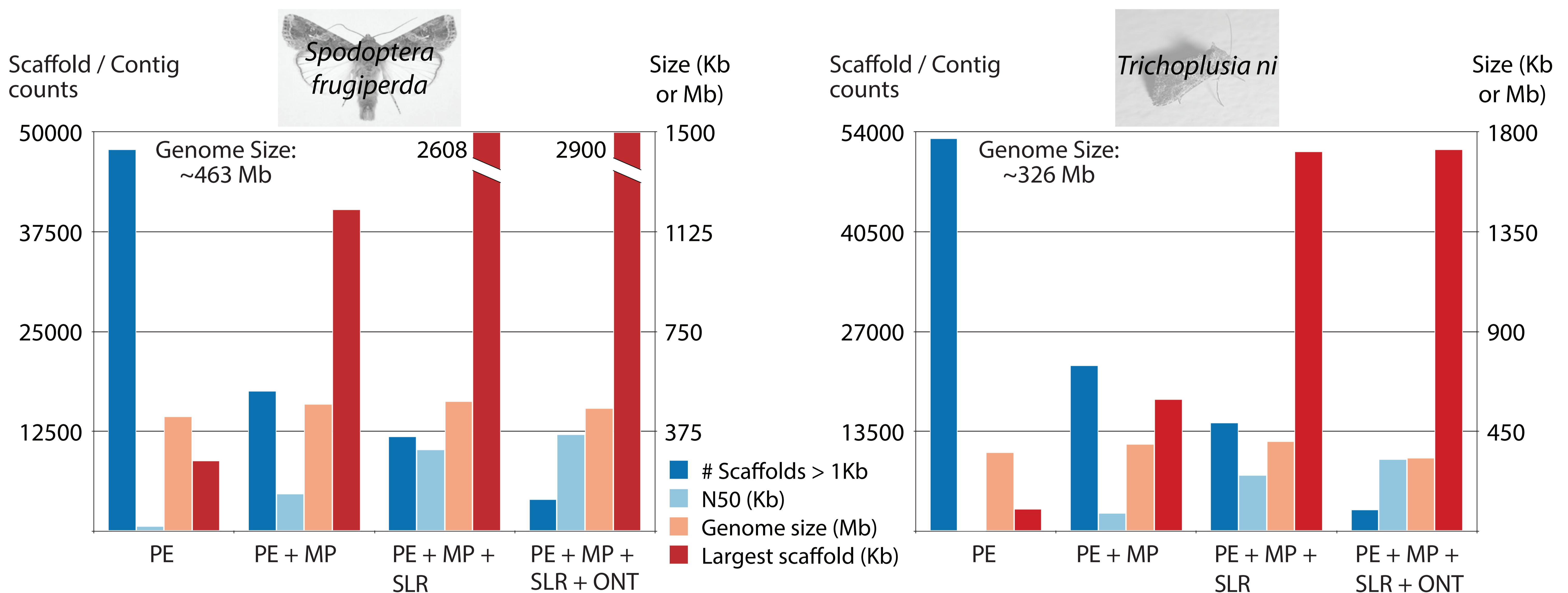
Multiple library type approach



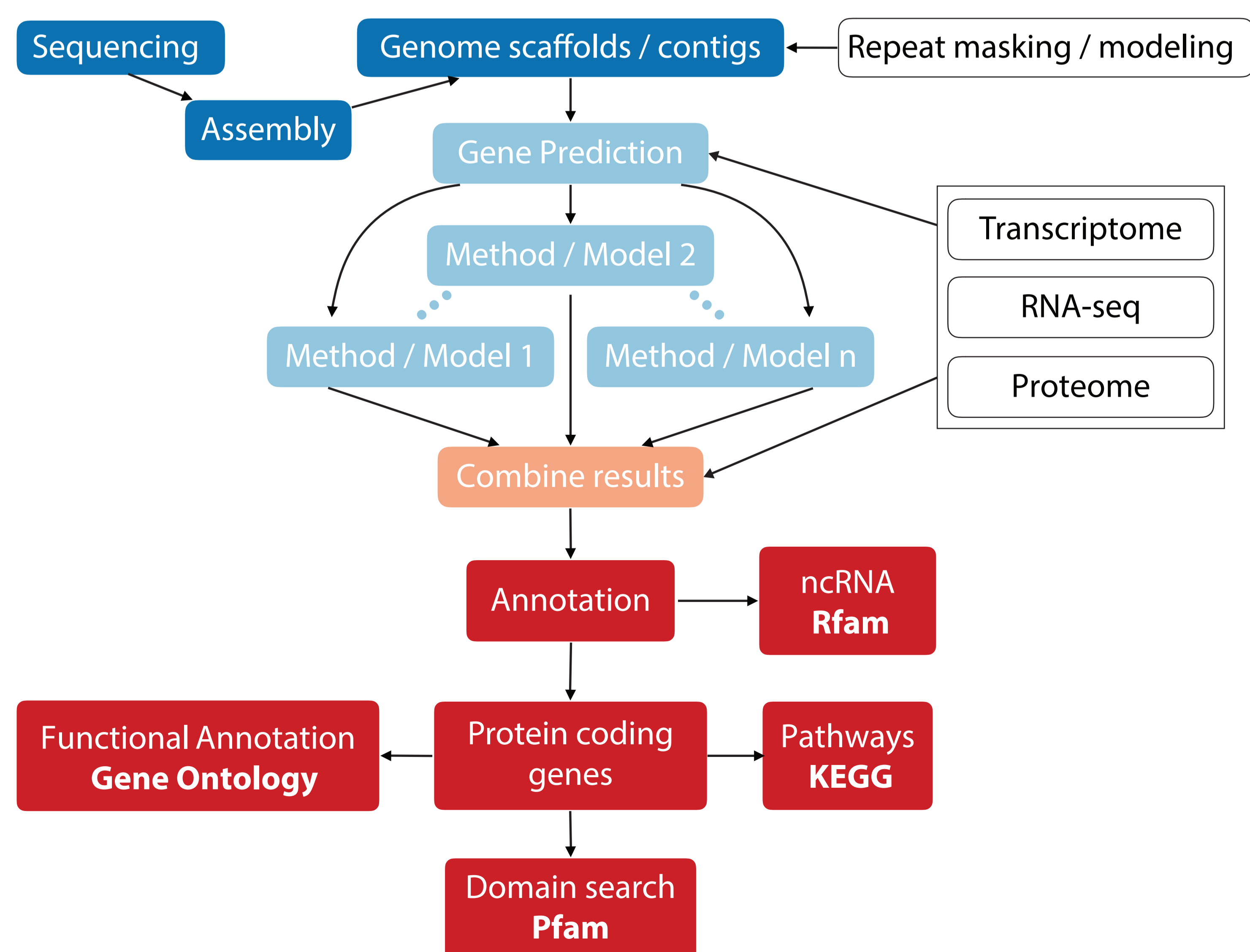
Datasets (Sf21)

Library	Input	Mode	Fragment size (bp)	Coverage seq. / spanning
1	DNA	Paired-End (PE)	~290	94.9 / 227.3
2	DNA	Paired-End (PE)	~590	1.4 / 12.7
3	DNA	Mate Pair (MP)	~4500	4.9 / 120.9
4	DNA	Mate Pair (MP)	~4500	13.9 / 320.1
5	DNA	Synthetic long reads (SLR)	~4900 (max 19 Kbp)	1.9 / NA
6	DNA	Oxford Nanopore (ONT)	~8000 (max 34 kbp)	0.2 / NA
7	RNA	Paired-End (PE)	~280	10.1 / 37.8

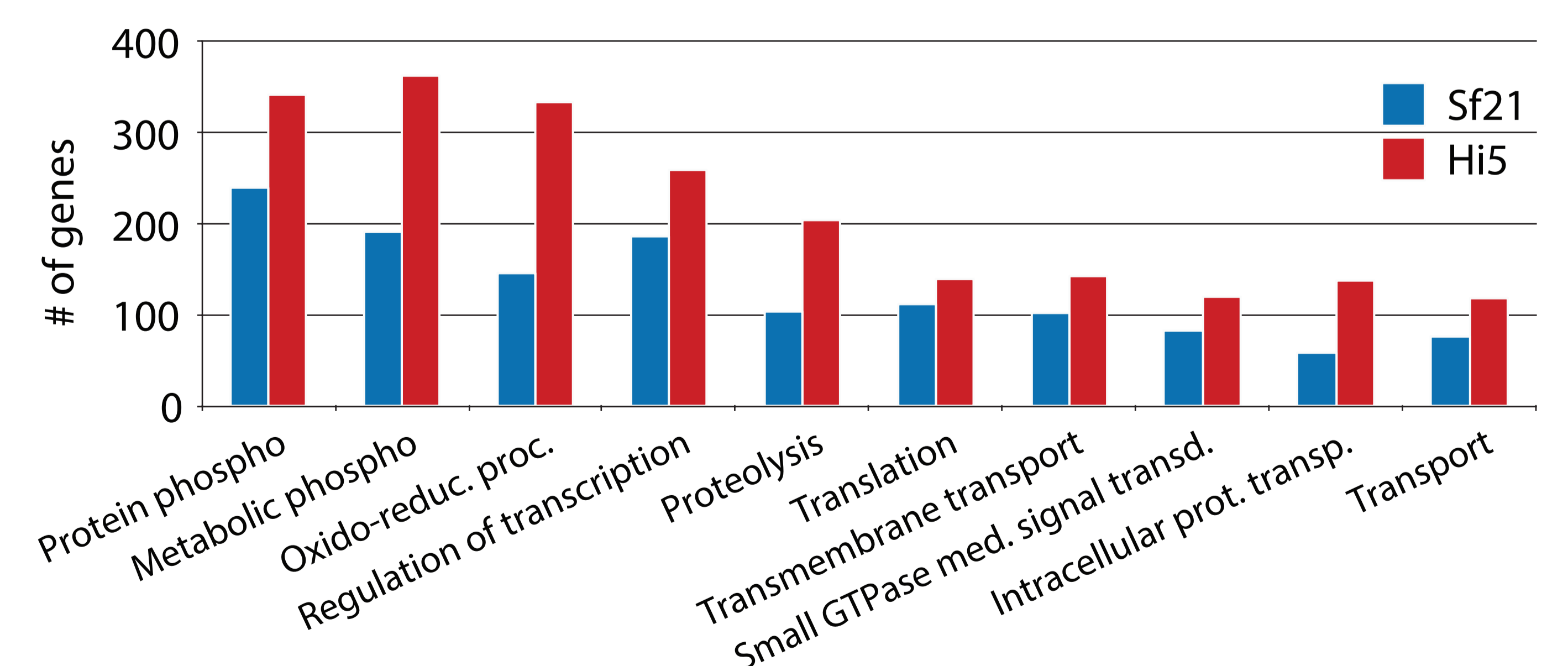
Genome assemblies



Annotation workflow



Genome function annotation



Conclusion:

This integrative approach results in assemblies at an unprecedented resolution. With these information available, Sf21 and Hi5 cells-lines will become an even better tools for protein expression and could be used in a wide range of applications, from promoter identifications to genome engineering and editing. We could in theory apply to any genomes and results to valuable resources.

* Cao, M.D., Nguyen, H.S., et al. Scaffolding and Completing Genome Assemblies in Real-time with Nanopore Sequencing. doi:10.1101/054783.