

Improvement of the assembly of heterozygous genomes of non-model organisms, a case study of the genomes of two *Spodoptera frugiperda* host strains

Anaïs GOUIN¹, Anthony BRETAUDEAU², Karine Labadie³, Jean-Marc Aury³, Emmanuelle d'Alençon⁴, Claire LEMAITRE¹ and Fabrice LEGEAIS^{1,2}

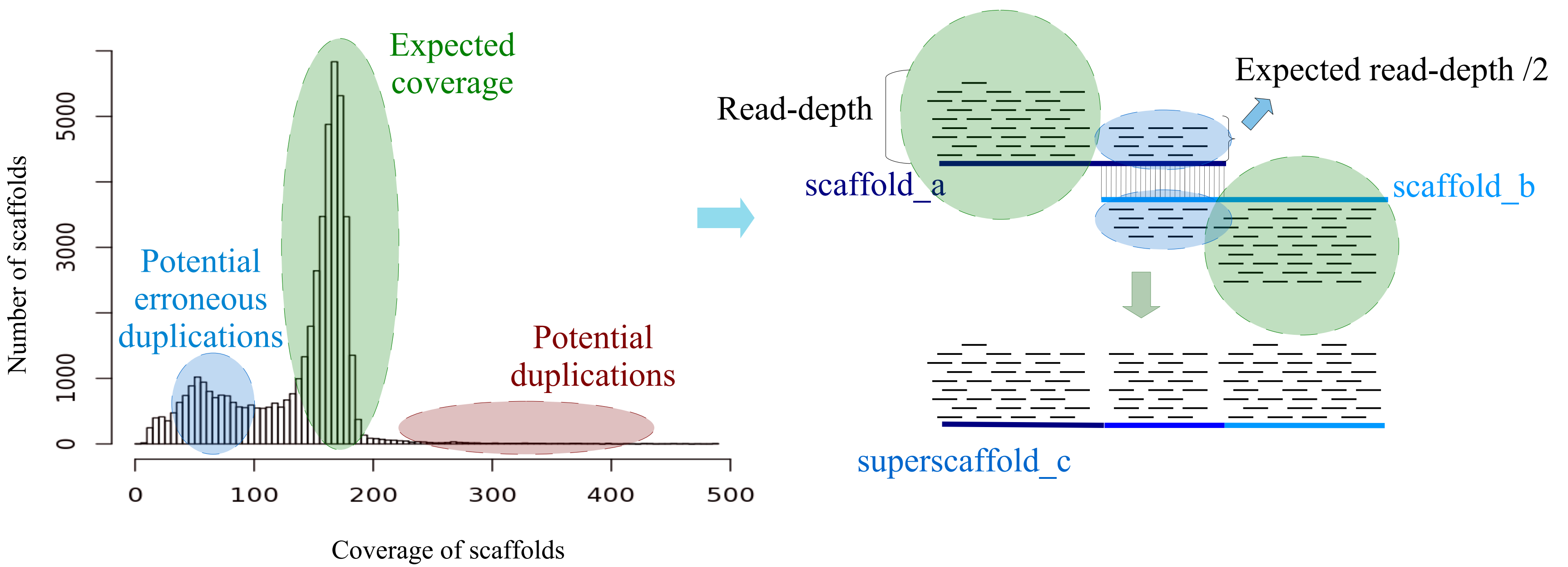


¹INRIA/IRISA/GenScale, Rennes, France
²INRA, Institut de Génétique, Environnement et Protection des Plantes (IGEPP), Rennes, France
³CEA Genoscope, Evry, France
⁴INRA, Diversité, Génomes et Interaction Micro-organismes (DGIMI), Montpellier France

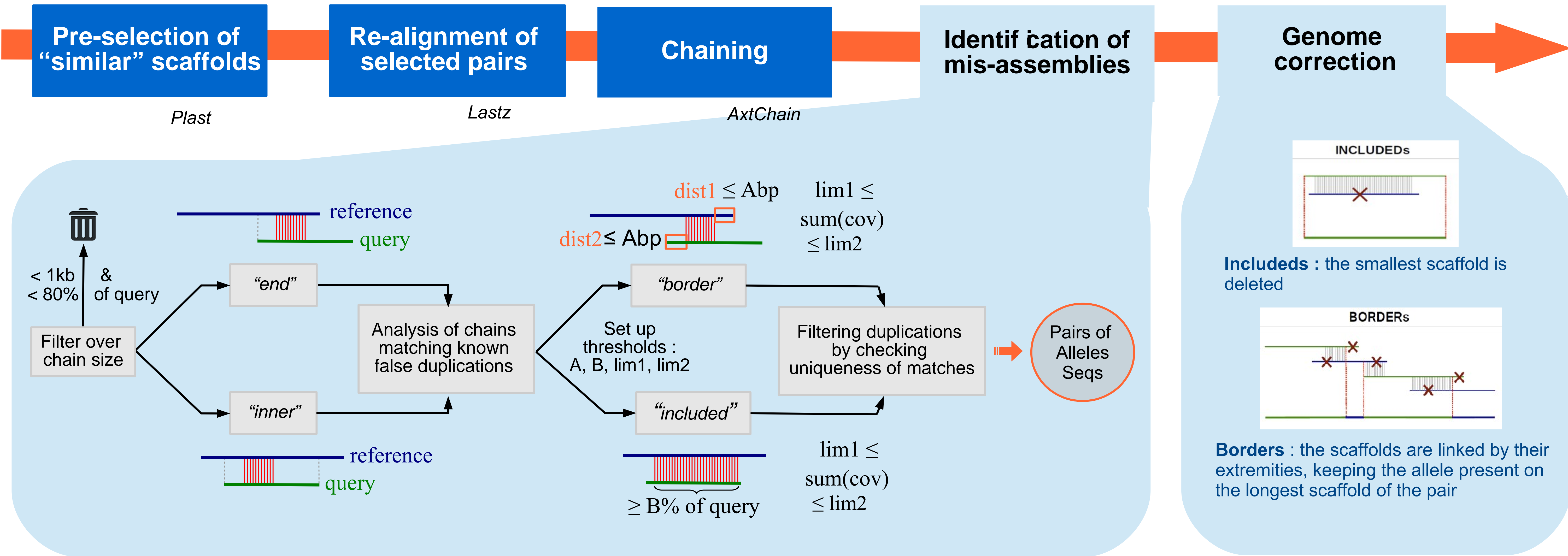


Motivation : Some heterozygous regions have a significant divergence between the two haplotypes and the assembly process can lead to the construction of two different contigs, instead of one consensus sequence.

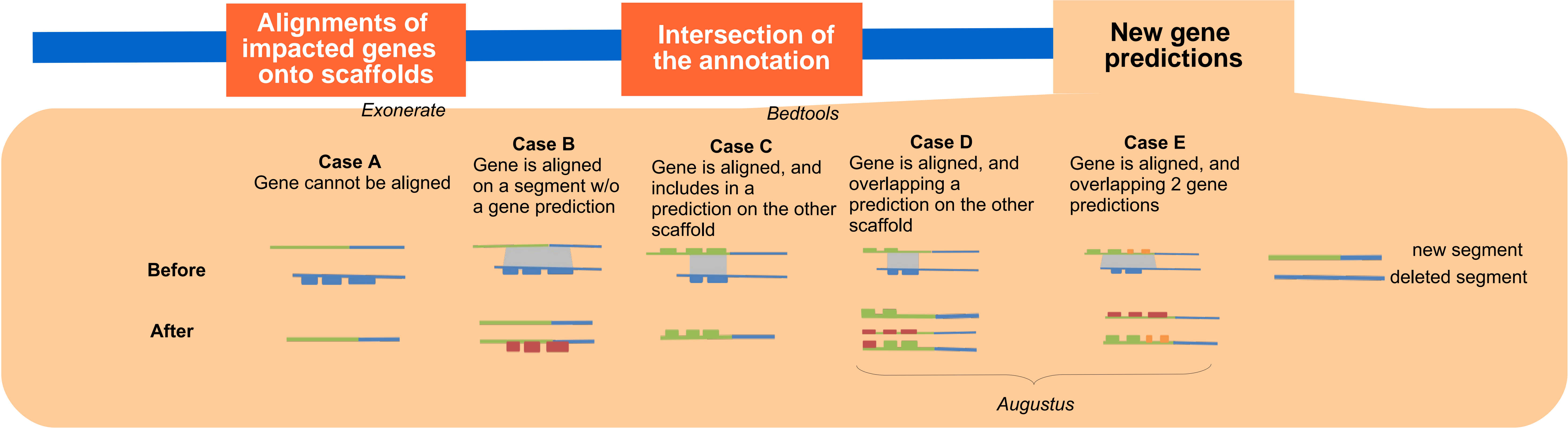
Objective : Set up a strategy to detect and correct false duplications in already-built assemblies.



RE-ASSEMBLY



RE-ANNOTATION



APPLICATION : *Spodoptera frugiperda* genomes

Assembly stats

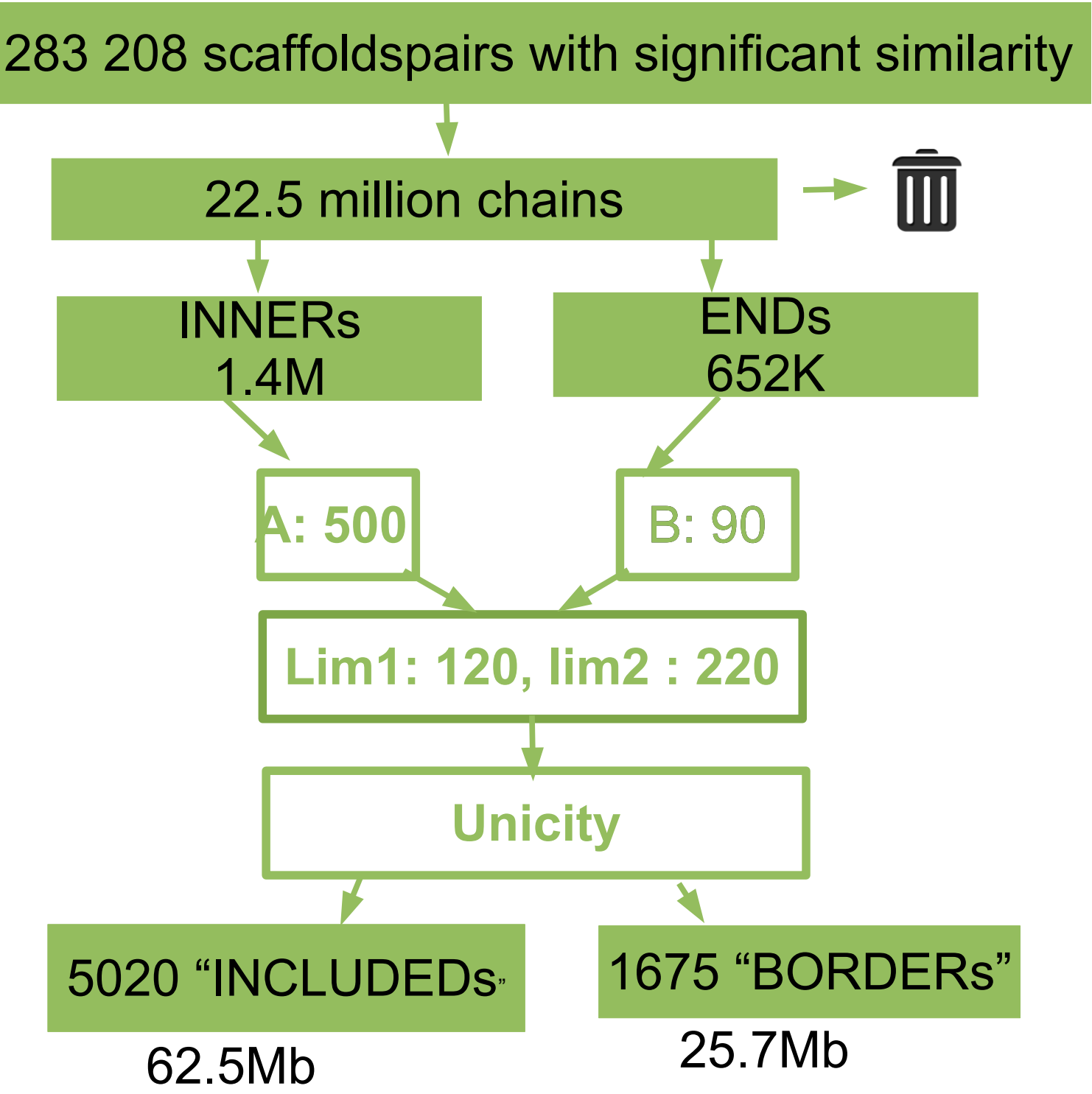
	Initial assembly <i>Allpaths</i>	<i>Platanus</i> assembly	Corrected assembly
Total size (Mb)	526.0	470.1	437.9
Nb. scaffolds	48 272	41 633	41 577
N50 (bp)	39 593	75 578	52 781
Total Gap length	13.6 Mb	56.1 Mb	11.4 Mb
BUSCO* stats	No hit	14	15
	Single hit	1497	1885
	Multi hit	782	393

Annotation stats

Previous release : 25,041 genes
3 746 genes to remap

	# genes	% success
Case A	34	0
Case B	747	45.4%
Case C	643	100.0%
Case D/E	2322	86.3%

New release : 21,578 genes



Comparison of 2 *Spodoptera* strains: the genome of the *Spodoptera frugiperda* "rice variant" has been sequenced and assembled with *Platanus*, and annotated with the help of *Maker*. The comparison of the two draft sequences leads to the identification of about 16,000 pairs of orthologous genes, and thousands of segmental rearrangements (deletions, duplications or inversions).

