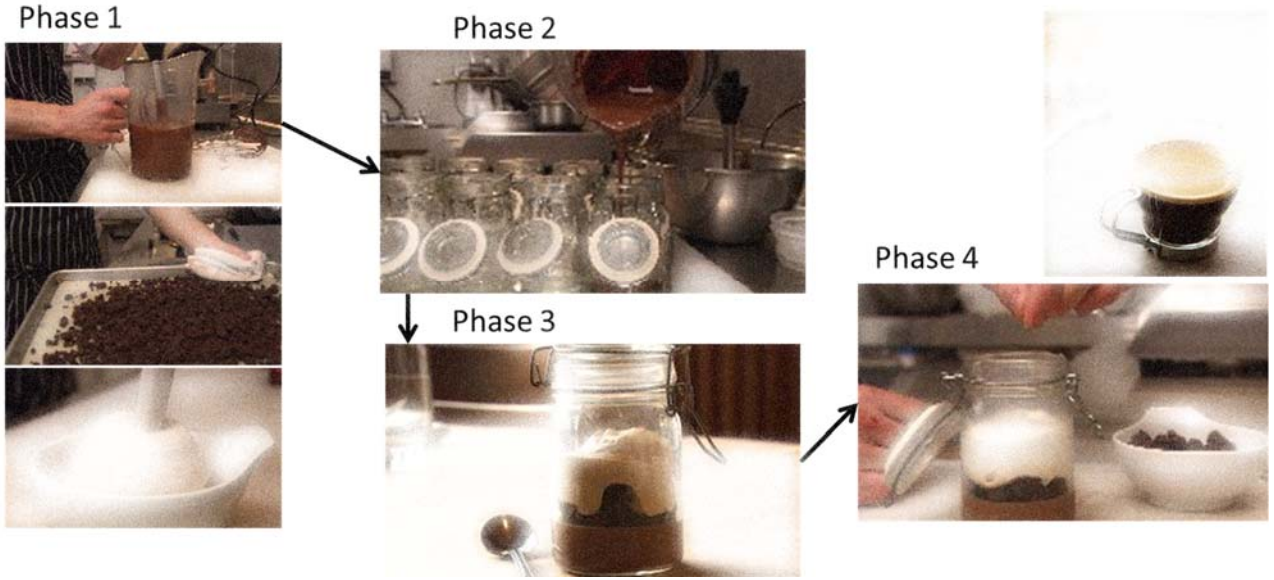
La genèse d’un pipeline d’assemblage de génomes bactériens

ou comment travailler en équipe pour optimiser une analyse à haut débit!



Aujourd'hui, on parle d'informatique!

Cystic Fibrosis Canada
Fibrose kystique Canada
Breathing life into the future®
Donnez le souffle de vie®

frontiers in Microbiology | Evolutionary and Genomic Microbiology

‹ Archive

PERSPECTIVE ARTICLE
Front. Microbiol., 29 September 2015 | http://dx.doi.org/10.3389/fmicb.2015.01036

Clinical utilization of genomics data produced by the international *Pseudomonas aeruginosa* consortium

Luca Freschi[1†], Julie Jeukens[1†], Irena Kukavica-Ibrulj[1†], Brian Boyle[1], Marie-Josée Dupont[1], Jérôme Laroche[1], Stéphane Larose[1], Halim Maaroufi[1], Joanne L. Fothergill[2], Matthew Moore[2], Geoffrey L. Winsor[3], Shawn D. Aaron[4], Jean Barbeau[5], Scott C. Bell[6], Jane L. Burns[7], Miguel Camara[8], André Cantin[9], Steve J. Charette[1,10,11], Ken Dewar[12], Éric Déziel[13], Keith Grimwood[14], Robert E. W. Hancock[15], Joe J. Harrison[16], Stephan Heeb[8], Lars Jelsbak[17], Baofeng Jia[18], Dervla T. Kenna[19], Timothy J. Kidd[20,21], Jens Klockgether[22], Joseph S. Lam[23], Iain L. Lamont[24], Shawn Lewenza[16], Nick Loman[25], François Malouin[9], Jim Manos[26], Andrew G. McArthur[18], Josie McKeown[8], Julie Milot[27], Hardeep Naghra[8], Dao Nguyen[12,28], Sheldon K. Pereira[18], Gabriel G. Perron[29], Jean-Paul Pirnay[30], Paul B. Rainey[31,32], Simon Rousseau[12], Pedro M. Santos[33], Anne Stephenson[34], Véronique Taylor[23], Jane F. Turton[19], Nicholas Waglechner[18], Paul Williams[8], Sandra W. Thrane[17], Gerard D. Wright[18], Fiona S. L. Brinkman[3], Nicholas P. Tucker[35], Burkhard Tümmler[22], Craig Winstanley[2] and Roger C. Levesque[1*]

[1]Institute for Integrative and Systems Biology, Université Laval, Quebec, QC, Canada
[2]Institute of Infection and Global Health, University of Liverpool, Liverpool, UK
[3]Department of Molecular Biology and Biochemistry, Simon Fraser University, Vancouver, BC, Canada
[4]Ottawa Hospital Research Institute, Ottawa, ON, Canada

Download Article
Export citation
3,460
TOTAL VIEWS
35
View Article Impact
Like 0
Comment 0
Share 1
SHARE ON
8 16 3 4 31
SUPPLEMENTAL DATA
Table 1.XLSX

… et d'histoire!: de fin 2013 à aujourd'hui

- Phase 1
  - Déménager sur Katak + paralléliser l'analyse
- Phase 2
  - Faciliter la vie de l'utilisateur
- Phase 3
  - Pipeline «intelligent», changement de serveur, etc.
- Phase 4

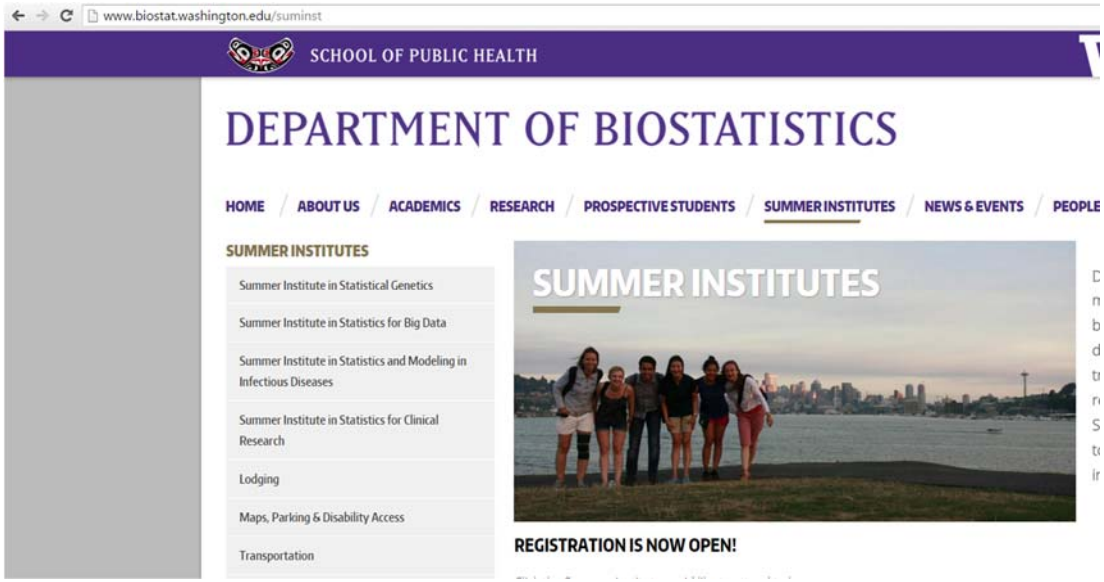
Objectifs

- Assembler plusieurs dizaines de génomes bactériens à la fois
- Aligner ces nouveaux génomes sur des génomes de référence
- Minimiser les interventions de l'utilisateur

Mes compétences en programmation (2005-2011)

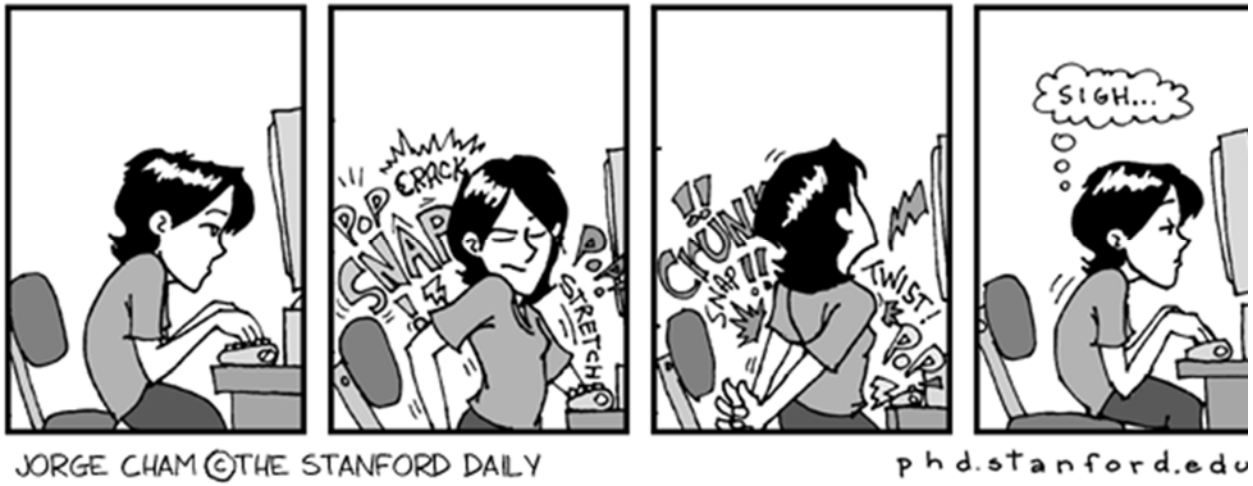
- (Avant 2005, quasi rien)
- Au Marchand depuis mai 2005!
- Quelques formations courtes sur R (analyses de données microarray)

- Un module sur R au SISG2007
- Un peu de bash (serveurs) et beaucoup de R durant mon doc (les débuts du RNA-Seq!)



Mes compétences en programmation (2011-2016)

- Pas mal tous les cours de bioinfo IBIS
- CSHL Advanced sequencing technologies nov. 2013
- Beaucoup de «Katak» depuis 1 an



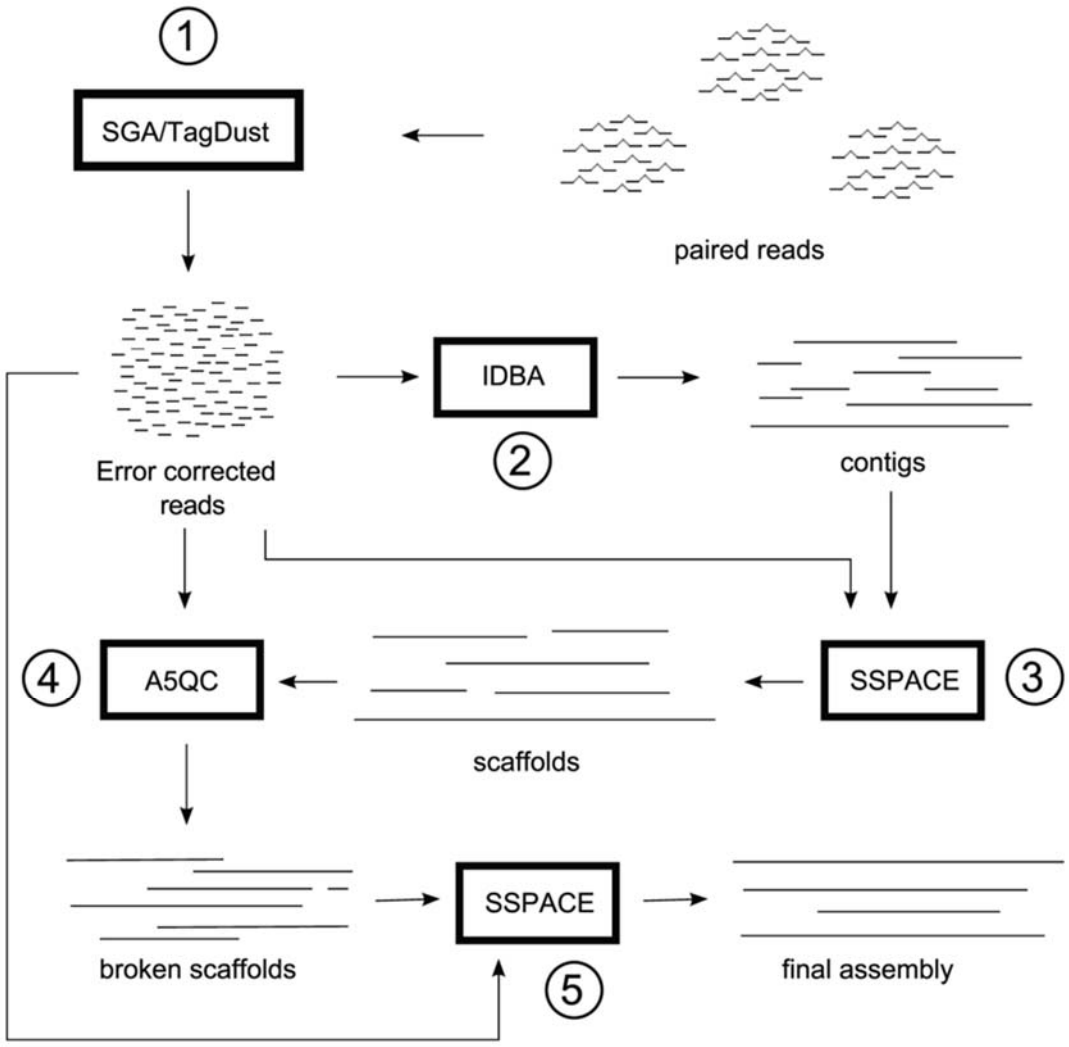
Phase 1, décembre 2013, Julie Jeukens

Google

- Projet pilote pour le projet 1000 génomes de *Pseudomonas*
- 20 génomes

Phase 1

- Test de différents assembleurs
- A5 pipeline (Tritt et al. 2012. PLOS ONE), disponible sur Katak(/prg, **modifié**)



Phase 1

- Sur mon portable (Linux)
- Fichier texte avec des instructions et des lignes de commandes en bash (à copier-coller)
- Tâches:
  - Assembler les génomes en série (2-3 h/génome)
  - Calculer/trier/extraire des statistiques d'assemblage
  - Aligner les nouveaux génomes sur des références (nucmer) et compiler les résultats


Taking a
Baby
Break . . .

Phase 2, printemps 2014, Jérôme Laroche

- Installation sur Katak
- Paralléliser: plusieurs génomes à la fois pour l'étape de l'assemblage avec A5 (la plus longue)
- Format: soumission avec SGE + 7 fichiers .sh
  - 01_PrepDir.sh
  - 02_A5assembly.sh
  - 03_ConcatCSV.sh
  - 04_ScaffoldDepth.sh
  - 05_AlignToRef.sh
  - 06_StatAlignRefs.sh
  - 07_MvResults.sh

Phase 3, hiver 2014-2015, Luca Freschi

- Fichier texte de configuration à éditer au préalable
- Analyse complète (incluant la soumission SGE) contrôlée par une seule commande: do_analysis.py
- Git: logiciel de gestion de versions

Phase 3:

- Ça marche… mais…
  - Bugs d’architecture de dossier
  - Copie manuelle de fichiers finaux
  - Mémoire
  - Etc.

Phase 4, janvier 2015, Luca et Julie

- Mise à jour après plusieurs mois d'utilisation et de prise de notes sur les manques et les irritants
- Nouvelle composante: projet Salmonelle

Phase 4

- Beaucoup de petites corrections dans les scripts existants
- Paralléliser toutes les étapes plus longues
- Choix du projet (pseudo ou salmo)
- Réduction au maximum de l'intervention de l'utilisateur
- Tout est contrôlé par A5-driver.pl

Non! Le A5-bus-driver!!

Nombre de passagers = nombre de génomes
(= nombre de cœurs ;)

Le résultat!

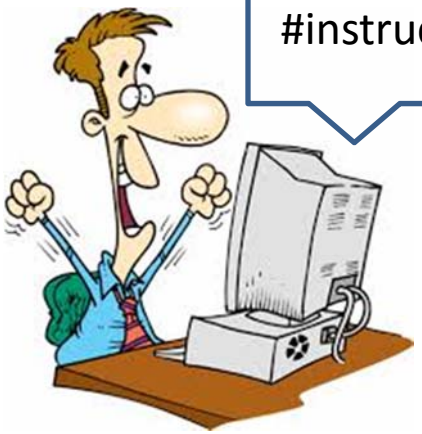
ipcd.ibis.ulaval.ca

salfos.ibis.ulaval.ca

- pipeline
- pseudo
  - données
    - RunX
  - assemblage
- salmo
  - données
  - assemblage

cd RunX
A5-bus-driver.pl

#instructions



Lancer l’assemblage

- jujeu2@katak:~/pseudo1000/miseq_data/Run12a> A5-bus-driver.pl
- ::I found the configuration.py file --OK

- ::Please choose a project [pseudo|salmo], default -- pseudo
- >>pseudo
- ::Project: pseudo

- ::I have detected that you are analyzing the run Run12a

- ::How many genomes do you want to assemble in parallel? -- default: 16
- >>16

- ::Writing all changes on configuration.py
- --sed -i -e 's@run=.*@run="Run12a\/"@' configuration.py
- --sed -i -e 's@dir_genomes=.*@dir_genomes="/project/rclevesq/pseudo/Pseudo1000/miseq_data/Run12a\/"@' configuration.py
- --sed -i -e 's@tasks=.*@tasks=16@' configuration.py
- --sed -i -e 's@dir_refs=.*@dir_refs="/project/rclevesq/pseudo/Pseudo1000/reference_genomes/"@' configuration.py
- --sed -i -e 's@dir_refs=.*@dir_refs="/project/rclevesq/pseudo/Pseudo1000/reference_genomes/"@' configuration.py

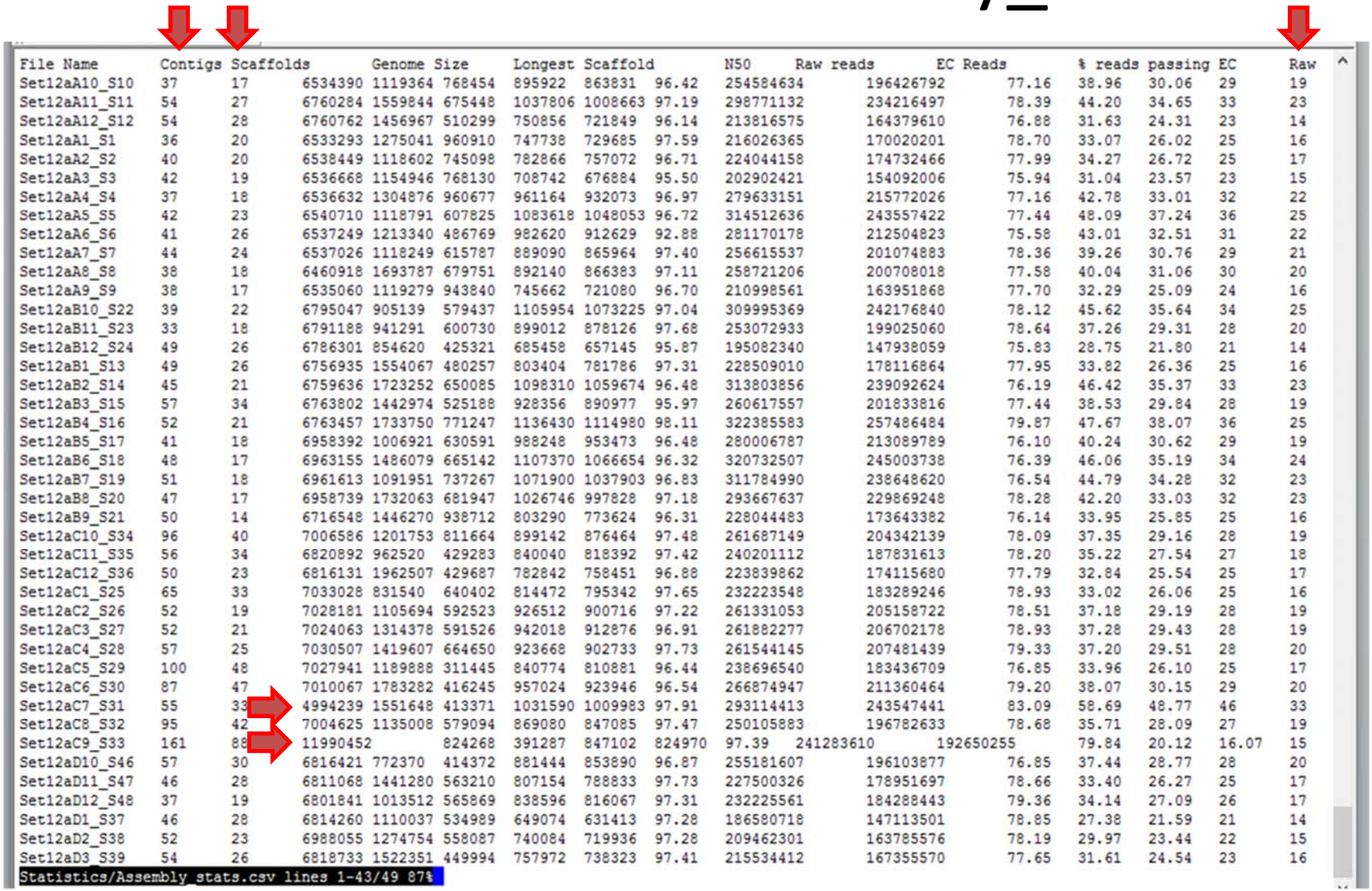
- ::Printing current variables -- so you can check if everything is ok
- dir_scripts="/project/rclevesq/users/lfreschi/pseudo/pipeline_lfreschi_current/"
- dir_genomes="/project/rclevesq/pseudo/Pseudo1000/miseq_data/Run12a/"
- dir_doc="/home/lfreschi/pseudo/doc/"
- dir_refs="/project/rclevesq/pseudo/Pseudo1000/reference_genomes/"
- tasks=16
- smp=1
- project="/project/rclevesq/pseudo/Pseudo1000/data_pseudo/"
- run="Run12a/"

- ::Do you want to start the analysis now [y/n]? -- default: n
- >>n
- ::Bye!

Regarder les résultats

```
jujeu2@katak:~/pseudo1000/miseq_data/Run12a> ls
configuration.py                    Set12aA9_S9                         Set12aB9_S21_L001_R2_001.fastq.gz   Set12aD10_S46_L001_R1_001.fastq.gz
configuration.pyc                   Set12aA9_S9_L001_R1_001.fastq.gz    Set12aC10_S34                       Set12aD10_S46_L001_R2_001.fastq.gz
FastqSummaryF1L1.txt                Set12aA9_S9_L001_R2_001.fastq.gz    Set12aC10_S34_L001_R1_001.fastq.gz  Set12aD11_S47
logs                                Set12aB10_S22                       Set12aC10_S34_L001_R2_001.fastq.gz  Set12aD11_S47_L001_R1_001.fastq.gz
queue                               Set12aB10_S22_L001_R1_001.fastq.gz  Set12aC11_S35                       Set12aD11_S47_L001_R2_001.fastq.gz
Set12aA10_S10                       Set12aB10_S22_L001_R2_001.fastq.gz  Set12aC11_S35_L001_R1_001.fastq.gz  Set12aD12_S48
Set12aA10_S10_L001_R1_001.fastq.gz  Set12aB11_S23                       Set12aC11_S35_L001_R2_001.fastq.gz  Set12aD12_S48_L001_R1_001.fastq.gz
Set12aA10_S10_L001_R2_001.fastq.gz  Set12aB11_S23_L001_R1_001.fastq.gz  Set12aC12_S36                       Set12aD12_S48_L001_R2_001.fastq.gz
Set12aA11_S11                       Set12aB11_S23_L001_R2_001.fastq.gz  Set12aC12_S36_L001_R1_001.fastq.gz  Set12aD1_S37
Set12aA11_S11_L001_R1_001.fastq.gz  Set12aB12_S24                       Set12aC12_S36_L001_R2_001.fastq.gz  Set12aD1_S37_L001_R1_001.fastq.gz
Set12aA11_S11_L001_R2_001.fastq.gz  Set12aB12_S24_L001_R1_001.fastq.gz  Set12aC1_S25                        Set12aD1_S37_L001_R2_001.fastq.gz
Set12aA12_S12                       Set12aB12_S24_L001_R2_001.fastq.gz  Set12aC1_S25_L001_R1_001.fastq.gz   Set12aD2_S38
Set12aA12_S12_L001_R1_001.fastq.gz  Set12aB1_S13                        Set12aC1_S25_L001_R2_001.fastq.gz   Set12aD2_S38_L001_R1_001.fastq.gz
Set12aA12_S12_L001_R2_001.fastq.gz  Set12aB1_S13_L001_R1_001.fastq.gz   Set12aC2_S26                        Set12aD2_S38_L001_R2_001.fastq.gz
Set12aA1_S1                         Set12aB1_S13_L001_R2_001.fastq.gz   Set12aC2_S26_L001_R1_001.fastq.gz   Set12aD3_S39
Set12aA1_S1_L001_R1_001.fastq.gz    Set12aB2_S14                        Set12aC2_S26_L001_R2_001.fastq.gz   Set12aD3_S39_L001_R1_001.fastq.gz
Set12aA1_S1_L001_R2_001.fastq.gz    Set12aB2_S14_L001_R1_001.fastq.gz   Set12aC3_S27                        Set12aD3_S39_L001_R2_001.fastq.gz
Set12aA2_S2                         Set12aB2_S14_L001_R2_001.fastq.gz   Set12aC3_S27_L001_R1_001.fastq.gz   Set12aD4_S40
Set12aA2_S2_L001_R1_001.fastq.gz    Set12aB3_S15                        Set12aC3_S27_L001_R2_001.fastq.gz   Set12aD4_S40_L001_R1_001.fastq.gz
Set12aA2_S2_L001_R2_001.fastq.gz    Set12aB3_S15_L001_R1_001.fastq.gz   Set12aC4_S28                        Set12aD4_S40_L001_R2_001.fastq.gz
Set12aA3_S3                         Set12aB3_S15_L001_R2_001.fastq.gz   Set12aC4_S28_L001_R1_001.fastq.gz   Set12aD5_S41
Set12aA3_S3_L001_R1_001.fastq.gz    Set12aB4_S16                        Set12aC4_S28_L001_R2_001.fastq.gz   Set12aD5_S41_L001_R1_001.fastq.gz
Set12aA3_S3_L001_R2_001.fastq.gz    Set12aB4_S16_L001_R1_001.fastq.gz   Set12aC5_S29                        Set12aD5_S41_L001_R2_001.fastq.gz
Set12aA4_S4                         Set12aB4_S16_L001_R2_001.fastq.gz   Set12aC5_S29_L001_R1_001.fastq.gz   Set12aD6_S42
Set12aA4_S4_L001_R1_001.fastq.gz    Set12aB5_S17                        Set12aC5_S29_L001_R2_001.fastq.gz   Set12aD6_S42_L001_R1_001.fastq.gz
Set12aA4_S4_L001_R2_001.fastq.gz    Set12aB5_S17_L001_R1_001.fastq.gz   Set12aC6_S30                        Set12aD6_S42_L001_R2_001.fastq.gz
Set12aA5_S5                         Set12aB5_S17_L001_R2_001.fastq.gz   Set12aC6_S30_L001_R1_001.fastq.gz   Set12aD7_S43
Set12aA5_S5_L001_R1_001.fastq.gz    Set12aB6_S18                        Set12aC6_S30_L001_R2_001.fastq.gz   Set12aD7_S43_L001_R1_001.fastq.gz
Set12aA5_S5_L001_R2_001.fastq.gz    Set12aB6_S18_L001_R1_001.fastq.gz   Set12aC7_S31                        Set12aD7_S43_L001_R2_001.fastq.gz
Set12aA6_S6                         Set12aB6_S18_L001_R2_001.fastq.gz   Set12aC7_S31_L001_R1_001.fastq.gz   Set12aD8_S44
Set12aA6_S6_L001_R1_001.fastq.gz    Set12aB7_S19                        Set12aC7_S31_L001_R2_001.fastq.gz   Set12aD8_S44_L001_R1_001.fastq.gz
Set12aA6_S6_L001_R2_001.fastq.gz    Set12aB7_S19_L001_R1_001.fastq.gz   Set12aC8_S32                        Set12aD8_S44_L001_R2_001.fastq.gz
Set12aA7_S7                         Set12aB7_S19_L001_R2_001.fastq.gz   Set12aC8_S32_L001_R1_001.fastq.gz   Set12aD9_S45
Set12aA7_S7_L001_R1_001.fastq.gz    Set12aB8_S20                        Set12aC8_S32_L001_R2_001.fastq.gz   Set12aD9_S45_L001_R1_001.fastq.gz
Set12aA7_S7_L001_R2_001.fastq.gz    Set12aB8_S20_L001_R1_001.fastq.gz   Set12aC9_S33                        Set12aD9_S45_L001_R2_001.fastq.gz
Set12aA8_S8                         Set12aB8_S20_L001_R2_001.fastq.gz   Set12aC9_S33_L001_R1_001.fastq.gz   Statistics
Set12aA8_S8_L001_R1_001.fastq.gz    Set12aB9_S21                        Set12aC9_S33_L001_R2_001.fastq.gz   WG_Alignments
Set12aA8_S8_L001_R2_001.fastq.gz    Set12aB9_S21_L001_R1_001.fastq.gz   Set12aD10_S46
jujeu2@katak:~/pseudo1000/miseq_data/Run12a>
```

2 fichiers résumés: Assembly_stats.csv



File Name	Contigs	Scaffolds		Genome Size			Longest Scaffold		N50	Raw reads		EC Reads		% reads passing EC		Raw
Set12aA10_S10	37	17	6534390	1119364	768454	895922	863831	96.42	254584634	196426792	77.16	38.96	30.06	29	19	
Set12aA11_S11	54	27	6760284	1559844	675448	1037806	1008663	97.19	298771132	234216497	78.39	44.20	34.65	33	23	
Set12aA12_S12	54	28	6760762	1456967	510299	750856	721849	96.14	213816575	164379610	76.88	31.63	24.31	23	14	
Set12aA1_S1	36	20	6533293	1275041	960910	747738	729685	97.59	216026365	170020201	78.70	33.07	26.02	25	16	
Set12aA2_S2	40	20	6538449	1118602	745098	782866	757072	96.71	224044158	174732466	77.99	34.27	26.72	25	17	
Set12aA3_S3	42	19	6536668	1154946	768130	708742	676884	95.50	202902421	154092006	75.94	31.04	23.57	23	15	
Set12aA4_S4	37	18	6536632	1304876	960677	961164	932073	96.97	279633151	215772026	77.16	42.78	33.01	32	22	
Set12aA5_S5	42	23	6540710	1118791	607825	1083618	1048053	96.72	314512636	243557422	77.44	48.09	37.24	36	25	
Set12aA6_S6	41	26	6537249	1213340	486769	982620	912629	92.88	281170178	212504823	75.58	43.01	32.51	31	22	
Set12aA7_S7	44	24	6537026	1118249	615787	889090	865964	97.40	256615537	201074883	78.36	39.26	30.76	29	21	
Set12aA8_S8	38	18	6460918	1693787	679751	892140	866383	97.11	258721206	200708018	77.58	40.04	31.06	30	20	
Set12aA9_S9	38	17	6535060	1119279	943840	745662	721080	96.70	210998561	163951868	77.70	32.29	25.09	24	16	
Set12aB10_S22	39	22	6795047	905139	579437	1105954	1073225	97.04	309995369	242176840	78.12	45.62	35.64	34	25	
Set12aB11_S23	33	18	6791188	941291	600730	899012	878126	97.68	253072933	199025060	78.64	37.26	29.31	28	20	
Set12aB12_S24	49	26	6786301	854620	425321	685458	657145	95.87	195082340	147938059	75.83	28.75	21.80	21	14	
Set12aB1_S13	49	26	6756935	1554067	480257	803404	781786	97.31	228509010	178116864	77.95	33.82	26.36	25	16	
Set12aB2_S14	45	21	6759636	1723252	650085	1098310	1059674	96.48	313803856	239092624	76.19	46.42	35.37	33	23	
Set12aB3_S15	57	34	6763802	1442974	525188	928356	890977	95.97	260617557	201833816	77.44	38.53	29.84	28	19	
Set12aB4_S16	52	21	6763457	1733750	771247	1136430	1114980	98.11	322385583	257486484	79.87	47.67	38.07	36	25	
Set12aB5_S17	41	18	6958392	1006921	630591	988248	953473	96.48	280006787	213089789	76.10	40.24	30.62	29	19	
Set12aB6_S18	48	17	6963155	1486079	665142	1107370	1066654	96.32	320732507	245003738	76.39	46.06	35.19	34	24	
Set12aB7_S19	51	18	6961613	1091951	737267	1071900	1037903	96.83	311784990	238648620	76.54	44.79	34.28	32	23	
Set12aB8_S20	47	17	6958739	1732063	681947	1026746	997828	97.18	293667637	229869248	78.28	42.20	33.03	32	23	
Set12aB9_S21	50	14	6716548	1446270	938712	803290	773624	96.31	228044483	173643382	76.14	33.95	25.85	25	16	
Set12aC10_S34	96	40	7006586	1201753	811664	899142	876464	97.48	261687149	204342139	78.09	37.35	29.16	28	19	
Set12aC11_S35	56	34	6820892	962520	429283	840040	818392	97.42	240201112	187831613	78.20	35.22	27.54	27	18	
Set12aC12_S36	50	23	6816131	1962507	429687	782842	758451	96.88	223839862	174115680	77.79	32.84	25.54	25	17	
Set12aC1_S25	65	33	7033028	831540	640402	814472	795342	97.65	232223548	183289246	78.93	33.02	26.06	25	16	
Set12aC2_S26	52	19	7028181	1105694	592523	926512	900716	97.22	261331053	205158722	78.51	37.18	29.19	28	19	
Set12aC3_S27	52	21	7024063	1314378	591526	942018	912876	96.91	261882277	206702178	78.93	37.28	29.43	28	19	
Set12aC4_S28	57	25	7030507	1419607	664650	923668	902733	97.73	261544145	207481439	79.33	37.20	29.51	28	20	
Set12aC5_S29	100	48	7027941	1189888	311445	840774	810881	96.44	238696540	183436709	76.85	33.96	26.10	25	17	
Set12aC6_S30	87	47	7010067	1783282	416245	957024	923946	96.54	266874947	211360464	79.20	38.07	30.15	29	20	
Set12aC7_S31	55	33	4994239	1551648	413371	1031590	1009983	97.91	293114413	243547441	83.09	58.69	48.77	46	33	
Set12aC8_S32	95	42	7004625	1135008	579094	869080	847085	97.47	250105883	196782633	78.68	35.71	28.09	27	19	
Set12aC9_S33	161	88	11990452		824268	391287	847102	824970	97.39	241283610	192650255	79.84	20.12	16.07	15	
Set12aD10_S46	57	30	6816421	772370	414372	881444	853890	96.87	255181607	196103877	76.85	37.44	28.77	28	20	
Set12aD11_S47	46	28	6811068	1441280	563210	807154	788833	97.73	227500326	178951697	78.66	33.40	26.27	25	17	
Set12aD12_S48	37	19	6801841	1013512	565869	838596	816067	97.31	232225561	184288443	79.36	34.14	27.09	26	17	
Set12aD1_S37	46	28	6814260	1110037	534989	649074	631413	97.28	186580718	147113501	78.85	27.38	21.59	21	14	
Set12aD2_S38	52	23	6988055	1274754	558087	740084	719936	97.28	209462301	163785576	78.19	29.97	23.44	22	15	
Set12aD3_S39	54	26	6818733	1522351	449994	757972	738323	97.41	215534412	167355570	77.65	31.61	24.54	23	16	

Statistics/Assembly_stats.csv lines 1-43/49 87%

2 fichiers résumés: Assembly_stats.csv

	N50	Raw reads	EC Reads	% reads passing EC			Raw nt	EC nt	% nt passing EC	Raw cov	EC cov	Median cov
.42	254584634	196426792	77.16	38.96	30.06	29	19	6530302	66.3	A5-miseq	20150522	
.19	298771132	234216497	78.39	44.20	34.65	33	23	6755151	66.2	A5-miseq	20150522	
.14	213816575	164379610	76.88	31.63	24.31	23	14	6749379	66.2	A5-miseq	20150522	
.59	216026365	170020201	78.70	33.07	26.02	25	16	6528122	66.3	A5-miseq	20150522	
.71	224044158	174732466	77.99	34.27	26.72	25	17	6532321	66.3	A5-miseq	20150522	
.50	202902421	154092006	75.94	31.04	23.57	23	15	6528430	66.3	A5-miseq	20150522	
.97	279633151	215772026	77.16	42.78	33.01	32	22	6532263	66.3	A5-miseq	20150522	
.72	314512636	243557422	77.44	48.09	37.24	36	25	6535386	66.3	A5-miseq	20150522	
.88	281170178	212504823	75.58	43.01	32.51	31	22	6532628	66.3	A5-miseq	20150522	
.40	256615537	201074883	78.36	39.26	30.76	29	21	6531581	66.3	A5-miseq	20150522	
.11	258721206	200708018	77.58	40.04	31.06	30	20	6456760	66.3	A5-miseq	20150522	
.70	210998561	163951868	77.70	32.29	25.09	24	16	6528845	66.3	A5-miseq	20150522	
.04	309995369	242176840	78.12	45.62	35.64	34	25	6790822	66.2	A5-miseq	20150522	
.68	253072933	199025060	78.64	37.26	29.31	28	20	6787216	66.2	A5-miseq	20150522	
.87	195082340	147938059	75.83	28.75	21.80	21	14	6777303	66.2	A5-miseq	20150522	
.31	228509010	178116864	77.95	33.82	26.36	25	16	6748960	66.2	A5-miseq	20150522	
.48	313803856	239092624	76.19	46.42	35.37	33	23	6755374	66.2	A5-miseq	20150522	
.97	260617557	201833816	77.44	38.53	29.84	28	19	6756349	66.2	A5-miseq	20150522	
.11	322385583	257486484	79.87	47.67	38.07	36	25	6759070	66.2	A5-miseq	20150522	
.48	280006787	213089789	76.10	40.24	30.62	29	19	6954285	66.0	A5-miseq	20150522	
.32	320732507	245003738	76.39	46.06	35.19	34	24	6956982	66.0	A5-miseq	20150522	
.83	311784990	238648620	76.54	44.79	34.28	32	23	6956261	66.0	A5-miseq	20150522	
.18	293667637	229869248	78.28	42.20	33.03	32	23	6954318	66.0	A5-miseq	20150522	
.31	228044483	173643382	76.14	33.95	25.85	25	16	6708790	66.1	A5-miseq	20150522	
.48	261687149	204342139	78.09	37.35	29.16	28	19	6993010	66.1	A5-miseq	20150522	
.42	240201112	187831613	78.20	35.22	27.54	27	18	6812171	66.2	A5-miseq	20150522	
.88	223839862	174115680	77.79	32.84	25.54	25	17	6809313	66.2	A5-miseq	20150522	
.65	232223548	183289246	78.93	33.02	26.06	25	16	7023698	66.0	A5-miseq	20150522	
.22	261331053	205158722	78.51	37.18	29.19	28	19	7021794	66.0	A5-miseq	20150522	
.91	261882277	206702178	78.93	37.28	29.43	28	19	7018395	66.0	A5-miseq	20150522	
.73	261544145	207481439	79.33	37.20	29.51	28	20	7022499	66.0	A5-miseq	20150522	
.44	238696540	183436709	76.85	33.96	26.10	25	17	7013581	66.0	A5-miseq	20150522	
.54	266874947	211360464	79.20	38.07	30.15	29	20	7000800	66.1	A5-miseq	20150522	
.91	293114413	243547441	83.09	58.69	48.77	46	33	4990458	55.3	A5-miseq	20150522	
.47	250105883	196782633	78.68	35.71	28.09	27	19	6992059	66.1	A5-miseq	20150522	
4970	97.39	241283610	192650255	79.84	20.12	16.07	15	9	11837307	61.6	A5-miseq	20150522
.87	255181607	196103877	76.85	37.44	28.77	28	20	6809910	66.2	A5-miseq	20150522	
.73	227500326	178951697	78.66	33.40	26.27	25	17	6805179	66.2	A5-miseq	20150522	
.31	232225561	184288443	79.36	34.14	27.09	26	17	6795791	66.2	A5-miseq	20150522	
.28	186580718	147113501	78.85	27.38	21.59	21	14	6803703	66.2	A5-miseq	20150522	
.28	209462301	163785576	78.19	29.97	23.44	22	15	6979343	65.9	A5-miseq	20150522	
.41	215534412	167355570	77.65	31.61	24.54	23	16	6808730	66.2	A5-miseq	20150522	
0418	97.66	241710840	193371384	80.00	20.47	16.37	15	9	11668381	61.6	A5-miseq	20150522

Connected to katak.ibis.ulaval.ca SSH2 - aes128-cbc - hmac-md5 - n(148x44

2 fichiers résumés: PercentAlignAll

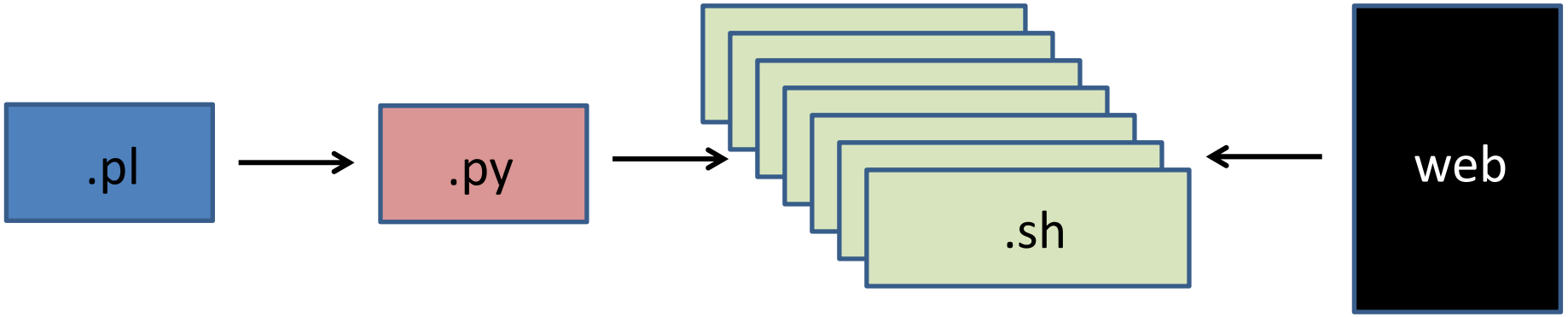
Genome	reference	alignment_length	genome_size_estimate	percent_align
Set12aA10_S10	A_xylosoxidans	0	6534390	0
Set12aA10_S10	LESB58	6.0299e+06	6534390	92.2795
Set12aA10_S10	PA01	6.00013e+06	6534390	91.8239
Set12aA10_S10	PA14	6.01993e+06	6534390	92.1269
Set12aA10_S10	PA7	1.66688e+06	6534390	25.5093
Set12aA11_S11	A_xylosoxidans	21960.5	6760284	0.324846
Set12aA11_S11	LESB58	6.10332e+06	6760284	90.2821
Set12aA11_S11	PA01	6.04801e+06	6760284	89.4638
Set12aA11_S11	PA14	6.05473e+06	6760284	89.5632
Set12aA11_S11	PA7	1.67438e+06	6760284	24.7679
Set12aA12_S12	A_xylosoxidans	21967.9	6760762	0.324932
Set12aA12_S12	LESB58	6.10759e+06	6760762	90.3388
Set12aA12_S12	PA01	6.05233e+06	6760762	89.5215
Set12aA12_S12	PA14	6.06128e+06	6760762	89.6538
Set12aA12_S12	PA7	1.66218e+06	6760762	24.5857
Set12aA1_S1	A_xylosoxidans	0	6533293	0
Set12aA1_S1	LESB58	6.0267e+06	6533293	92.246
Set12aA1_S1	PA01	5.99529e+06	6533293	91.7652
Set12aA1_S1	PA14	6.0159e+06	6533293	92.0807
Set12aA1_S1	PA7	1644287	6533293	25.1678
Set12aA2_S2	A_xylosoxidans	0	6538449	0
Set12aA2_S2	LESB58	6.03081e+06	6538449	92.2362
Set12aA2_S2	PA01	6.00091e+06	6538449	91.7788
Set12aA2_S2	PA14	6.02069e+06	6538449	92.0813
Set12aA2_S2	PA7	1.6479e+06	6538449	25.2032
Set12aA3_S3	A_xylosoxidans	0	6536668	0
Set12aA3_S3	LESB58	6.03408e+06	6536668	92.3112
Set12aA3_S3	PA01	5.99903e+06	6536668	91.7751
Set12aA3_S3	PA14	6.01937e+06	6536668	92.0863
Set12aA3_S3	PA7	1.64507e+06	6536668	25.1667
Set12aA4_S4	A_xylosoxidans	0	6536632	0
Set12aA4_S4	LESB58	6.02181e+06	6536632	92.124
Set12aA4_S4	PA01	5.99495e+06	6536632	91.7132
Set12aA4_S4	PA14	6.01607e+06	6536632	92.0362
Set12aA4_S4	PA7	1.68619e+06	6536632	25.796
Set12aA5_S5	A_xylosoxidans	0	6540710	0
Set12aA5_S5	LESB58	6.02461e+06	6540710	92.1095
Set12aA5_S5	PA01	5.99946e+06	6540710	91.7248
Set12aA5_S5	PA14	6.01715e+06	6540710	91.9953
Set12aA5_S5	PA7	1.6554e+06	6540710	25.3091
Set12aA6_S6	A_xylosoxidans	0	6537249	0
Set12aA6_S6	LESB58	6.02363e+06	6537249	92.1431

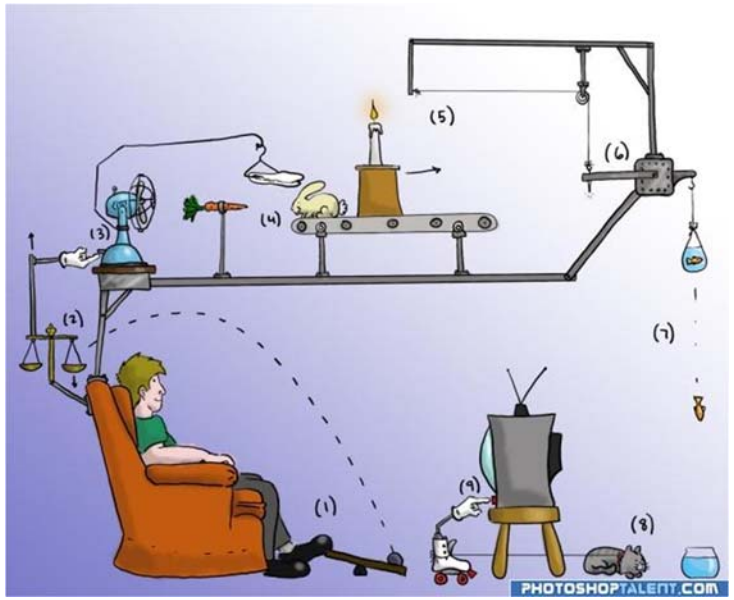
Statistics/PercentAlignAll.txt lines 1-43/241 17%

2 fichiers résumés: PercentAlignAll

```
jujeu2@katak:~/pseudo1000/miseq_data/Run12a> less -S Statistics/PercentAlignAll.txt | grep PA01
Set12aA10_S10 PA01 6.00013e+06 6534390 91.8239
Set12aA11_S11 PA01 6.04801e+06 6760284 89.4638
Set12aA12_S12 PA01 6.05233e+06 6760762 89.5215
Set12aA1_S1 PA01 5.99529e+06 6533293 91.7652
Set12aA2_S2 PA01 6.00091e+06 6538449 91.7788
Set12aA3_S3 PA01 5.99903e+06 6536668 91.7751
Set12aA4_S4 PA01 5.99495e+06 6536632 91.7132
Set12aA5_S5 PA01 5.99946e+06 6540710 91.7248
Set12aA6_S6 PA01 6.00059e+06 6537249 91.7907
Set12aA7_S7 PA01 5.99754e+06 6537026 91.7472
Set12aA8_S8 PA01 6.00183e+06 6460918 92.8944
Set12aA9_S9 PA01 5.99628e+06 6535060 91.7556
Set12aB10_S22 PA01 6.01031e+06 6795047 88.4514
Set12aB11_S23 PA01 6.00864e+06 6791188 88.477
Set12aB12_S24 PA01 6.01125e+06 6786301 88.5792
Set12aB1_S13 PA01 6.05126e+06 6756935 89.5563
Set12aB2_S14 PA01 6.05162e+06 6759636 89.5258
Set12aB3_S15 PA01 6.05261e+06 6763802 89.4853
Set12aB4_S16 PA01 6.05548e+06 6763457 89.5323
Set12aB5_S17 PA01 6.01884e+06 6958392 86.4976
Set12aB6_S18 PA01 6.02111e+06 6963155 86.471
Set12aB7_S19 PA01 6.01579e+06 6961613 86.4137
Set12aB8_S20 PA01 6.02104e+06 6958739 86.5249
Set12aB9_S21 PA01 6.01936e+06 6716548 89.6199
Set12aC10_S34 PA01 6.01855e+06 7006586 85.8985
Set12aC11_S35 PA01 6.05669e+06 6820892 88.7962
Set12aC12_S36 PA01 6.05103e+06 6816131 88.7752
Set12aC1_S25 PA01 6.02662e+06 7033028 85.6902
Set12aC2_S26 PA01 6.02509e+06 7028181 85.7276
Set12aC3_S27 PA01 6.02855e+06 7024063 85.8271
Set12aC4_S28 PA01 6.02495e+06 7030507 85.6973
Set12aC5_S29 PA01 6.01926e+06 7027941 85.6476
Set12aC6_S30 PA01 6.01588e+06 7010067 85.8178
Set12aC7_S31 PA01 882 4994239 0.0176603
Set12aC8_S32 PA01 6.01192e+06 7004625 85.8279
Set12aC9_S33 PA01 6.01616e+06 11990452 50.1746
Set12aD10_S46 PA01 6.05769e+06 6816421 88.8691
Set12aD11_S47 PA01 6.04606e+06 6811068 88.7682
Set12aD12_S48 PA01 6.00959e+06 6801841 88.3524
Set12aD1_S37 PA01 6.05286e+06 6814260 88.8264
Set12aD2_S38 PA01 6.02566e+06 6988055 86.228
Set12aD3_S39 PA01 6.04958e+06 6818733 88.72
```

Mais...



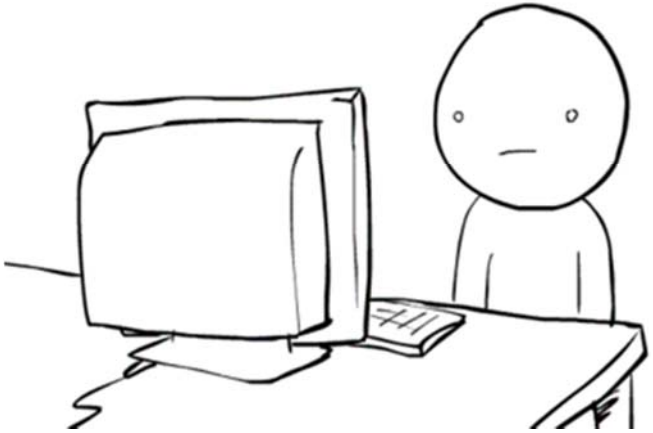


Conclusion

- Objectifs atteints!

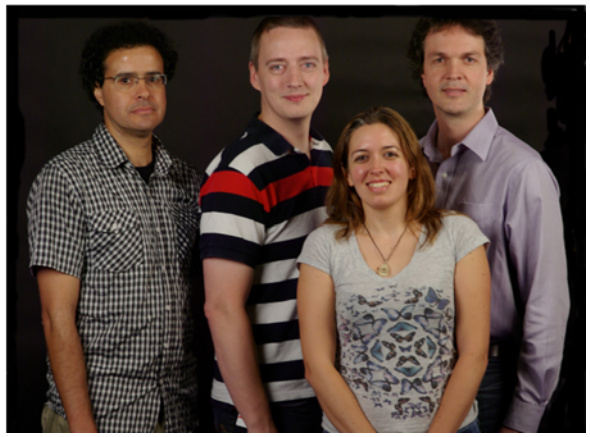
En fait…

- Ma faute: Mise à jour de A5
- Pas ma faute: soumission avec Slurm



- Autres ajouts potentiels au pipeline

Merci!



Nouveau site web!!!
http://rclevesque.ibis.ulaval.ca/fr/accueil-2/