

BisonEx

Un pipeline bioinformatique de ré-interprétation
d'analyses constitutionnelles d'exome

Laboratoire de Génétique, CHU Minjoz, Besançon

Alexis Praga, 18/04/2024



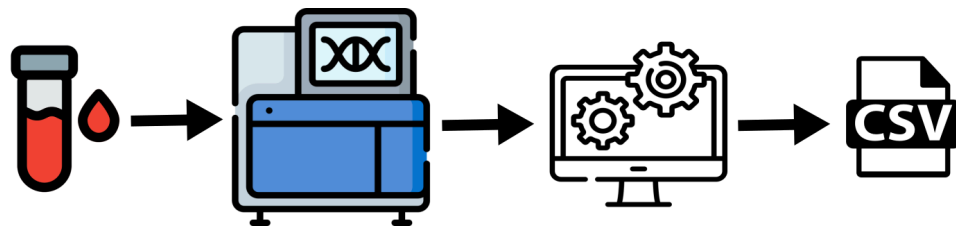
Contexte

- Consultations de maladies rares (Centre de Génétique Humaine)
- *Exome* souvent prescrit après un premier bilan
 - 1% de l'ADN
 - rendement diagnostic 30-50% (selon indication)
 - sous-traité à un laboratoire privé accrédité

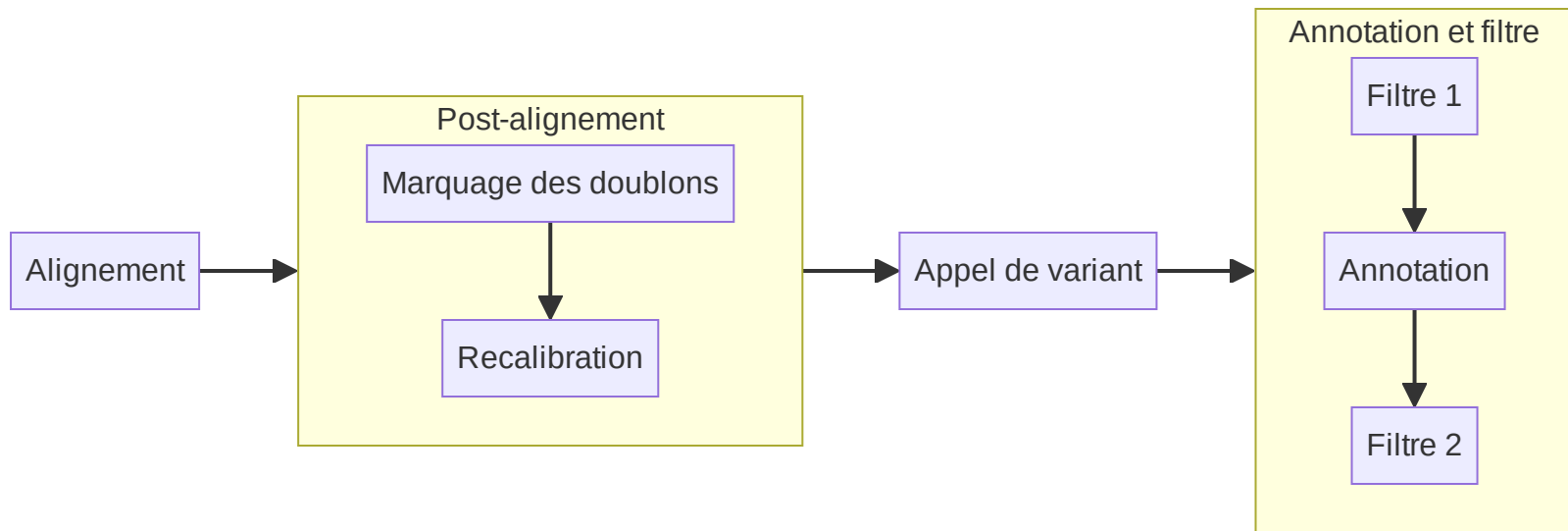
Contexte

Patients en errance diagnostique

- ré-interprétation à la demande
- données brutes **depuis 2022**
- **pipeline maison** (v0.1 par Dr. A. Overs)



Pipeline



Objectifs de cette thèse

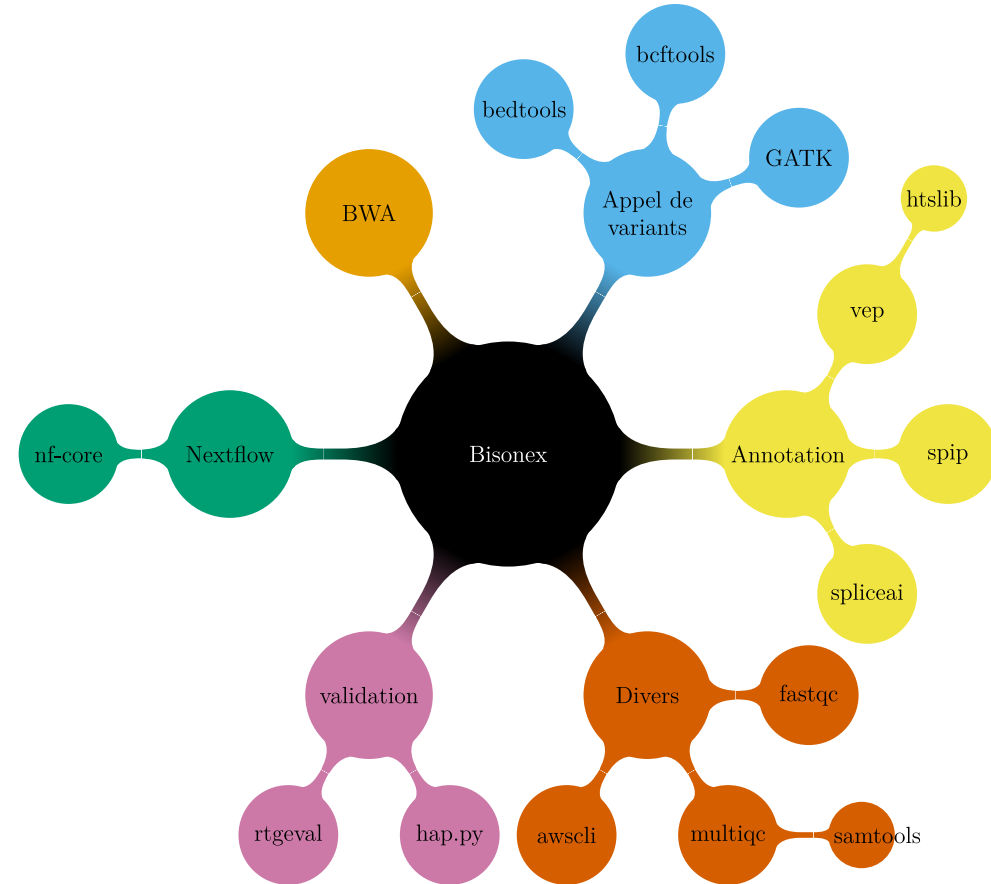
1. Pipeline reproductible, portable, performant
2. Validation (patients de référence, *in silico*)
3. Ré-analyse des exomes négatifs

1. Reproductibilité

Comment assurer au COFRAC des résultats reproductibles ?

Nix

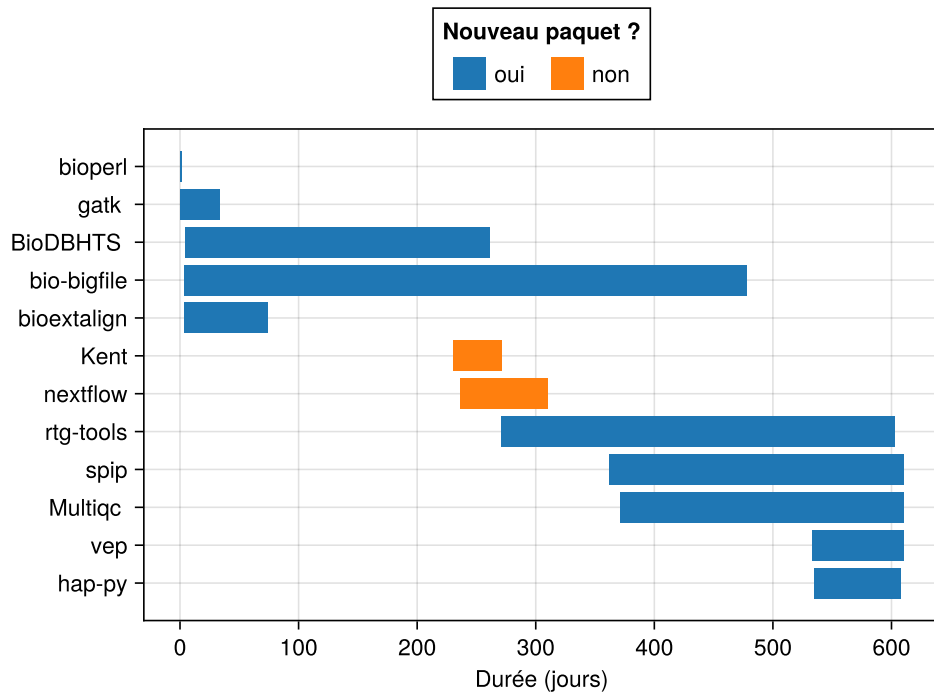
1. bloque la version de tous les logiciels
2. quelque soit l'ordinateur (Linux +/- OSX)
3. environnement logiciel de "production" identique à celui validation



1. Reproductibilité

Incorporation dans Nix :

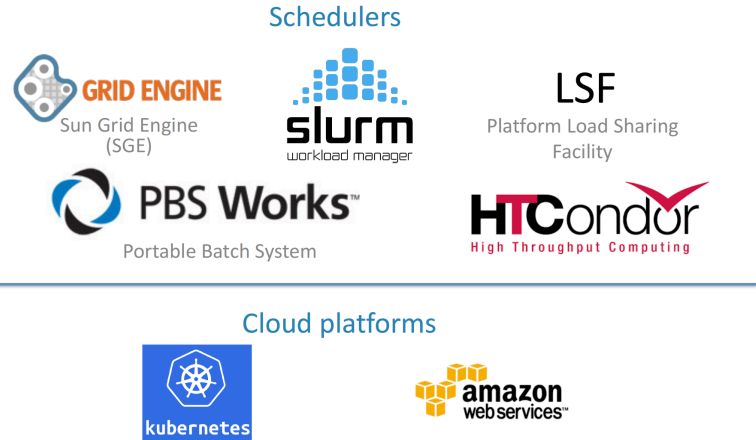
- 6 logiciels + 3 dépendences
- 3/6 déjà utilisables par la communauté (nixpkgs)



1. Portabilité

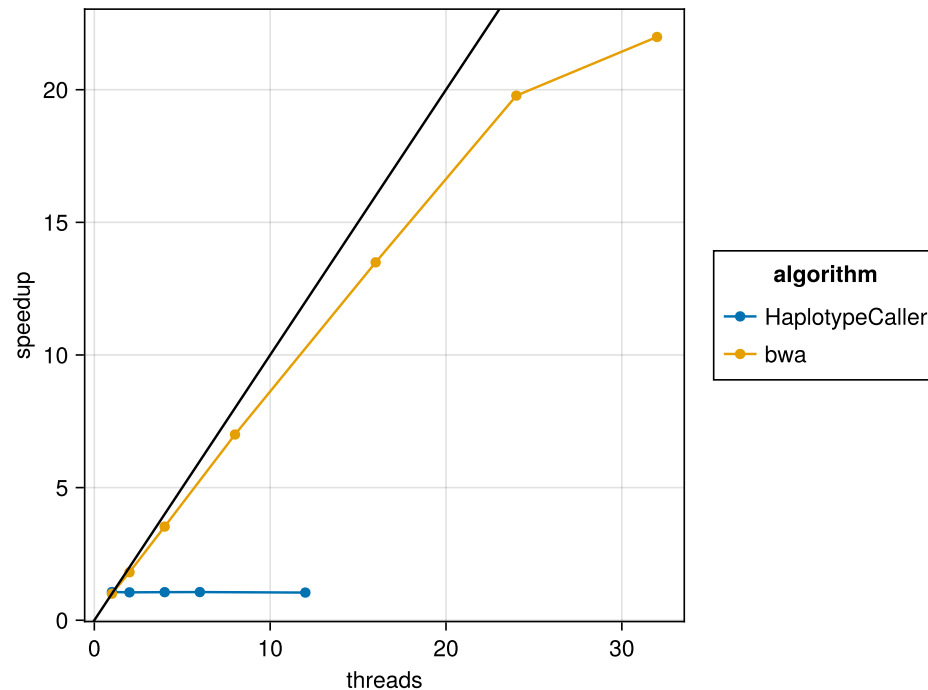
Nextflow

- Maîtrise des risques liés au matériel
- Exécution sur de multiples architectures (super-calculateur...)



1. Performances

- Mésocentre de Franche-Comté
- Parallélisation de l'alignement
- 1 patient = 5h de calcul
- 20 patients/jour



2. Validation

- "Gold standard" sur patients de référence (Genome In A Bottle Consortium)
- *In silico*

2. Validation: séquencage et pipeline

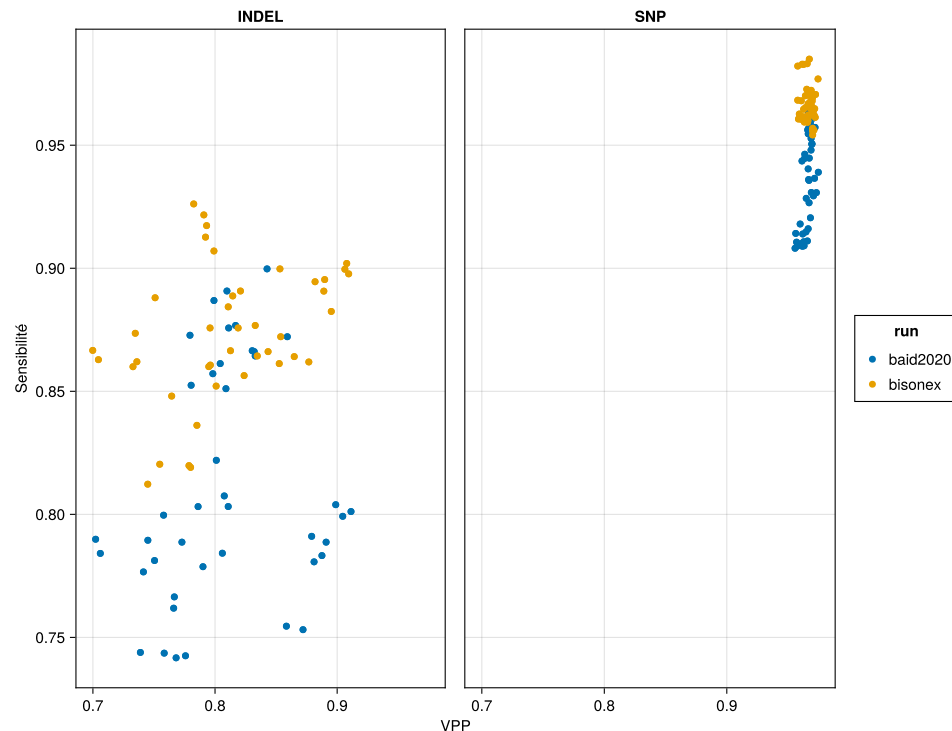
Patient NA12878

- ADN commandé à Coriell
- séquencé à Centogène
- analysé par notre pipeline

Type	Sensibilité	Valeur Prédictive Positive
Indel	0.954	0.775
SNV	0.983	0.965

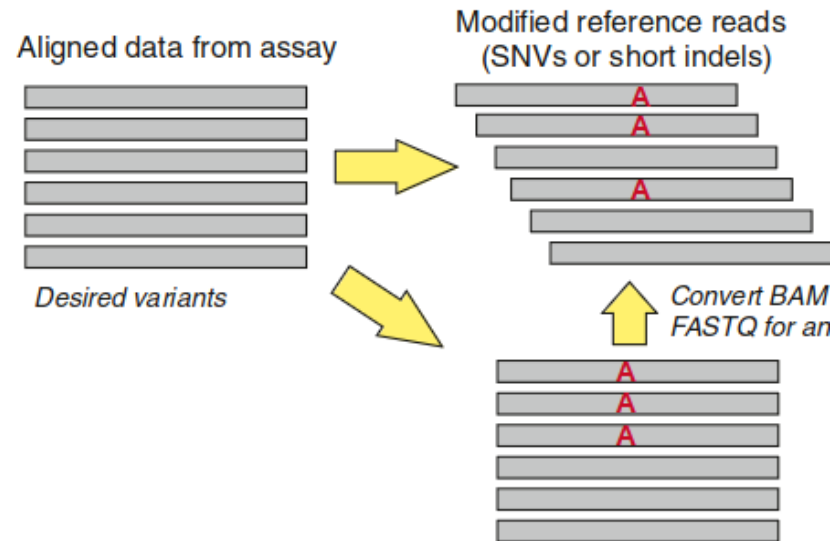
2. Validation : pipeline

- 7 Patients GIAB
- Google (Baid et al. 2022)
 - 3 kits de capture
 - 2 séquenceurs (HiSeq4000, Novaseq)
 - alignement et appel de variant identiques



2. Validation : *in silico*

- Patient de synthèse avec **varben**
- 126 variants confirmés en Sanger
 - 1 non inséré
 - 1 filtré sur la profondeur ($21 < 30$)
 - 124 retrouvés

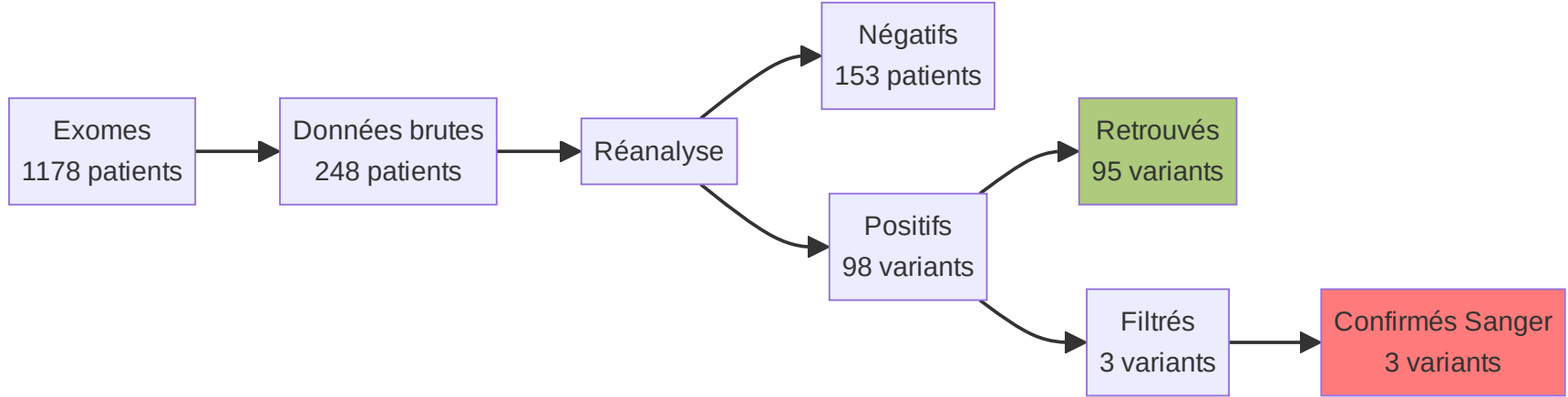


2. Validation : *in silico*

Données simulées avec **simuscop**

- 126 même variants
 - 1 non appelé (3 read sur 34)
 - 3 non insérés
 - 122 retrouvés

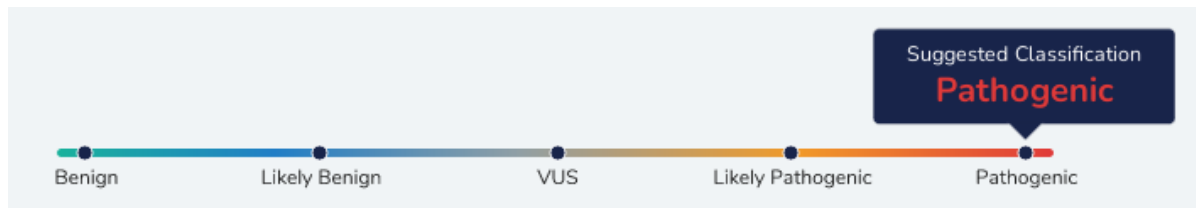
3. Réinterprétation: non-infériorité



- 2 variants filtrés sur la profondeur (*CHD3*, *RRAS2*)
- 1 variant filtré sur le nombre de reads porteurs (*PITX3*)

3. Réinterprétation: nouveaux diagnostics

1 probable diagnostic sur *ADNP*



- très peu fréquent gnomAD
- tronquant sur gène perte de fonction
- clinvar pathologique
- syndrome Helsmoortel-Van der Aa

Predictive scores

pLI	1.00	?
LOEUF	0.12	?
sHet	0.321	?
pHaplo	1.00	?
pTriplo	1.00	?

Germline

Classification ?
★★★★☆ (3) ?

Pathogenic/Likely pathogenic
criteria provided, multiple submitters, no conflicts

3. Réinterprétation: nouveaux diagnostics

- 2 variants intéressants *KCNQ2*, *MED13*
- 1 signification indéterminée (*SPEN*)

Conclusion

- Pipeline reproductible, portable, performant
- Validation : résultats encourageants
- 3 variants prometteurs
- Nombreux axes d'amélioration (T2T, oligénisme...)

Merci de votre attention !



