

Un exemple d'application du logiciel pepline:

ANNOTATION DU GENOME DE L' ALGUE VERTE CHLAMYDOMONAS REINHARDTII

Marianne Tardif #, Annie Adrait #, Norbert Rolland *, Jérôme Garin #

Laboratoire de Chimie des Protéines , ERM Inserm 0201 , DRDC/CP - CEA/Grenoble

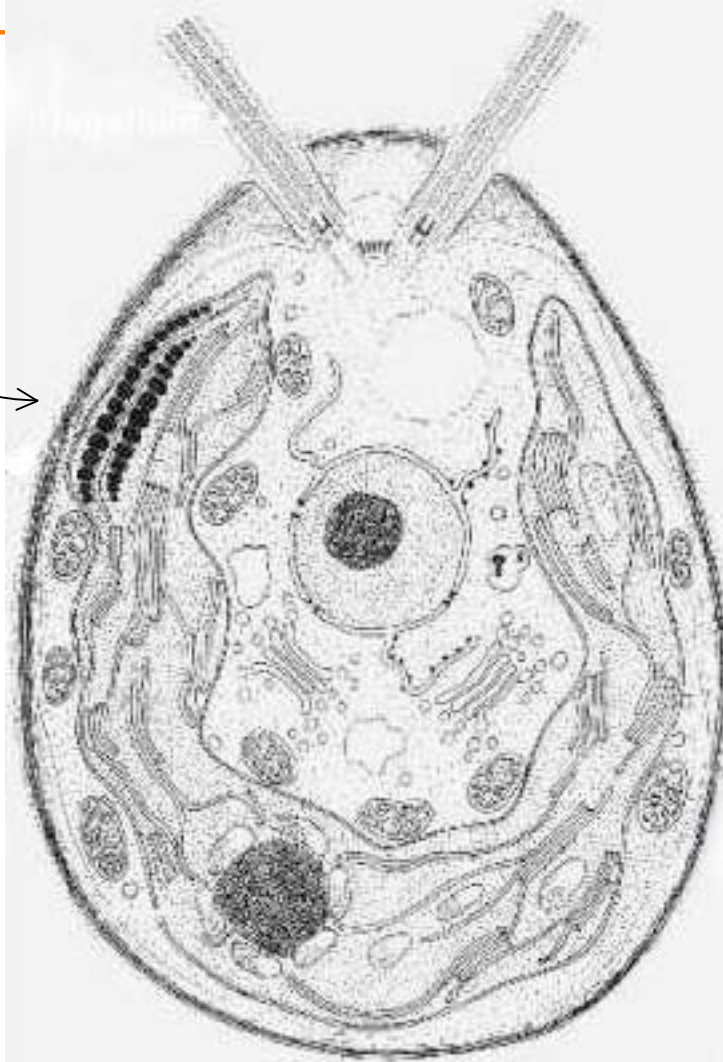
* Laboratoire de Physiologie Cellulaire Végétale , UMR 5019, DRDC/ PCV –CEA/Grenoble
17 rue des martyrs F-38054 Grenoble

Chlamydomonas reinhardtii: aspects bio

Rhône-Alpes

Genopole

oeil!
oeil!



- photosynthèse
- biogénèse (chloroplaste)
- mobilité
- tropisme (phototactisme)
- autres métabolismes, d'intérêt agronomique (synthèse amidon)
- cycles circadiens
- réponses au stress

Protéines du chloroplaste = 50% des protéines totales

Chlamydomonas reinhardtii: aspects bioinfo

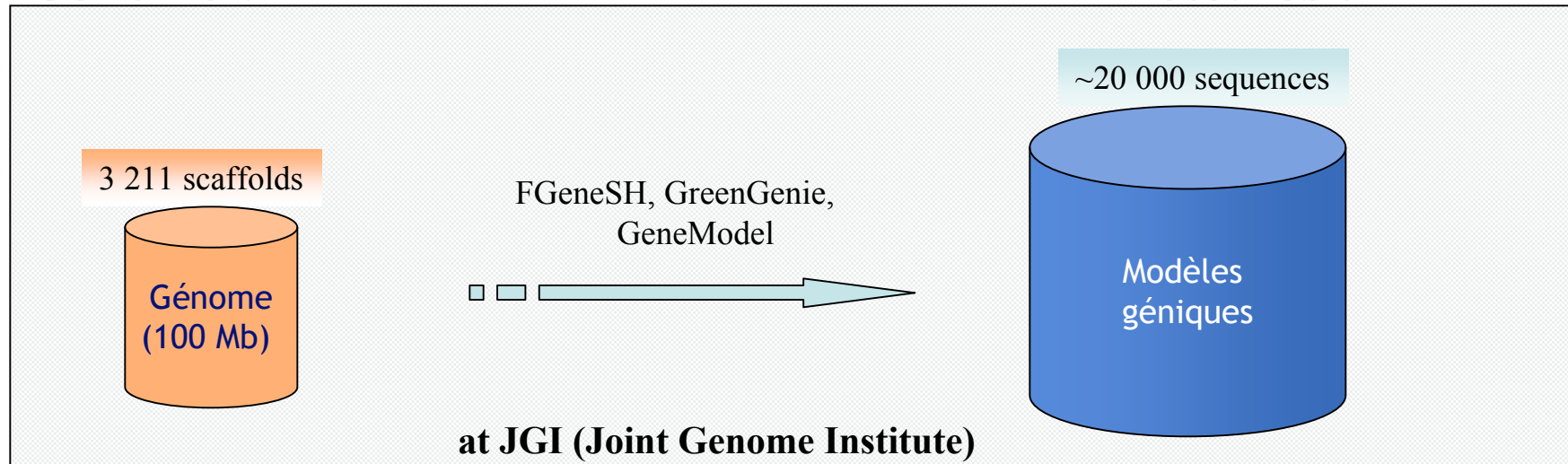
<http://genome.jgi-psf.org/chlre2/chlre2.home.html>

Rhône-Alpes

Genopole

Genome :

Protéines :



- taille : 100 Mb
- 17 chromosomes
version1: 4 000 scaffolds
version2: 3 211 scaffolds
- Biais GC (62%)
- Annotation au JGI

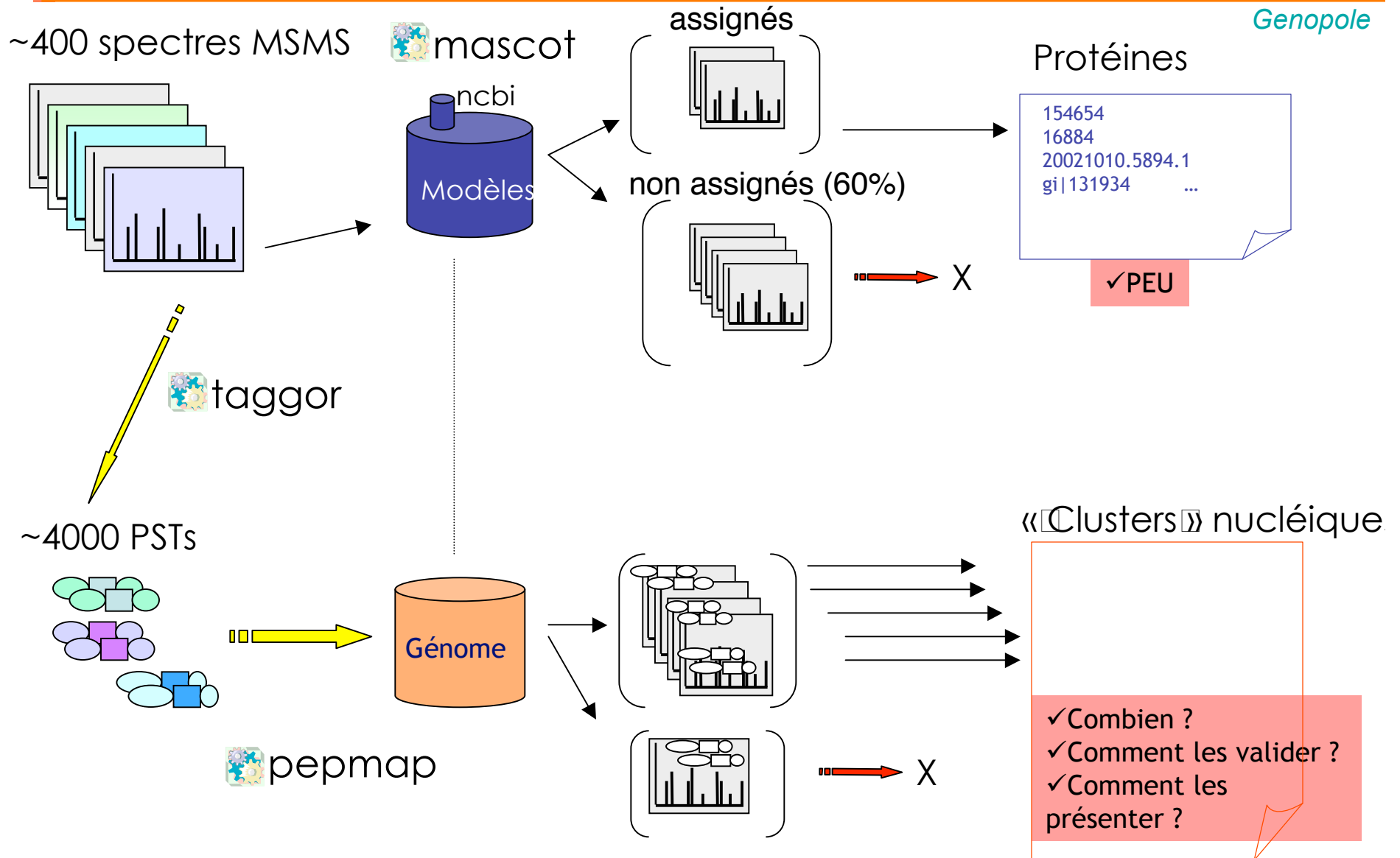
~3 000 sequences (C. Hauser)



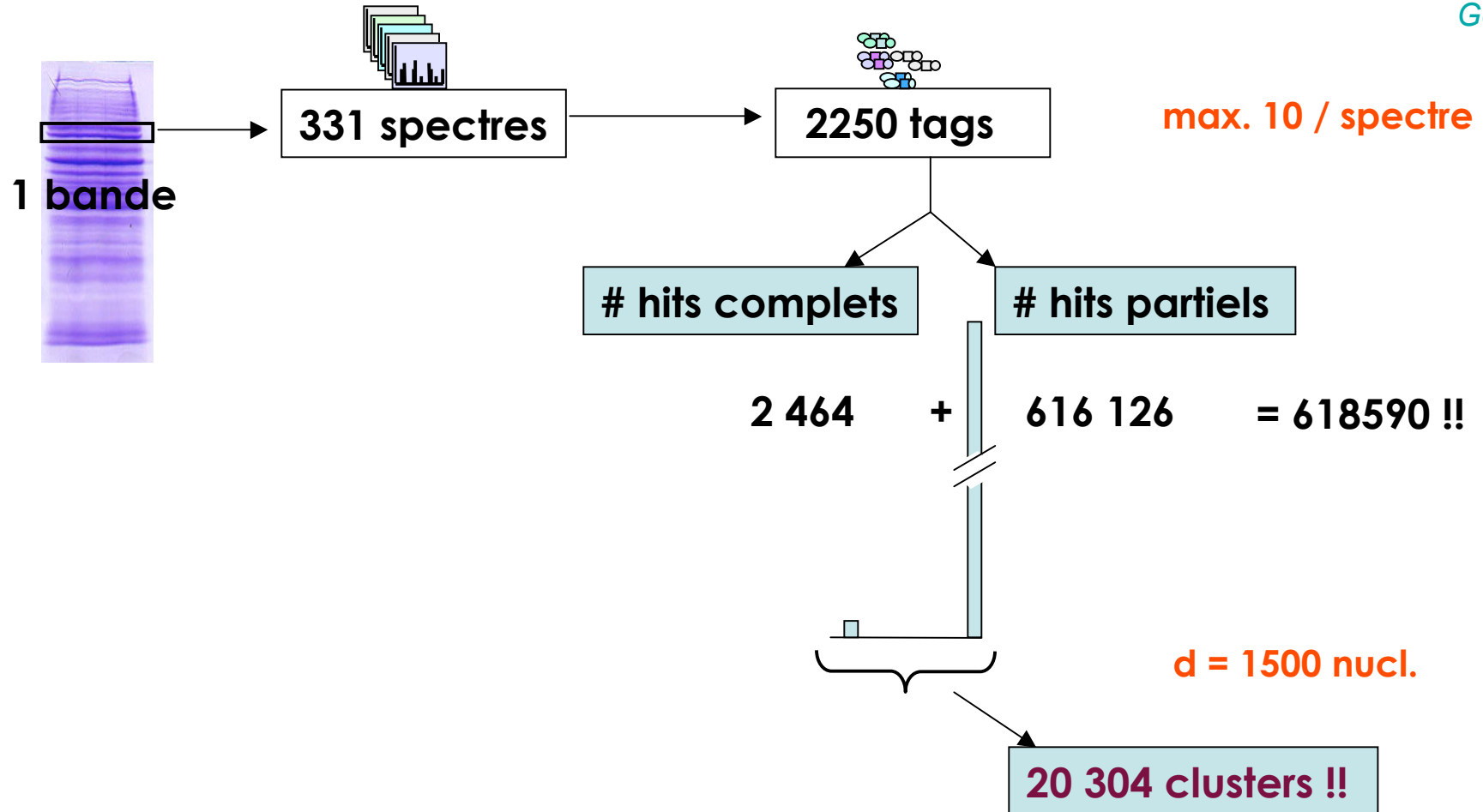
~700 sequences « Chlamy » (@NCBI)



Pepline versus Mascot

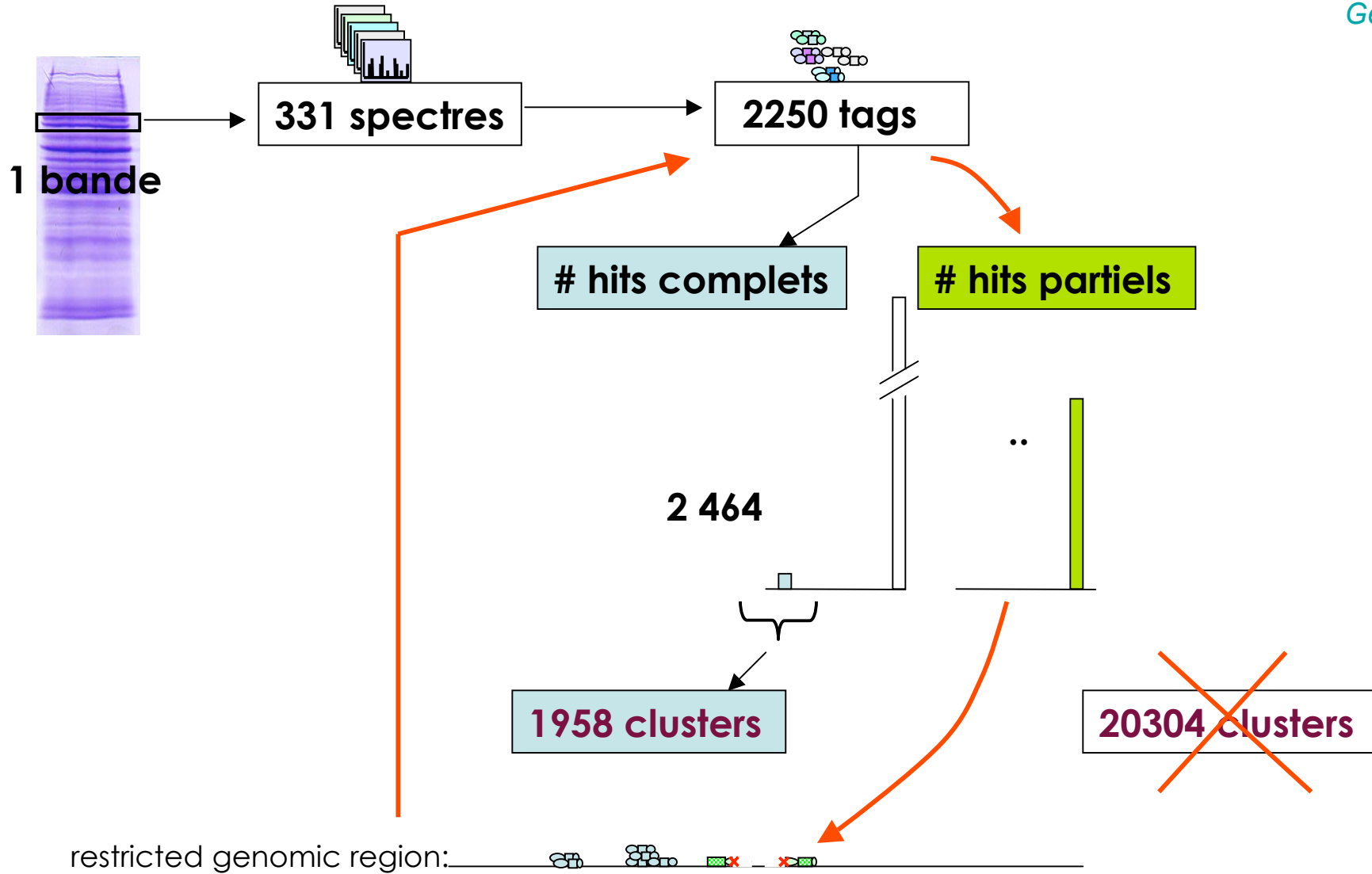


nombre de clusters générés -1

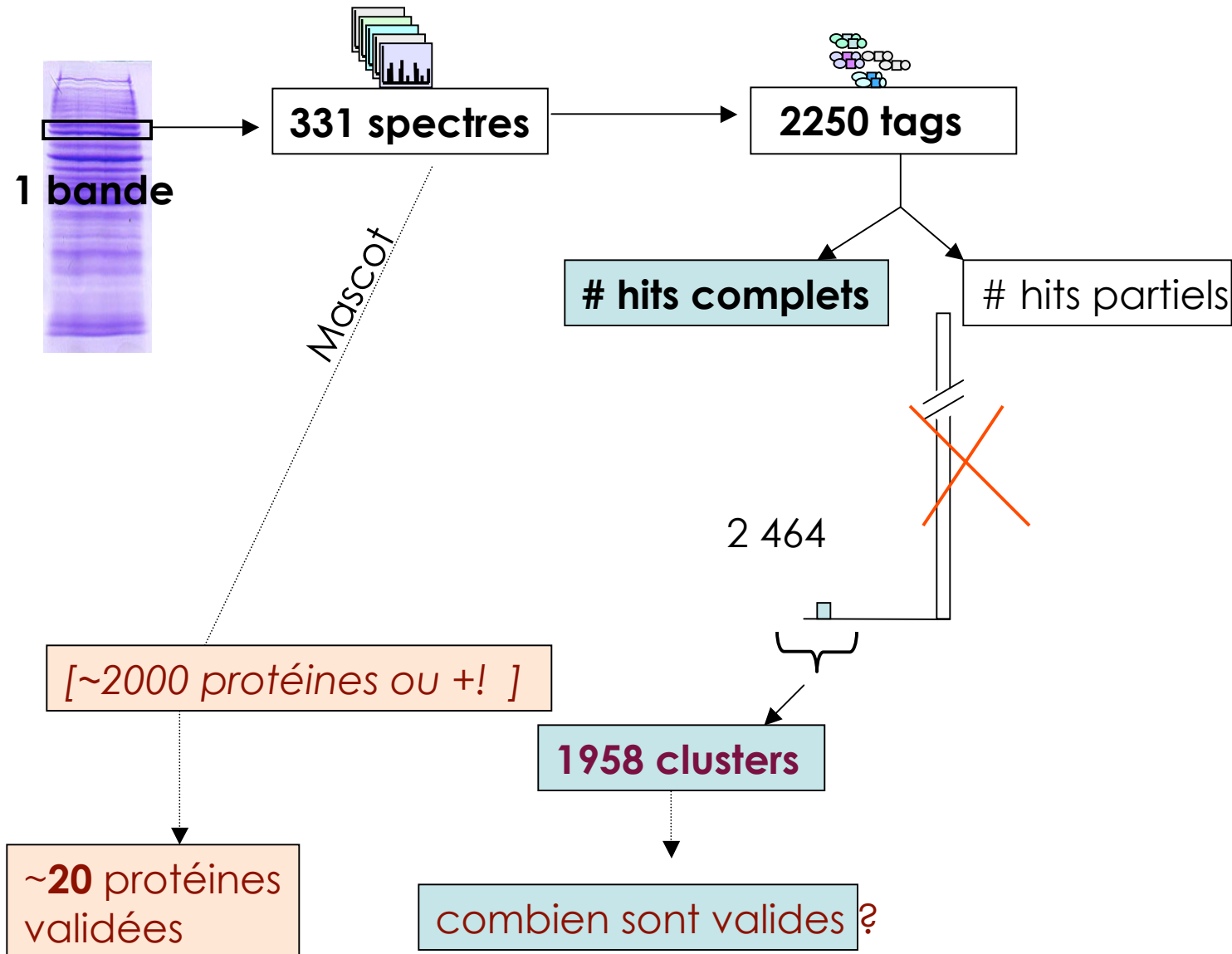


clusters « géants »
ex.30 kb...

nombre de clusters générés -2

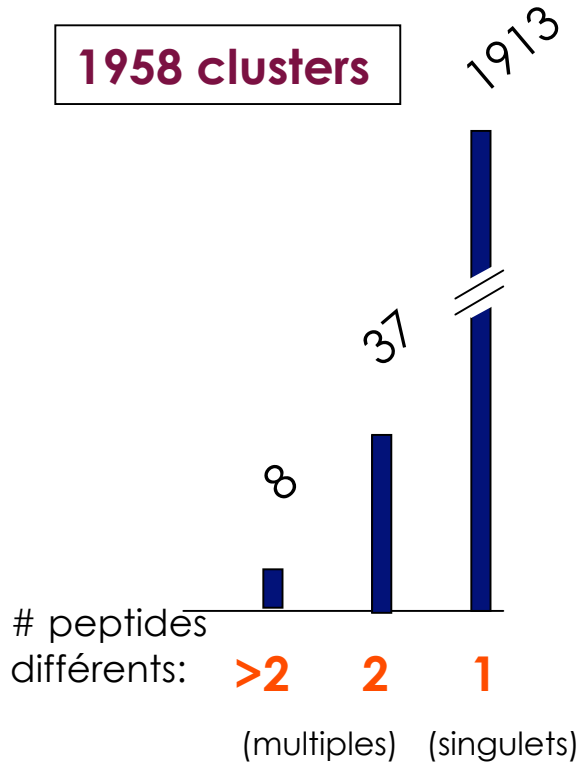


nombre de clusters générés -3



validation des clusters - 1

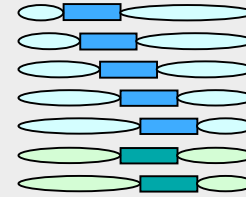
1958 clusters



« *singulet* »



7 hits

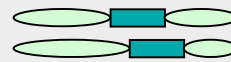


(EAEAISVDVTSK)

multiple



3 hits



(EAEAISVDVTSK)



(AVLIGDESDFER)

Valides ?

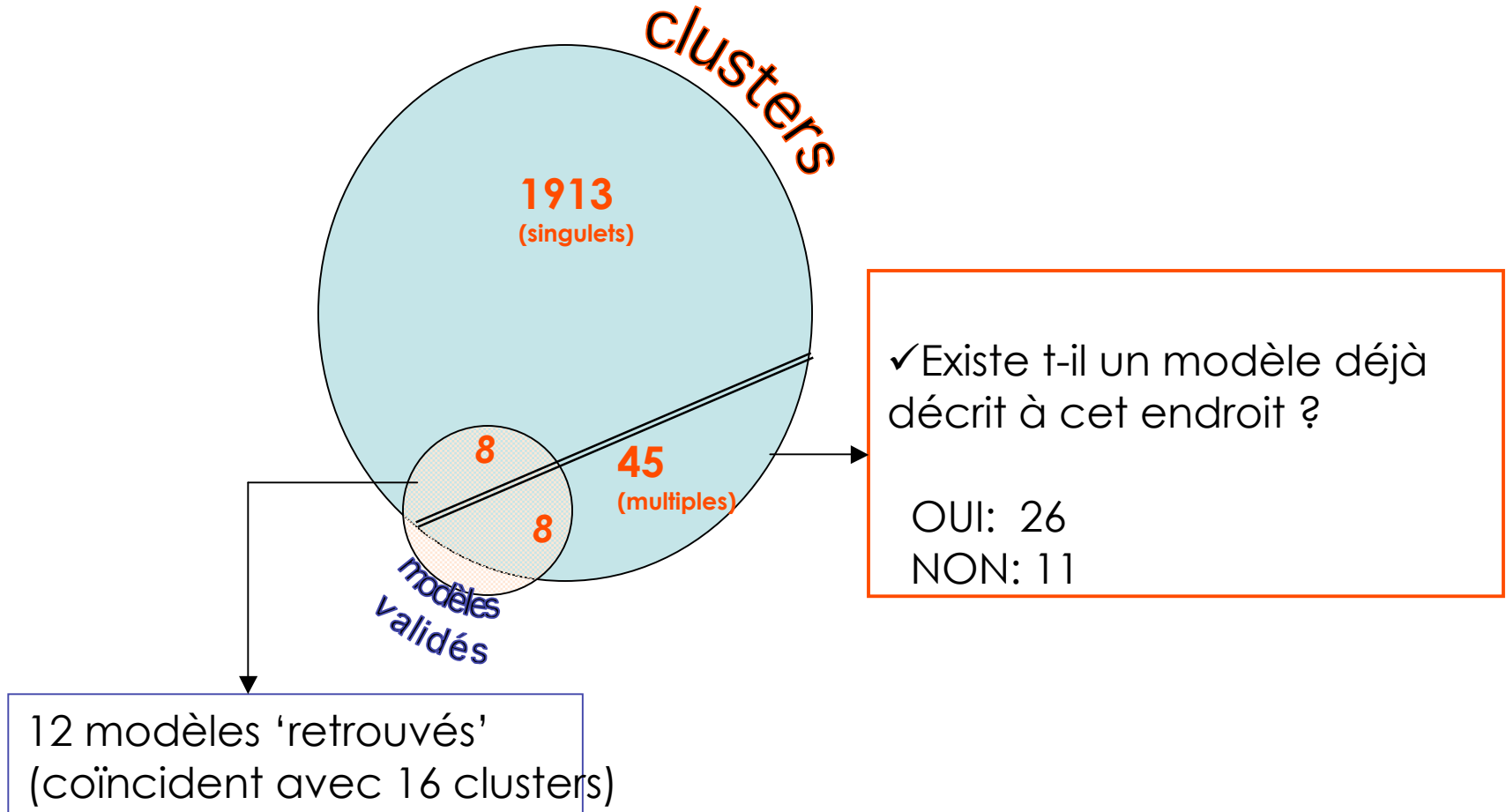
??

Bruit ?

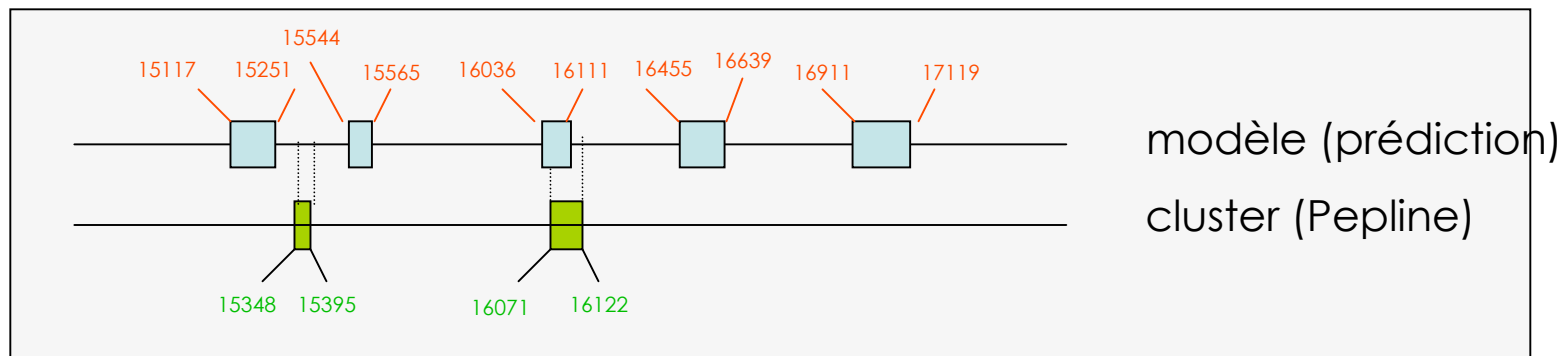
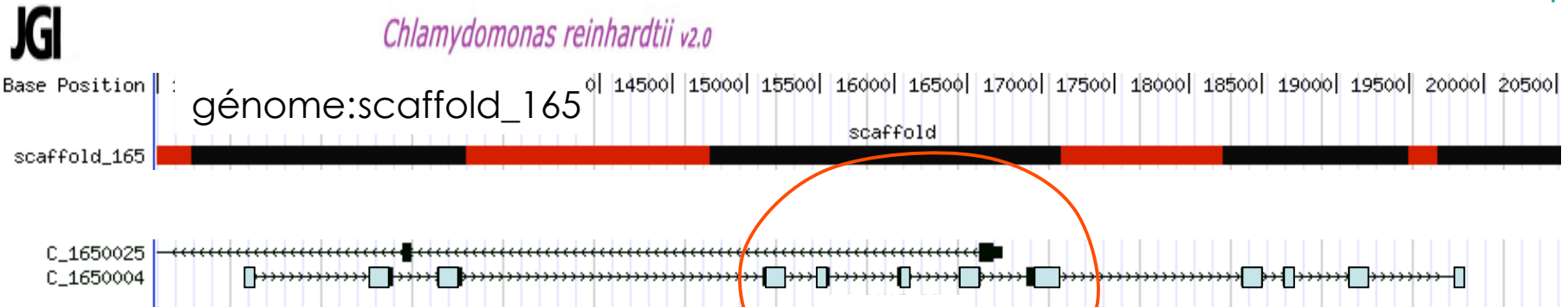
→ fonction de score ?

validation des clusters -2

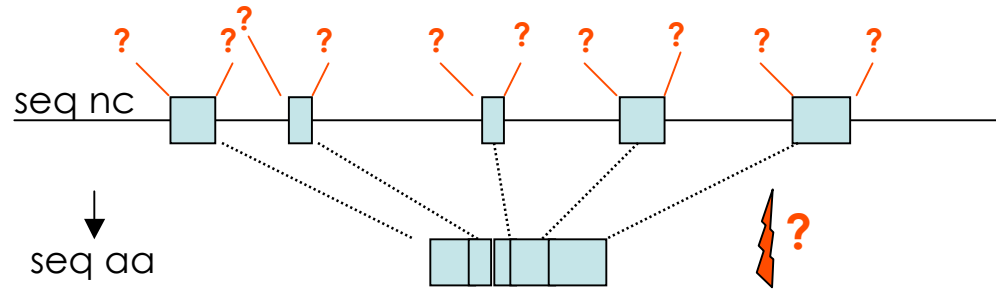
Validation indirecte par comparaison avec les résultats Mascot



Exemple



Formalisme-1



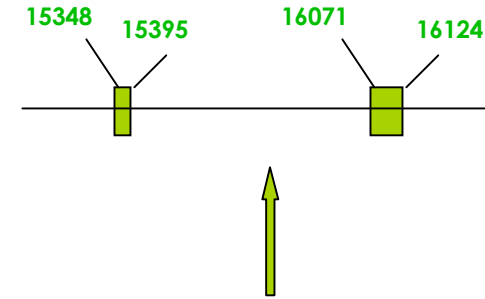
modèle (id 157017)

>157017

```
MATDPAGCAAACPAGPLASVLGAAPPGGYDPDQVAAVTFWS
FALAPWREVGYNRTMVKATELVALLSNESYLSKPVNASFDY
NTTYMREAVRTVYSTAQLPITAAAVPLADGETDIITRLKRN
REPIRGNYCPEGSSKPSIVCEGGYYCPNATSQIICPAGYYC
KPQSIYPVACPPVVACPEGTESPDGRPLAGLIFAFIVVGMY
LLYWVTELGMMVGERIIRHLSLVARIRKKNLANLGQIAGITE
TQTEEHKEMEKVQEAARARNEVGFQMATMKTSPWWELNENT
NIRFRDVRGRMPKKGKSGRLENLEYSARLKLPTVPRVYRHG
IIDDTLRMLGMYDKQDRLTGSVENKVISGGERKRVSIGVEI
VGKPPILFMDEPTSGLDAARSELCTLMSNLAAASKTNI IA
VIHQPRYSVRC*
```

>157018

```
MSNALRGLARRVGAQYLRNQRVRRGGGGEYPGGSFWSEGTQT
GKNGFLFGEVPIGQPRKTLWWEPYWYAGFGGMGVGVYLIY
HAKPLEALDIKYWAAPRAAKELETEMRMLDKLNERPDLKER
LVAVCKDLNMIIEDEAYDLVLMRNEYKVKLGMHTRVPEDLK
AIYEELEA*
```



Cluster Pepline (id 520)

```
scaffold_165@D1 777 15348 16124 102 ...
```

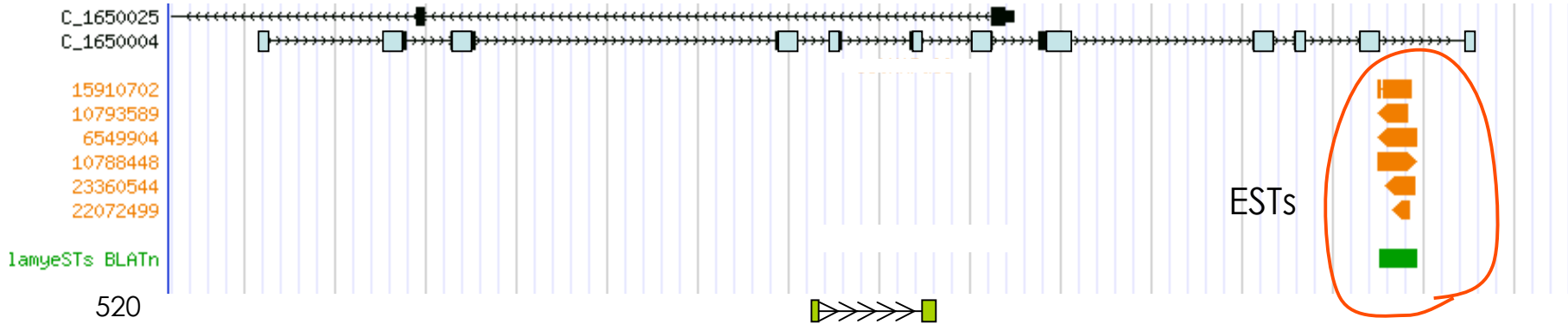
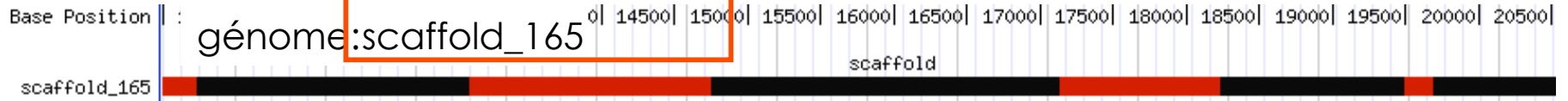
777	LONGUEUR
15348	nucl DEBUT
16124	nucl FIN
102	COUVERTURE (# nucl)
2	NBRE DE PEPTIDES DIFF
...	SCORES
...	
...	COORD DES HITS
etc	

Formalisme-2



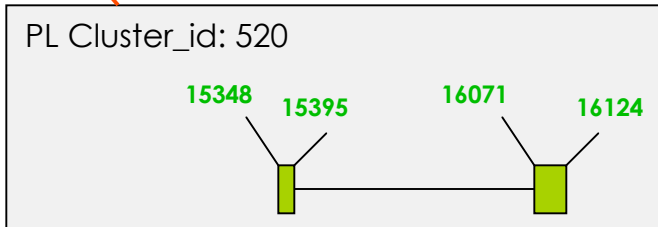
Chlamydomonas reinhardtii v2.0

persistence des résultats d'annotation ?



comparaison

nouvelle piste sur le « browser »





Chlamy and Co.

Rhône-Alpes
Genopole

http://www.biology.duke.edu/chlamy_genome/
<http://genome.jgi-psf.org/chlre2/chlre2.home.html>

LA MASSE, LA BIOINFO:

. CP /DRDC/ CEA Grenoble

Marianne Tardif

Annie Adrait

Jérôme Garin

LES ECHANTILLONS, LA BIO:

. PCV UMR 5019 /DRDC/ CEA Gre.

Norbert Rolland

LES COLLABORATEURS:

. UMR 8576 CNRS, Villeneuve d'Ascq

Steven Ball

. UPR-CNRS 1261 /IBPC de Paris/

Francis-André Wollman

Olivier Vallon

LES CONTACTS AMERICAINS:

Stanford Univ & Duke Univ:

Arthur R. Grossman Stanford University

Charles Hauser Duke University

Joint Genome Institute (JGI):

Dan Rokhsar Lawrence Berkeley Lab

Diego Martinez Los Alamos Nat. Lab

Et

Les personnes citées pour Pepline !