DE LA RECHERCHE À L'INDUSTRIE

NGL_BI

INTERFACES DE SUIVI ET DE VALIDATION DES RUNS, READSETS ET ANALYSES

DOCUMENTATION UTILISATEURS

www.cea.fr

Actualisée pour NGL-BI 1.9 – Février 2015

INTRODUCTION

NGL : « Next Generation LIMS »

5 applications :

- NGL-P : Gestion de Projet *[en cours de développement]*
- NGL-S : Gestion des échantillons (Sample) *[en cours de développement]*
- NGL-SQ : Gestion des expériences de Sequençage *[en cours de développement]*
- NGL-BI : Gestion des analyses primaires de Bio-Informatique *[en production au CNG depuis janvier 2014 et au CNS depuis fin avril 2014]*
- NGL-SUB : Gestion des soumissions *[en cours de développement]*

NGL-BI :

- Interfaces de suivi et d'évaluation des **RUNS, READSETS** (lots de séquences) et **ANALYSES** ;
- Visualisation des statistiques issues des pipelines **NGS-RG, NGS-QC** et **NGS-BA** ;
- **URL** : http://ngl-bi.cng.fr et http://ngl-bi.genoscope.cns.fr → Authentification requise ;
- **Navigateur** recommandé : Chrome
- **Documentation utilisateur** dans Alfresco : https://intra.institut-genomique.fr/share/page/ site NGL > espace documentaire > NGL-BI doc utilisateur ;

SOMMAIRE

- S'authentifier et utiliser les raccourcis de la page d'accueil p.4
- Rechercher des runs, readsets ou analyses p.7
- Visualiser les détails et les traitements :
  - d'un run *(et Run Status Flow)* p.14
  - d'un readset *(et Readset Status Flow)* et export PDF des résultats p.20
  - d'une analyse p.39
- Visualiser des readsets ou analyses selon une configuration de tableau spécifique p.41
- Effectuer du reporting sur les readsets *(colonnes supplémentaires)* p.43
- Générer des représentations graphiques de statistiques sur les readsets p.45
- Évaluer des runs, readsets ou analyses et générer des highlights via des critères d'évaluation :
  - Évaluation runs et highlights p.50
  - Évaluation readsets et highlights p.56
  - Évaluation analyses et highlights p.62
- Fonctionnalités du tableau de résultats p.66
  - Éditer dans un tableau p.68
  - Exporter les résultats en CSV p.71
  - « Grouper » des résultats p.73
- Renseigner des runs / des readsets :
  - *Run : A conserver ?* p.77
  - *Readset : Envoyé au CCRT ? / Envoyé au collaborateur ?* p.78
- Changer l'état d'un run, readset ou analyse (se référer au workflow) p.79
- Workflow run, readset et analyse p.85
  - Procédure d'import SAV p.92
- ANNEXES *(Lexique, Liens vers documents report config. et critères évaluation)* p.104

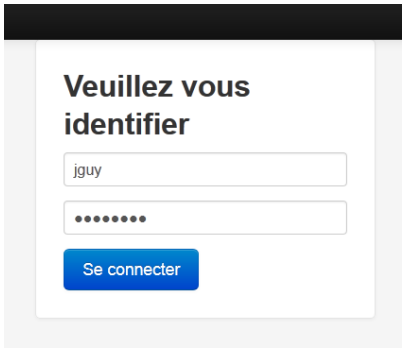
S’authentifier et utiliser les raccourcis de la page d’accueil

HOME - AUTHENTIFICATION

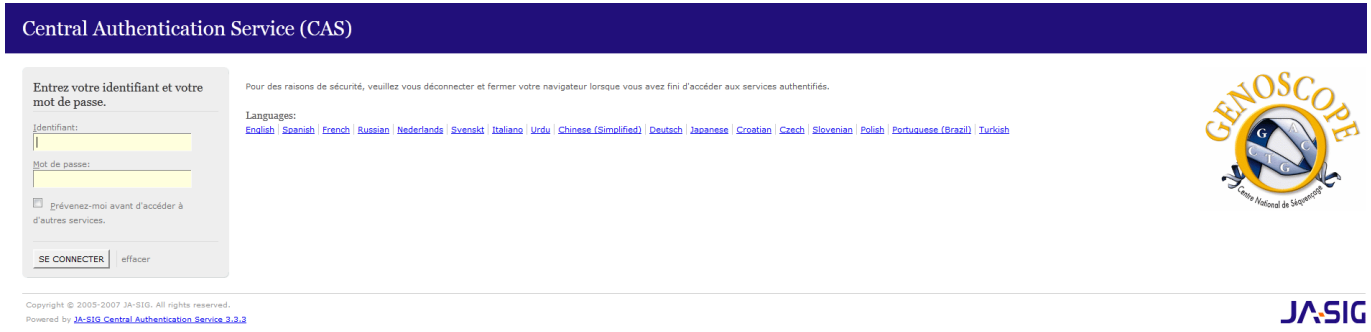
Pour se connecter à NGL-BI, se rendre sur l'url suivante et renseigner son login et mot de passe :

- CNG : http://ngl-bi.cng.fr
- CNS : http://ngl-bi.genoscope.cns.fr

CNG



CNS



HOME - RACCOURCIS

Une fois connecté, des raccourcis sont déjà disponibles sur la page d'accueil pour visualiser la liste :

- des runs *en cours de séquençage* ;
- des runs *en cours de transfert* ;
- des runs *en attente d'évaluation* ;
- des runs taggés *« à conserver ».*

NGL-BI Evaluation Runs Readsets Archives jguy

Bienvenue *login* sur NGL - Bio Informatique

NGL-BI est un LIMS simple et flexible pour le suivi des données bioinformatique issues du séquençage de nouvelle génération.

Consulter la liste directement sur la page d'accueil ou en utilisant le lien html.

Runs en cours de séquençage

1 2 3 Taille (5) 12 Résultat(s)

Code	Type	Date Run	Etat	Valide ?
140207_HISEQ3_C3CFLACXX	RHS2000	07/02/2014	Séquençage en cours	---
140207_HISEQ3_C3CDCACXX	RHS2000	07/02/2014	Séquençage en cours	---
140218_HISEQ1_C3U3TACXX	RHS2000	18/02/2014	Séquençage en cours	---
140218_HISEQ1_C3URVACXX	RHS2000	18/02/2014	Séquençage en cours	---
140220_HISEQ2_C3FMGACXX	RHS2000	20/02/2014	Séquençage en cours	---

Runs en cours de read generation

Taille (5) 4 Résultat(s)

Code	Type	Date Run	Etat	Valide ?
111004_HISEQ6_D08J4ACXX	RHS2500	04/10/2011	Read generation en cours	---
111005_HISEQ8_D08HNACXX	RHS2500	05/10/2011	Read generation en cours	---
140214_HISEQ6_C3EYWACXX	RHS2000	14/02/2014	Read generation en cours	---
140214_HISEQ6_C3FFCACXX	RHS2000	14/02/2014	Read generation en cours	---

Runs en attente d'évaluation

1 2 3 Taille (5) 13 Résultat(s)

Code	Type	Date Run	Etat	Valide ?
140129_HISEQ10_H84N2ADXX	RHS2500R	29/01/2014	Evaluation en attente	---
140204_HISEQ11_C3C52ACXX	RHS2500	04/02/2014	Evaluation en attente	---
140204_HISEQ11_C3C1PACXX	RHS2500	04/02/2014	Evaluation en attente	---
140205_HISEQ10_C3CE4ACXX	RHS2500	05/02/2014	Evaluation en attente	---
140205_HISEQ10_C3CD4ACXX	RHS2500	05/02/2014	Evaluation en attente	---

Runs à conserver

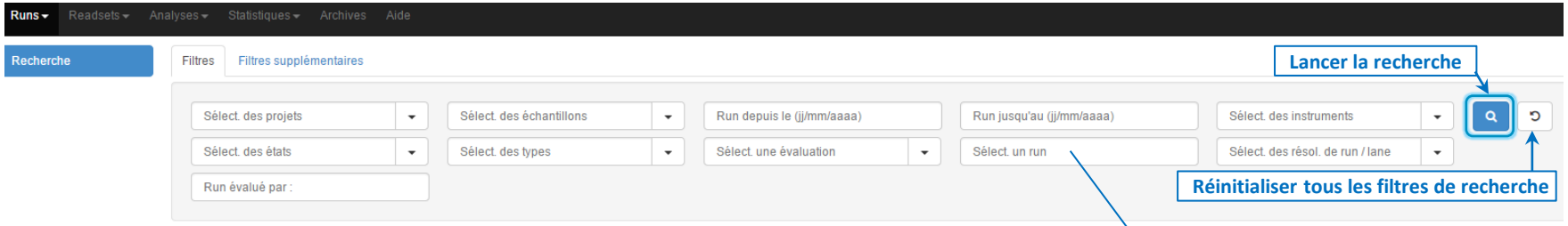
Taille (5) 0 Résultat(s)

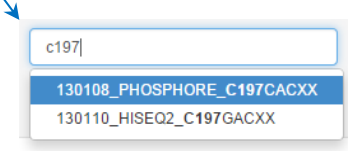
Code	Type	Date Run	Etat	Valide ?

Rechercher des runs, readsets ou analyses

RECHERCHER DES RUNS

Pour rechercher un ou plusieurs run(s), aller sur Runs > Recherche.





Plusieurs filtres de recherche sont disponibles :

- Sélectionnez des projets *(nomenclature CNS : « AQI » | nomenclature CNG : « ECTOPIA_432 ») ;*
- Sélectionnez des échantillons *(ne peut être utilisé que si un ou plusieurs projets sont sélectionnés)*;
- Date run depuis le (jj/mm/aaaa) ;
- Date run jusqu'au (jj/mm/aaaa) ;
- Sélectionnez des instruments *(Mercure, Platine, Hiseq10, …) ;*
- Sélectionnez des états pour le **run** *(séquençage en cours, RG en cours, terminé…) ;*
- Sélectionnez des types de run *(RHS2000, RHS2500R, …) ;*
- Sélectionnez une évaluation de run *(oui, non , -) ;*
- Sélectionnez un run *(en tapant des lettres qui le composent) ;*
- Sélectionnez des résolutions de run / lane *(Problème machine : fluidique, Problème réactifs : PE module, …).*
- Run évalué par *(login utilisateur)*

NB : **Les runs** qui s'affichent dans le tableau résultat sont ceux qui contiennent les résolutions recherchées : **soit dans le compte-rendu RUN soit dans le compte-rendu de leurs LANES.**

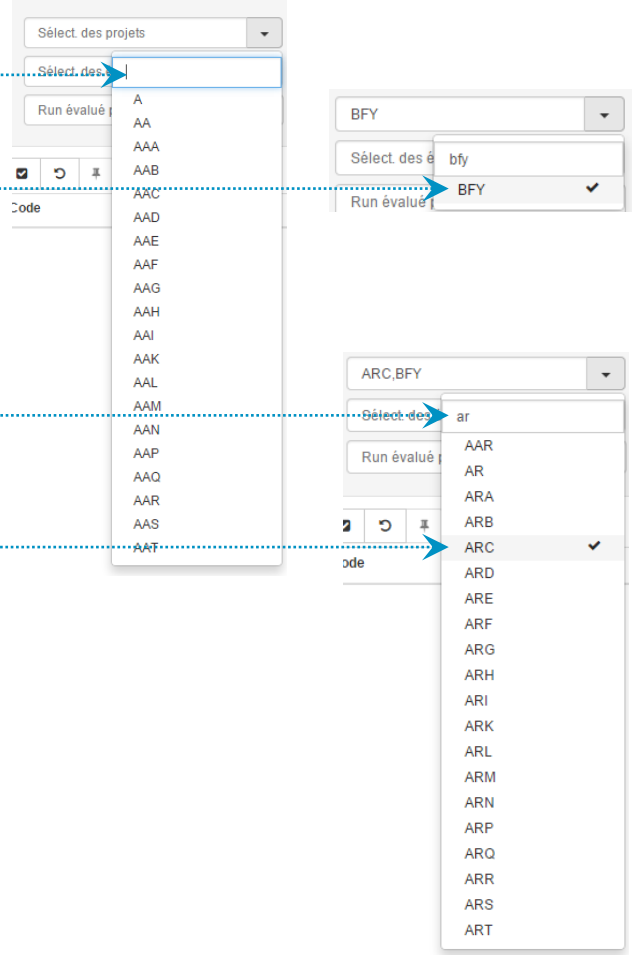
Filtres supplémentaires :

- Type de lectures *(SR / PE) ;*
- Nb cycles Read 1 *(saisie libre : nb exact de cycles R1 après transfert, ex : 101, 151, 201, 301…)*

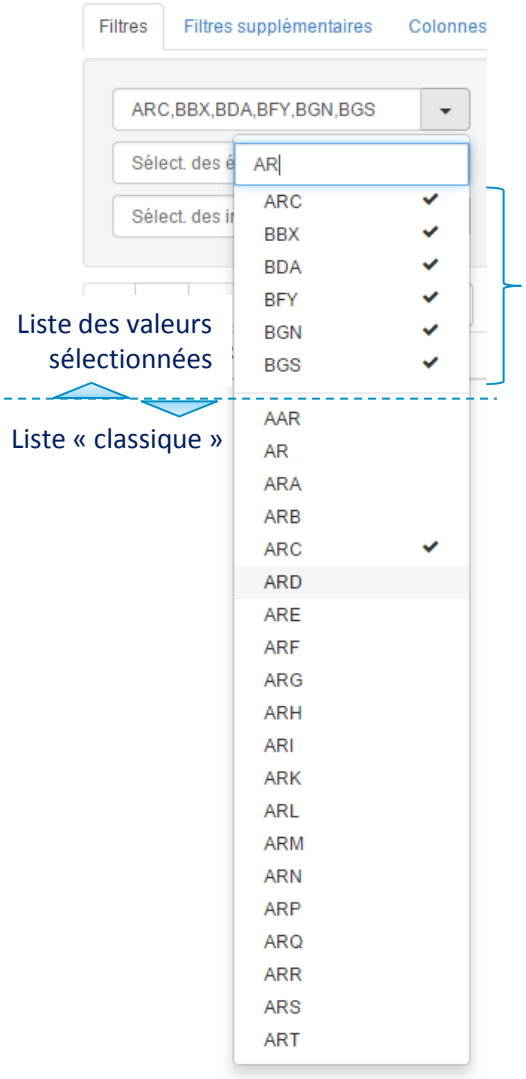
De nouveaux filtres de recherche peuvent avoir été ajoutés après la rédaction de cette documentation.

FILTRES DE RECHERCHE MULTI-PROJETS / MULTI-ECHANTILLONS

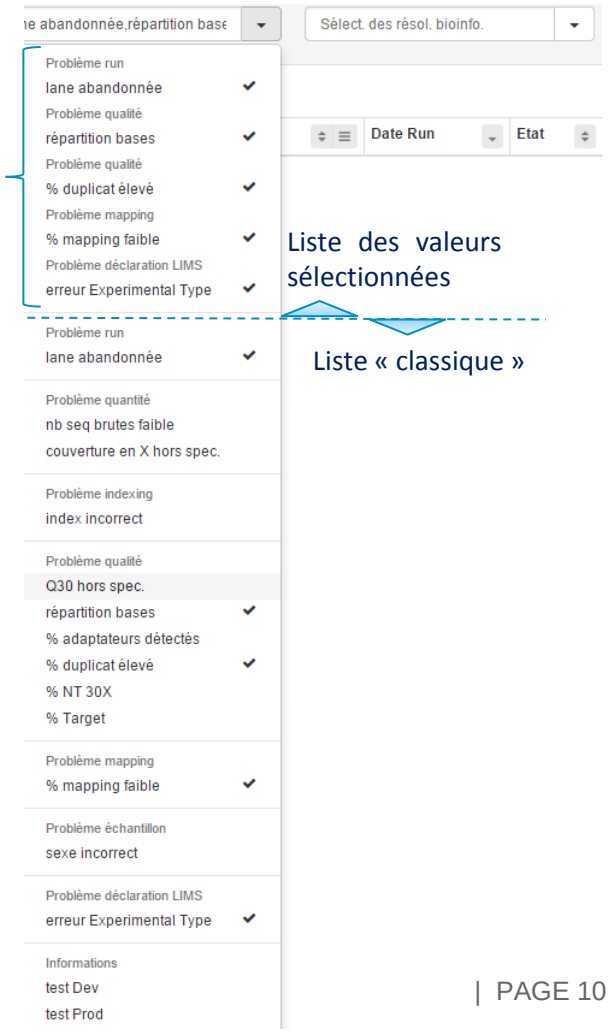
- Pour sélectionner un ou plusieurs projet(s), taper des lettres qui composent le nom / code du projet dans la cellule située en haut de la liste déroulante.
- Cliquer sur le projet souhaité, dans la liste.
- Revenir sur le bandeau du haut pour taper les lettres d'un autre projet.
- Le sélectionner dans la liste.
- Etc.
- Le fonctionnement est le même pour la recherche d'échantillons.



EVOLUTION DES FILTRES DE RECHERCHE MULTI-VALEURS DANS LA VERSION 1.6

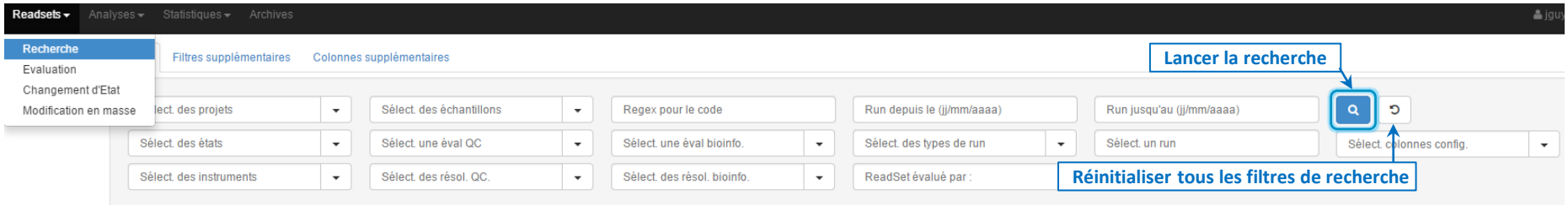


Lorsqu'un projet / échantillon / ou toute autre valeur d'une liste déroulante est sélectionnée, elle réapparaît en haut de la liste, afin de pouvoir facilement désélectionner une valeur si besoin.



RECHERCHER DES READSETS

Pour rechercher un ou plusieurs readset(s), aller sur Readsets > Recherche.



Plusieurs filtres de recherche sont disponibles :

- Sélectionnez des projets *(nomenclature CNS : « AQI » | nomenclature CNG : « ECTOPIA_432 ») ;*
- Sélectionnez des échantillons *(ne peut être utilisé que si un ou plusieurs projets sont sélectionnés)*;
- Regex pour le code *(taper un motif pour retrouver tous les readsets qui contiennent ce motif dans leur code ; exemple CNG : _RA_ permet de retrouver tous les readsets qui ont pour experimental type « RNAseq »)*
- Date run depuis le (jj/mm/aaaa) ;
- Date run jusqu'au (jj/mm/aaaa) ;
- Sélectionnez des états pour le **readset** *(RG en cours, évaluation en attente, disponible, indisponible…) ;*
- Sélectionnez une évaluation QC *(oui, non , -) ;*
- Sélectionnez une évaluation bioinfo *(oui, non , -) ;*
- Sélectionnez des types de run *(RHS2000, RHS2500R, …) ;*
- Sélectionnez un run *(en tapant des lettres qui le composent) ;*
- Sélectionnez des instruments *(Mercure, Platine, Hiseq10, …) ;*
- Sélectionnez des résolutions QC *(Problème qualité : répartition bases ; Problème taxon : conta manip…) ;*
- Sélectionnez des résolutions bioinfo.
- Readset évalué par *(login utilisateur)*

Filtres supplémentaires de recherche :

- [CNG] Envoyé CCRT ? *(oui | non) ;*
- [CNG] Envoyé Collaborateur ? *(oui | non) ;*
- [CNS] Type échantillon
- [CNS] Type Processus Banque
- [CNS] Porosité Tara (code)
- [CNS] Profondeur Tara (code)
- [CNS] Station Tara

De nouveaux filtres de recherche peuvent avoir été ajoutés après la rédaction de cette documentation.

RECHERCHER DES READSETS – REGEX SUR CODE

Quelques exemples de recherche avec la Regex pour le code readset :

Regex pour le code *(taper un motif pour retrouver tous les readsets qui contiennent ce motif dans leur code ; exemple CNG : _RA_ permet de retrouver tous les readsets qui ont pour experimental type « RNAseq »)*

Filtres | Filtres supplémentaires | Colonnes supplémentaires

Sélect. des projets | Sélect. des échantillons | OSF_

Sélect. des états | Sélect. une éval QC | Sélect. une éval bioinfo.

Sélect. des instruments | Sélect. des résol. QC. | Sélect. des résol. bioinfo.

Code	Run	N° Piste	Projet	Echantillon	Date Run
BFY_ACNHOSF_1_ABCD3.IND28	141017_MELISSE_ABCD3	1	BFY	BFY_ACNH	17/10/2014
BKB_INOSF_1_ABAV8.IND21	141017_MIMOSA_ABAV8	1	BKB	BKB_IN	17/10/2014
BKB_AKHOSF_1_ABAV8.IND22	141017_MIMOSA_ABAV8	1	BKB	BKB_AKH	17/10/2014
BKB_AIBOSF_1_ABAV8.IND19	141017_MIMOSA_ABAV8	1	BKB	BKB_AIB	17/10/2014

Tous les readsets contenant « **OSF_** » dans leur code apparaissent dans le tableau résultat.

Filtres | Filtres supplémentaires | Colonnes supplémentaires

Sélect. des projets | Sélect. des échantillons | .IND2

Code	Run	N° Piste	Projet	Echantillon	Date Run
BFY_ACLBOSF_1_ABCD3.IND2	141017_MELISSE_ABCD3	1	BFY	BFY_ACLB	17/10/2014
BKB_ADNOSF_1_ABAV8.IND24	141017_MIMOSA_ABAV8	1	BKB	BKB_ADN	17/10/2014

Tous les readsets contenant « **.IND2** » dans leur code apparaissent dans le tableau résultat. Un readset avec l'index IND24 apparaîtra également.

Filtres | Filtres supplémentaires | Colonnes supplémentaires

Sélect. des projets | Sélect. des échantillons | .IND2$ **$: se termine par le caractère précédent**

Code	Run	N° Piste	Projet	Echantillon	Date R
BGT_DOSW_1_ABAV8.IND2	141017_MIMOSA_ABAV8	1	BGT	BGT_D	17/10/2
BFY_ACLBOSF_1_ABCD3.IND2	141017_MELISSE_ABCD3	1	BFY	BFY_ACLB	17/10/2
BFY_ACCBOSF_1_AAG63.IND2	141013_MELISSE_AAG63	1	BFY	BFY_ACCB	13/10/2

Seuls les readsets dont le code **se termine par** « **.IND2** » apparaissent.

RECHERCHER DES ANALYSES

Pour certains projets, le contrôle qualité « classique » ne va pas être suffisant pour évaluer (éval. bioinfo) les readsets.
Sur ces projets, tagués au préalable comme devant comporter des analyses biologiques, vont être générées des « analyses », effectuées à partir d'un ou plusieurs readsets.

Pour rechercher un ou plusieurs analyse(s), aller sur Analyses > Recherche.

Plusieurs filtres de recherche sont disponibles :

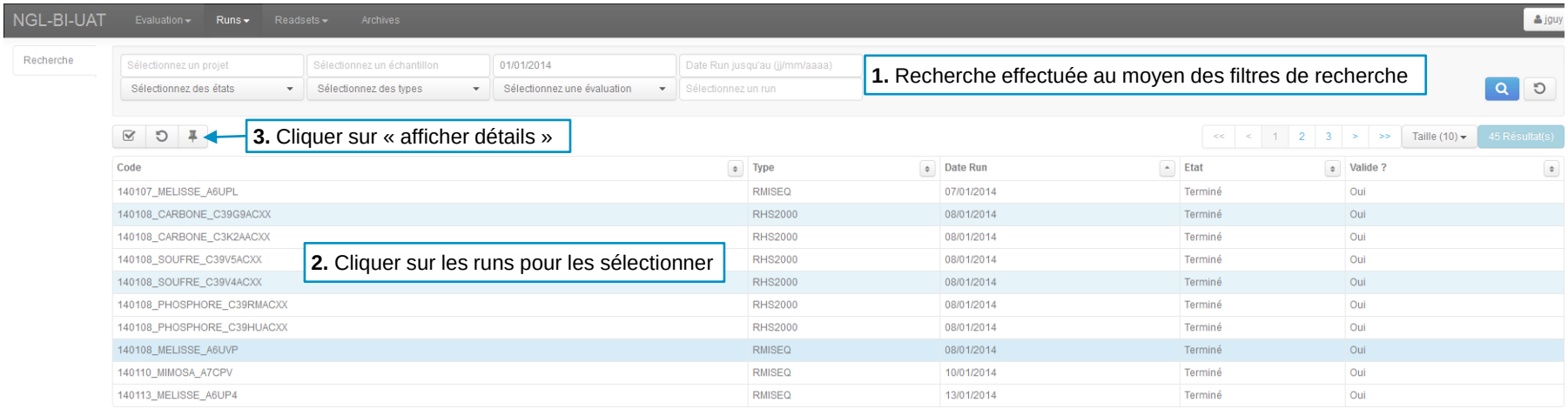
- Sélectionnez des projets *(nomenclature CNS : « AQI » | nomenclature CNG : « ECTOPIA_432 ») ;*
- Sélectionnez des échantillons *(ne peut être utilisé que si un ou plusieurs projets sont sélectionnés)* ;
- Regex pour le code *(taper un motif pour retrouver toutes les analyses qui contiennent ce motif dans leur code) ;*
- Sélectionnez des types *(type de l'analyse, exemple : BAC pool assembly) ;*
- Sélectionnez des états pour l'**analyse** *(Analyse BI en cours, Evaluation en attente, Evaluation terminée) ;*
- Sélectionnez une évaluation *(oui, non , -) ;*
- Analyse évaluée par *(login utilisateur) ;*
- Sélectionnez des résolutions *(Merging : % merging, Scaffolding : N50…)*

De nouveaux filtres de recherche peuvent avoir été ajoutés après la rédaction de cette documentation.

Visualiser les détails et les traitements d’un run

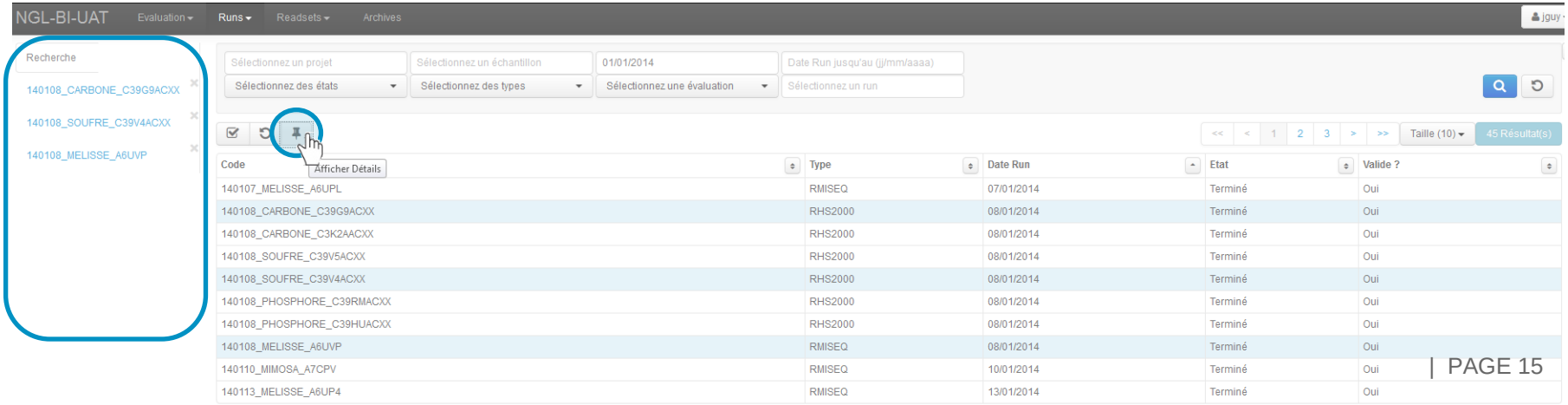
SÉLECTIONNER DES RUNS A VISUALISER

Pour visualiser les détails d'un ou de plusieurs run(s) afin de consulter les différentes statistiques relatives aux différents traitements, aller sur **Runs > Recherche**, sélectionner le(s) run(s) en cliquant sur les lignes du tableau, puis cliquer sur « afficher détails ».



Les runs à consulter s'affichent alors dans une barre à gauche.

Cette barre reste visible lorsque vous consultez un run, vous permettant ainsi de naviguer d'un run à l'autre, ou même de revenir à la recherche initiale en cliquant sur « recherche ». Vous pouvez supprimer un run de la sélection en cliquant sur la petite croix à droite du nom du run.



VISUALISER UN RUN

Cliquer sur le run à consulter.

NGL-BI | Evaluation | Runs | Readsets | Analyses | Statistiques | Archives

Recherche
- 141027_PLATINE_H072KAMXX
- 141024_MIMOSA_ABCJT
- 141023_MELISSE_AAUNG
- 141022_CARBONE_C4VM6ACXX

Sélect. des projets | Sélect. des échantillons | 01/10/2014 | Run jusqu'au (jj/mm/aaaa) | Sélect. des instruments
Sélect. des états | Sélect. des types | Sélect. une évaluation | Sélect. un run | Sélect. des résol. de run / lane
Run évalué par :

<< < 1 2 3

Code	Type	Date Run	Date fin RG	Etat	Valide ?
141027_PLATINE_H072TAMXX	RHS2500R	27/10/2014		Séquençage en cours	---
141027_PLATINE_H072KAMXX	RHS2500R	27/10/2014		Séquençage en cours	---
141027_MELISSE_AAULU	RMISEQ	27/10/2014		Séquençage en cours	---
141024_MIMOSA_ABCJT	RMISEQ	24/10/2014	26/10/2014	Evaluation terminée	Oui
141023_MELISSE_AAUNG	RMISEQ	23/10/2014	24/10/2014	Evaluation en attente	---
141022_CARBONE_C4VM6ACXX	RHS2000	22/10/2014		Séquençage en cours	---
141020_FLUOR_HAMUUADXX	RHS2500R	20/10/2014	22/10/2014	Evaluation terminée	Oui

Recherche
- 141027_PLATINE_H072KAMXX
- 141024_MIMOSA_ABCJT
- 141023_MELISSE_AAUNG
- 141022_CARBONE_C4VM6ACXX

141027_PLATINE_H072KAMXX Séquençage en cours

Code	141027_PLATINE_H072KAMXX	Type	RHS2500R	Nb Cycles		Code Instrument	PLATINE
Etat	Séquençage en cours	Nb Clusters (total)		Ligne Contrôle		Type d'Instrument	HISEQ2500
Valide ?	---	% Clusters filt. (moyenne)		Code Flowcell	H072KAMXX	Date Run	27/10/2014
Comptes Rendus		Nb Clusters filt. (total)		Position Flowcell		Date fin RG	
Critères		Nb Bases (total)		Version RTA			
Evalué par	()	A conserver ?	☐	Version Flowcell			
		Supprimé	Non				
Détails évaluation							

Pistes

#	Valide ?	Comptes Rendus
1	---	
2	---	

Readsets (0)

INTERFACE RUN – DÉTAILS

Code couleur : Run valide, Run non valide, Run à « - »

NOM du RUN → 141008_SOUFRE_C4A65ACXX Evaluation terminée ← Etat (cf. workflow)

Permet d'éditer un run *(renseigner les champs : valide ?, Comptes rendus, critères, détails évaluation, etc.)* en dehors du mode « évaluation » Editer

Résultat évaluation RUN

Code	141008_SOUFRE_C4A65ACXX	Type	RHS2000	Nb Cycles	209	Code Instrument	SOUFRE
Etat	Evaluation terminée	Nb Clusters (total)	1 520 393 260	Ligne Contrôle	4	Type d'Instrument	HISEQ2000
Valide ?	Oui	% Clusters filt. (moyenne)	88,48	Code Flowcell	C4A65ACXX	Date Run	08/10/2014
Comptes Rendus		Nb Clusters filt. (total)	1 345 304 718	Position Flowcell	A	Date fin RG	18/10/2014 00:00:00
Critères	RHS2000 2x101	Nb Bases (total)	281 168 686 062	Version RTA	1.13.48		
Evalué par	mlepretre (20/10/2014)	A conserver ?	☐	Version Flowcell	HiSeq Flow Cell v3		
		Supprimé	Non				

Visualiser workflow (cf. page suiv.)

Le choix d'un critère d'évaluation permet l'highlighting

Données ngsrg [RUN]

Highlights

Détails évaluation

Onglets traitements [LANES] → NGS-RG SAV

Statistiques associées au traitement [LANES]

#	Nb Cycles	Nb Clusters	% Clusters filtrés (illumina)	Nb Cluster filtrés (illumina)	% Séquences filtrées (interne)	Nb Séquences (filtre illumina + interne)	Nb Bases (filtre illumina + interne)	% Perte
1	101,101	185 062 377	83,49	154 511 690	99,18	153 238 982	30 954 274 364	0,87
2	101,101	162 709 695	86,11	140 115 620	98,99	138 693 875	28 016 162 750	1,42
3	101,101	171 680 243	85,13	146 149 689	99,09	144 819 792	29 253 597 984	0,78
4	101,101	209 817 789	90,25	189 367 172	99,46	188 347 830	38 046 261 660	1,62
5	101,101	198 935 026	80,64	160 427 921	99,31	159 313 261	32 181 278 722	1,08
6	101,101	206 696 480	92,92	192 058 252	99,44	190 975 094	38 576 968 988	0,71
7	101,101	189 724 327	94,2	178 721 208	99,36	177 575 901	35 870 332 002	0,74
8	101,101	195 767 323	93,97	183 953 166	99,35	182 754 663	36 916 441 926	0,7

Pistes

Résultat évaluation LANES

#	Valide ?	Comptes Rendus
1	Oui	
2	Oui	
3	Oui	
4	Oui	
5	Oui	
6	Oui	
7	Oui	
8	Oui	

Readsets (26)

Sélect. des lanes

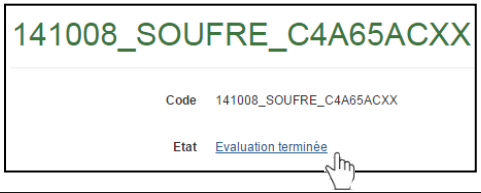
Voir Readsets

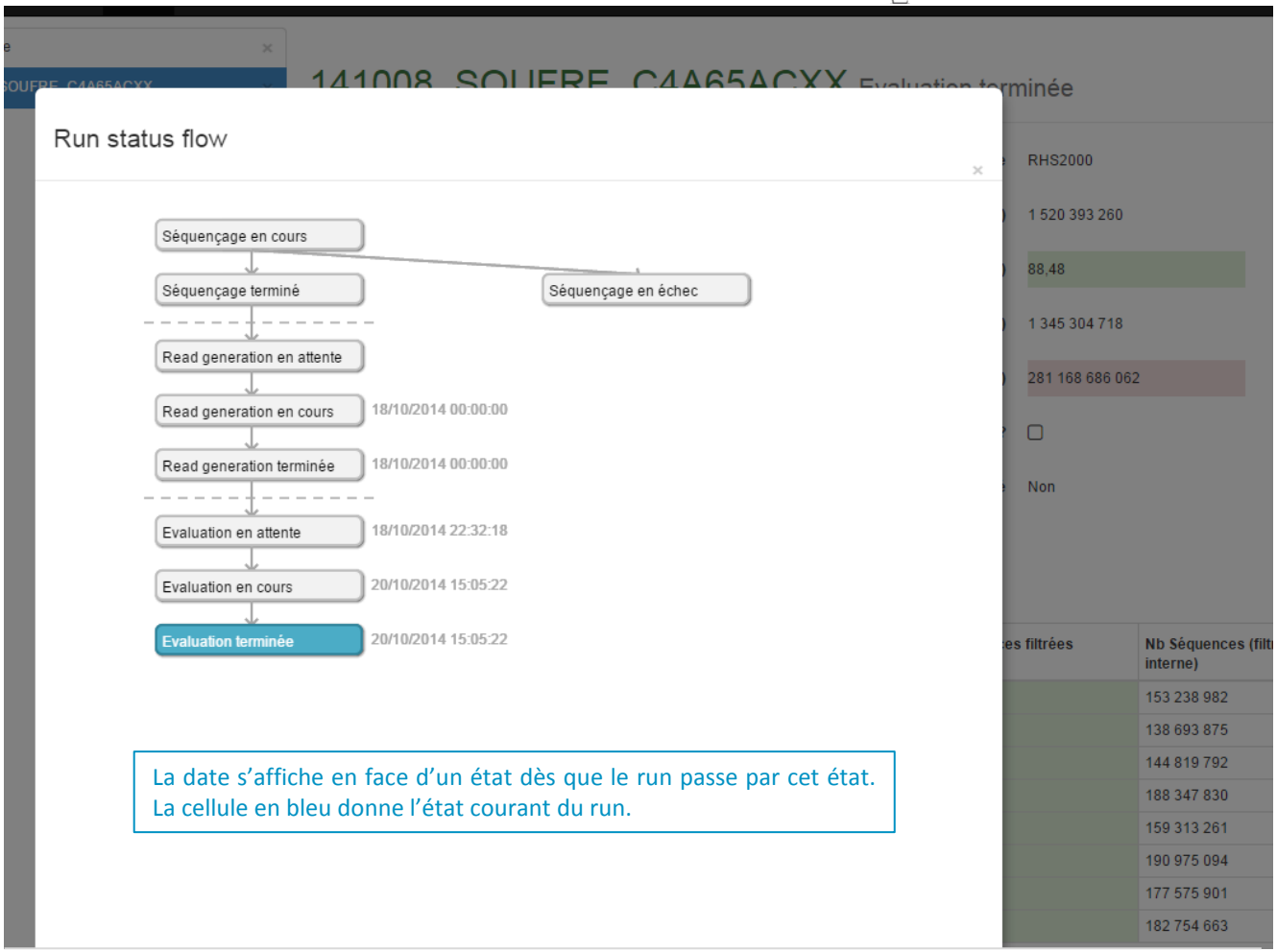
Récapitulatif des lots de séquences sur la FC et métriques importantes

N° Piste	Code	Etat	% déposé	% Séquences valides / piste	Nb Séquences valides	Nb Bases	% >= Q30	Score Qualité moyen	Valide QC ?	Valide BioInfo ?
1	BEZ_ABMAOSW_1_C4A65ACXX.IND4	Disponible	100,00	99,13	151 904 236	30 684 655 672	90,41	35,08	Oui	Oui
2	BEZ_ABQAOSW_2_C4A65ACXX.IND2	Disponible	100,00	98,58	136 722 846	27 618 014 892	91,69	35,42	Oui	Oui
3	BEZ_ABSAOSW_3_C4A65ACXX.IND1	Disponible	100,00	99,22	143 692 620	29 025 909 240	90,95	35,24	Oui	Oui
4	BCM_CIDOSW_4_C4A65ACXX.IND8	Disponible	10,00	10,50	19 775 254	3 994 601 308	92,20	35,98	Oui	Oui
4	BCM_CLLOSW_4_C4A65ACXX.IND5	Disponible	10,00	7,56	14 245 439	2 877 578 678	91,00	35,62	Oui	Oui
4	BCM_CLROSW_4_C4A65ACXX.IND10	Disponible	10,00	8,49	15 987 572	3 229 489 544	91,64	35,80	Oui	Oui
4	BCM_CLMOSW_4_C4A65ACXX.IND6	Disponible	10,00	11,19	21 084 325	4 259 033 650	89,75	35,25	Oui	Oui

VISUALISER UN RUN DANS SON WORKFLOW

Cliquer sur l'état du run pour le visualiser dans son workflow.





La date s'affiche en face d'un état dès que le run passe par cet état.
La cellule en bleu donne l'état courant du run.

RACCOURCIS RUNS – READSETS

Après avoir sélectionné un run et affiché ses détails, se rendre en bas de la page pour visualiser le tableau récapitulatif des readsets qui le composent. A partir de ce tableau, vous pouvez :

- Choisir de **n'afficher que certaines lanes** ;
- Cliquer sur « **Voir readsets** » pour afficher dans une nouvelle page le tableau des readsets tel qu'il apparaît à l'écran et consulter leurs détails ;
- Cliquer sur « **Evaluer readsets** » (bouton disponible quand le run est en mode édition) pour afficher dans une nouvelle page le tableau des readsets tel qu'il apparaît à l'écran et les valider / invalider en masse.

NGS-RG

#	Nb Cycles	Nb Clusters	% Clusters filtrés (illumina)	Nb Cluster filtrés (illumina)	% Séquences filtrées (interne)	Nb Séquences (filtre illumina + interne)	Nb Bases (filtre illumina + interne)	% Perte
1	101,101	189 116 299	93,4	176 641 897	99,26	175 327 073	35 416 068 746	1,9
2	101,101	153 970 984	93,99	144 714 120	98,99	143 245 849	28 935 661 498	7
3	101,101	147 022 449	94,36	138 725 497	96,36	133 673 081	27 001 962 362	7,07
4	101,101	143 428 210	94,28	135 228 393	98,63	133 376 759	26 942 105 318	6,94
5	101,101	161 139 741	93,65	150 911 677	98,8	149 102 517	30 118 708 434	6,83

Pistes

#	Valide ?	Comptes Rendus
1	---	
2	---	
3	---	
4	---	
5	---	

Readsets (10)

Sélect. des lanes

Voir Readsets | Evaluer Readsets

N° Piste	Code	Etat	% déposé	% Séquences valides / piste	Nb Séquences valides	Nb Bases	% >= Q30	Score Qualité moyen	Valide QC ?	Valide BioInfo ?
1	E421_CB_B00EUIT_1_D1DAGACXX.IND4	Disponible		38,05	66 712 820	13 475 989 640	92,81	36,20	---	---
1	E421_CB_B00EPPR_1_D1DAGACXX.IND2	Disponible		37,72	66 141 643	13 360 611 886	92,70	36,15	---	---

A partir du tableau récapitulatif des readsets, visible sur l'interface du run, vous pouvez cliquer sur un readset pour consulter ses détails. Une nouvelle page s'ouvre directement sur l'interface du readset.

ReadSets (10)

Permet de filtrer sur les lanes qu'on veut visualiser.

1,6

1 ✔
2
3
4
5
6 ✔
7
8

EUIT_1_D1DAGACXX.IND4
EPPR_1_D1DAGACXX.IND2
EUJ1_1_D1DAGACXX.IND5
P3N_2_D1DAGACXX.IND5
PLS_3_D1DAGACXX.IND5
E410_PA_B00FPM3_4_D1DAGACXX.IND5

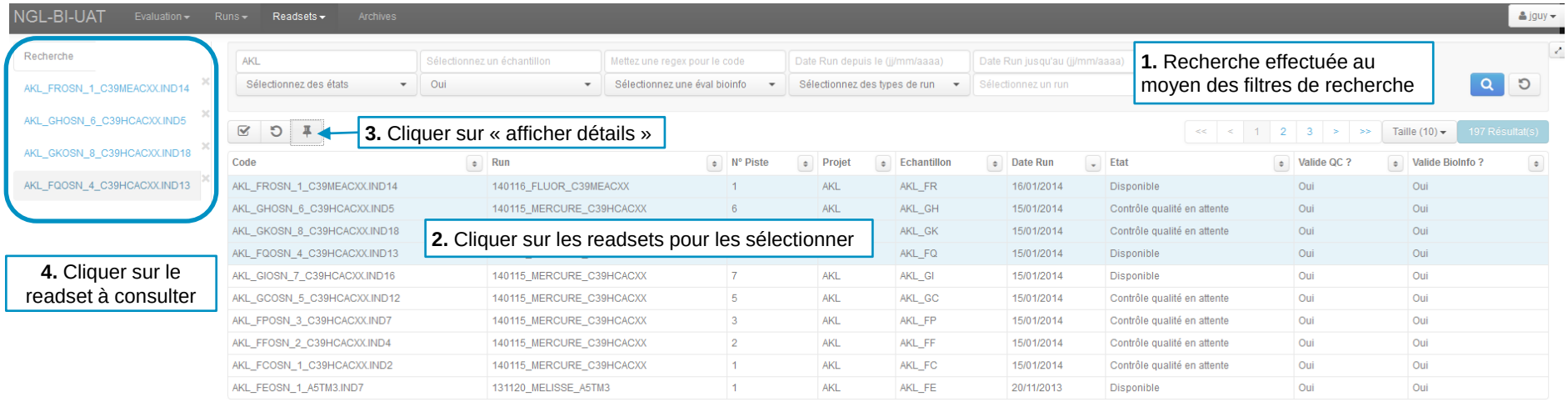
Permet d'ouvrir une page de recherche de readset donnant directement sur la sélection visualisée dans ce tableau.

Permet d'ouvrir une page d'évaluation en masse de readset donnant directement sur la sélection visualisée dans ce tableau.

Visualiser les détails et les traitements d’un readset

SÉLECTIONNER DES READSETS A VISUALISER

Pour visualiser les détails d'un ou de plusieurs readset(s) afin de consulter les différentes statistiques relatives aux différents traitements, aller sur **Readsets > Recherche**, sélectionner le(s) readset(s) en cliquant sur les lignes du tableau, puis cliquer sur « afficher détails ».



INTERFACE READSET – DÉTAILS (1/14)

Code couleur : valide, non valide, « - »

NOM du lot de séquence → **BAT_AAKOSW_2_C39G9ACXX.IND2** Disponible ← Etat (cf. workflow)

Permet d'éditer un readset *(renseigner les champs : valide ?, Comptes rendus, critères, détails évaluation, etc.)* en dehors du mode « évaluation »

Editer

Général | Avancé

Code	BAT_AAKOSW_2_C39G9ACXX.IND2	Nb Séquences utiles	200,311,306	Run / N° Piste	140108_CARBONE_C39G9ACXX / 2
Etat	Disponible	Nb Bases utiles	39,920,324,496	Type de Run	RHS2000
Valide QC ?	Non	Valide BioInfo ?	Oui	Nb Cycles	209
Comptes Rendus QC	Problème qualité Q30 duplicat > 30	Comptes Rendus BioInfo			
Critères QC	Default	Critères BioInfo	Low		
Evalué par	ngsrg (10/03/2014)	Evalué par	ngsrg (10/03/2014)		
Détails évaluation	commentaires suppl. sur l'évalution				

Résultat évaluation lot de séquence

Notion de validité bioinfo : un lot de séquence peut être invalidé par l'équipe QC mais utilisable pour le bio informaticien.

Stat. globales

Raccourci pour aller sur le RUN

Onglets traitements [READSET] → NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | Trimming (vs RAW) | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) | Read Quality (vs CLEAN) | Duplicates (vs CLEAN) | Mapping (vs CLEAN) | Merging (vs CLEAN)

% déposé	% Séquences valides / piste	Nb Séquences valides	Nb Bases	% >= Q30	Score Qualité moyen
	96.35	202,710,191	40,947,458,582	92.91	36.29

Statistiques associées au traitement [READSET]

INTERFACE READSET – DÉTAILS (2/14)

Code couleur : valide, non valide, « - »

NOM du lot de séquence → **BAT_AAKOSW_2_C39G9ACXX.IND2** Disponible ← Etat (cf. workflow)

Editer

Général | Avancé

SSID	1205772468
Date de l'archive	21/01/2014 00:00:00
Chemin	/env/cns/proj/projet_BAT/AAK/RunsSolexa/140108_CARBONE_C39G9ACXX/
Envoyé au CCRT ?	☐
Envoyé au Collaborateur ?	☐

← **Informations pouvant être renseignées (individuellement ou en masse)**

Nom du fichier	Type de fichier	Utilisable	Label	Encodage ASCII	Clé codage md5
BAT_AAKOSW_2_1_C39G9ACXX.IND2.fastq	RAW	Non	READ1	33	
BAT_AAKOSW_2_2_C39G9ACXX.IND2.fastq	RAW	Non	READ2	33	
BAT_AAKOSW_2_2_C39G9ACXX.IND2_clean.fastq.gz	CLEAN	Oui	READ2	33	32f6ea384723278ea73c31ab07038005
BAT_AAKOSW_2_1_C39G9ACXX.IND2_clean.fastq.gz	CLEAN	Oui	READ1	33	f1708c84ec898f637346fe42cb4a1ae3
BAT_AAKOSW_2_C39G9ACXX.IND2_single_clean.fastq.gz	CLEAN	Non	SINGLETON	33	925023c398ded4ceabbe63a9cd976e08

Informations sur les fichiers

NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | Trimming (vs RAW) | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) | Read Quality (vs CLEAN) | Duplicates (vs CLEAN) | Mapping (vs CLEAN) | Merging (vs CLEAN)

% déposé	% Séquences valides / piste	Nb Séquences valides	Nb Bases	% >= Q30	Score Qualité moyen
	96.35	202,710,191	40,947,458,582	92.91	36.29

INTERFACE READSET – DÉTAILS (3/14)

Code couleur : valide, non valide, « - »

NOM du lot de séquence → **BAT_AAKOSW_2_C39G9ACXX.IND2** Disponible ← Etat (cf. workflow)

Editer

Général | Avancé

Code	BAT_AAKOSW_2_C39G9ACXX.IND2	**Nb Séquences utiles**	200,311,306	**Run / N° Piste**	140108_CARBONE_C39G9ACXX / 2
Etat	Disponible	**Nb Bases utiles**	39,920,324,496	**Type de Run**	
Valide QC ?	Non	**Valide BioInfo ?**	Oui	**Nb Cycles**	
Comptes Rendus QC	**Problème qualité** Q30 duplicat > 30	**Comptes Rendus BioInfo**			
Critères QC	Default	**Critères BioInfo**	Low		
Evalué par	ngsrg (10/03/2014)	**Evalué par**	ngsrg (10/03/2014)		
Détails évaluation	commentaires suppl. sur l'évalution				

Onglets traitements [READSET] →

NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | Trimming (vs RAW) | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) | Read Quality (vs CLEAN) | Duplicates (vs CLEAN) | Mapping (vs CLEAN) | Merging (vs CLEAN)

Statistiques associées au traitement [READSET]

% déposé	% Séquences valides / piste	Nb Séquences valides	Nb Bases	% >= Q30	Score Qualité moyen
	96.35	202,710,191	40,947,458,582	92.91	36.29

INTERFACE READSET – DÉTAILS (4/14)

Code couleur : valide, non valide, « - »

NOM du lot de séquence → **BAT_AAKOSW_2_C39G9ACXX.IND2** Disponible ← Etat (cf. workflow)

Editer

Général | Avancé

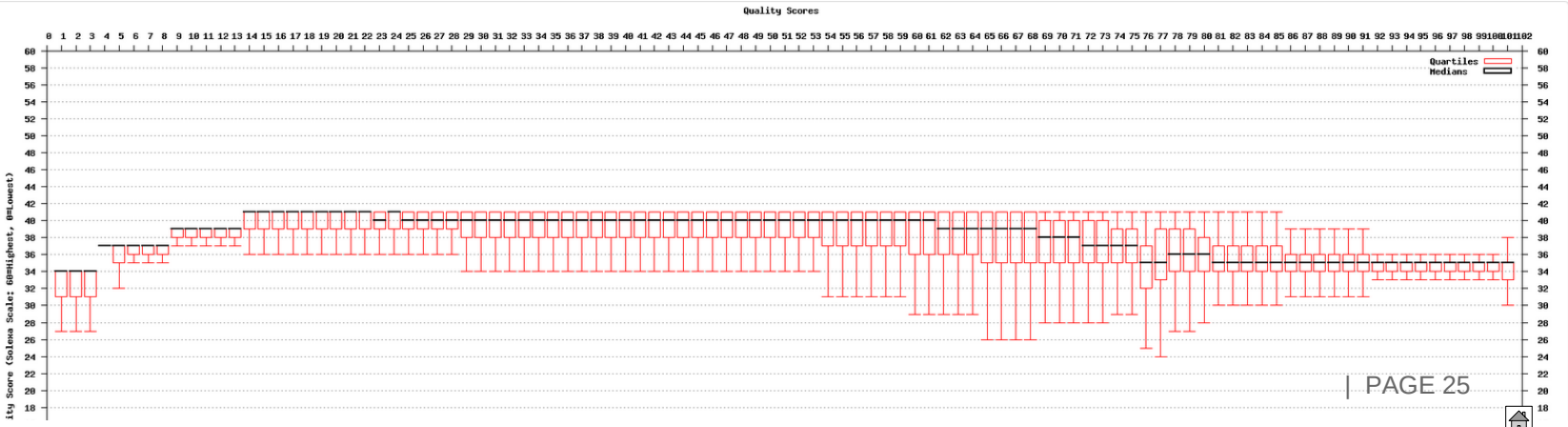
Code	BAT_AAKOSW_2_C39G9ACXX.IND2	Nb Séquences utiles	200,311,306	Run / N° Piste	140108_CARBONE_C39G9ACXX / 2
Etat	Disponible	Nb Bases utiles	39,920,324,496	Type de Run	RHS2000
Valide QC ?	Non	Valide BioInfo ?	Oui	Nb Cycles	209
Comptes Rendus QC	Problème qualité Q30 duplicat > 30	Comptes Rendus BioInfo			
Critères QC	Default	Critères BioInfo	Low		
Evalué par	ngsrg (10/03/2014)	Evalué par	ngsrg (10/03/2014)		
Détails évaluation	commentaires suppl. sur l'évalution				

Onglets traitements [READSET] →

NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | Trimming (vs RAW) | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) | Read Quality (vs CLEAN) | Duplicates (vs CLEAN) | Mapping (vs CLEAN) | Merging (vs CLEAN)

Statistiques associées au traitement [READSET]

Read 1 | Read 2

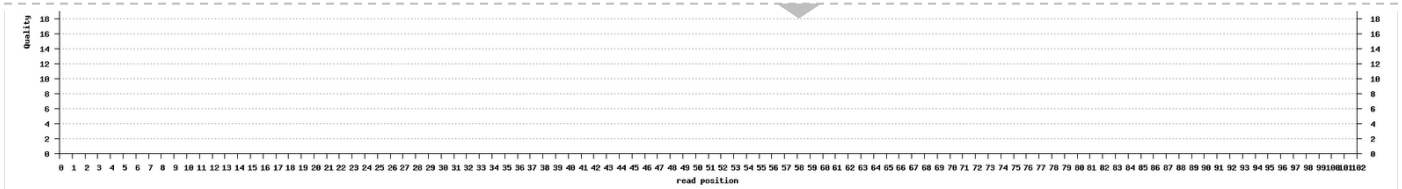


INTERFACE READSET – DÉTAILS (5/14)

Onglets traitements [READSET] →

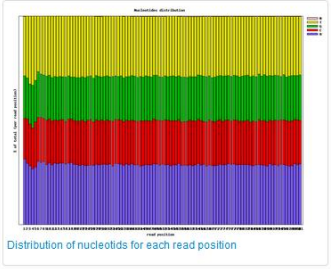
NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | Trimming (vs RAW) | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) | Read Quality (vs CLEAN) | Duplicates (vs CLEAN) | Mapping (vs CLEAN) | Merging (vs CLEAN)

Read 1 | Read 2

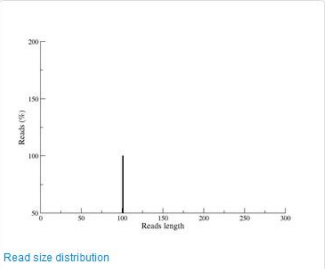


Distribution of nucleotids for each read position

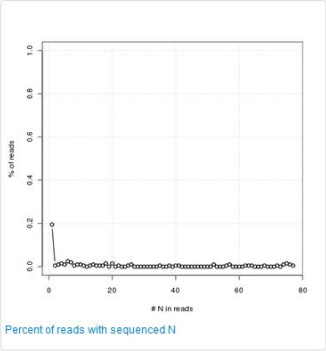
Statistiques associées au traitement [READSET]



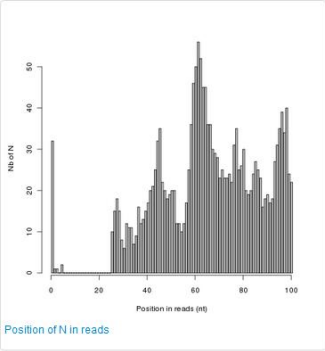
Distribution of nucleotids for each read position



Read size distribution

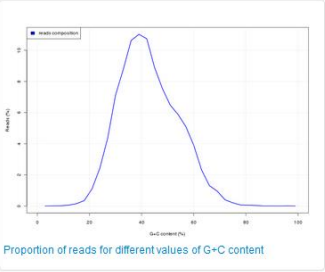


Percent of reads with sequenced N



Position of N in reads

Adapters vs run cycles



Proportion of reads for different values of G+C content

INTERFACE READSET – DÉTAILS (6/14)

Code couleur : valide, non valide, « - »

NOM du lot de séquence → **BAT_AAKOSW_2_C39G9ACXX.IND2** Disponible ← Etat (cf. workflow)

Editer

Général | Avancé

Code	BAT_AAKOSW_2_C39G9ACXX.IND2	**Nb Séquences utiles**	200,311,306	**Run / N° Piste**	140108_CARBONE_C39G9ACXX / 2
Etat	Disponible	**Nb Bases utiles**	39,920,324,496	**Type de Run**	RHS2000
Valide QC ?	Non	**Valide BioInfo ?**	Oui	**Nb Cycles**	209
Comptes Rendus QC	**Problème qualité** Q30 duplicat > 30	**Comptes Rendus BioInfo**			
Critères QC	Default	**Critères BioInfo**	Low		
Evalué par	ngsrg (10/03/2014)	**Evalué par**	ngsrg (10/03/2014)		
Détails évaluation	commentaires suppl. sur l'évalution				

Onglets traitements [READSET] →

NGS-RG | Read Quality (vs RAW) | **Duplicates (vs RAW)** | Trimming (vs RAW) | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) | Read Quality (vs CLEAN) | Duplicates (vs CLEAN) | Mapping (vs CLEAN) | Merging (vs CLEAN)

Statistiques associées au traitement [READSET]

	Read 1	Read 2	Paired
Estimation de lectures dupliquées	24,629,289 (12.15 %) 82.21 % observé 2 fois 8.27 % observé 3 fois 1.46 % observé 4 fois 0.75 % observé 5 fois 0.40 % observé 6 fois 0.31 % observé 7 fois 0.28 % observé 8 fois 0.25 % observé 9 fois 0.20 % observé 10 fois 5.87 % observé 11 fois ou plus	23,220,453 (11.45 %) 83.15 % observé 2 fois 7.85 % observé 3 fois 1.52 % observé 4 fois 0.73 % observé 5 fois 0.40 % observé 6 fois 0.37 % observé 7 fois 0.32 % observé 8 fois 0.22 % observé 9 fois 0.16 % observé 10 fois 5.28 % observé 11 fois ou plus	5,331,279 (2.63 %) 94.98 % observé 2 fois 4.47 % observé 3 fois 0.28 % observé 4 fois 0.08 % observé 5 fois 0.04 % observé 6 fois 0.04 % observé 7 fois 0.03 % observé 8 fois 0.01 % observé 10 fois 0.09 % observé 11 fois
Estimation de lectures uniques	178,080,902 (87.85 %)	179,489,738 (88.54 %)	197,378,912 (97.37 %)

INTERFACE READSET – DÉTAILS (7/14)

Code couleur : valide, non valide, « - »

NOM du lot de séquence → **BAT_AAKOSW_2_C39G9ACXX.IND2** Disponible ← Etat (cf. workflow)

Editer

Général | Avancé

Code	BAT_AAKOSW_2_C39G9ACXX.IND2	**Nb Séquences utiles**	200,311,306	**Run / N° Piste**	140108_CARBONE_C39G9ACXX / 2
Etat	Disponible	**Nb Bases utiles**	39,920,324,496	**Type de Run**	RHS2000
Valide QC ?	Non	**Valide BioInfo ?**	Oui	**Nb Cycles**	209
Comptes Rendus QC	**Problème qualité** Q30 duplicat > 30	**Comptes Rendus BioInfo**			
Critères QC	Default	**Critères BioInfo**	Low		
Evalué par	ngsrg (10/03/2014)	**Evalué par**	ngsrg (10/03/2014)		
Détails évaluation	commentaires suppl. sur l'évalution				

Onglets traitements [READSET] →

NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | **Trimming (vs RAW)** | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) | Read Quality (vs CLEAN) | Duplicates (vs CLEAN) | Mapping (vs CLEAN) | Merging (vs CLEAN)

Statistiques associées au traitement [READSET]

	Size range	Reads input	Reads output	Reads trim.	% Reads trim.	Nucleotide trim.	Rejected (short)	Rejected (length 0)
Read 1	30-101	202,710,191	202,341,180	15,250,138	7.52 %	252,655,799	366,680	2,331
Read 2	30-101	202,710,191	200,683,116	20,299,742	10.01 %	520,284,778	1,451,851	575,224

Stored Pairs	Stored Singleton	Rejected Pairs
200,335,982	2,352,332	21,877

INTERFACE READSET – DÉTAILS (8/14)

Code couleur : valide, non valide, « - »

NOM du lot de séquence → **BAT_AAKOSW_2_C39G9ACXX.IND2** Disponible ← Etat (cf. workflow)

Editer

Général | Avancé

Code	BAT_AAKOSW_2_C39G9ACXX.IND2	**Nb Séquences utiles**	200,311,306	**Run / N° Piste**	140108_CARBONE_C39G9ACXX / 2
Etat	Disponible	**Nb Bases utiles**	39,920,324,496	**Type de Run**	RHS2000
Valide QC ?	Non	**Valide BioInfo ?**	Oui	**Nb Cycles**	209
Comptes Rendus QC	**Problème qualité** Q30 duplicat > 30	**Comptes Rendus BioInfo**			
Critères QC	Default	**Critères BioInfo**	Low		
Evalué par	ngsrg (10/03/2014)	**Evalué par**	ngsrg (10/03/2014)		
Détails évaluation	commentaires suppl. sur l'évalution				

Onglets traitements [READSET] →

NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | Trimming (vs RAW) | **Contam. PhiX (vs TRIM)** | Taxonomy (vs CLEAN) | Read Quality (vs CLEAN) | Duplicates (vs CLEAN) | Mapping (vs CLEAN) | Merging (vs CLEAN)

Statistiques associées au traitement [READSET]

	Reads input	Removed reads	Remaining reads	% removed reads
Read 1 + Read 2	200,335,982	24,676	200,311,306	0.01 %
Single	2,352,332	2,242	2,350,090	0.10 %

INTERFACE READSET – DÉTAILS (9/14)

Code couleur : valide, non valide, « - »

NOM du lot de séquence → **BAT_AAKOSW_2_C39G9ACXX.IND2** Disponible ← Etat (cf. workflow)

Editer

Général | Avancé

Code	BAT_AAKOSW_2_C39G9ACXX.IND2	Nb Séquences utiles	200,311,306	Run / N° Piste	140108_CARBONE_C39G9ACXX / 2
Etat	Disponible	Nb Bases utiles	39,920,324,496	Type de Run	RHS2000
Valide QC ?	Non	Valide BioInfo ?	Oui	Nb Cycles	209
Comptes Rendus QC	Problème qualité Q30 duplicat > 30	Comptes Rendus BioInfo			
Critères QC	Default	Critères BioInfo	Low		
Evalué par	ngsrg (10/03/2014)	Evalué par	ngsrg (10/03/2014)		
Détails évaluation	commentaires suppl. sur l'évalution				

Onglets traitements [READSET] →

NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | Trimming (vs RAW) | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) * | Read Quality (vs CLEAN) | Duplicates (vs CLEAN) | Mapping (vs CLEAN) | Merging (vs CLEAN)

Statistiques associées au traitement [READSET]

Bilans | Krona | Arbre phylogenetique

Nb Seq. (échantillonnage)	Organisme	Taxonomie
20,000	Capra hircus	cellular organisms; Eukaryota; Fungi/Metazoa group; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Laurasiath

Bilan par taxon

Taxon	Nb Seq.	%
Unknown::No hits	12,162	60.81 %
Bos taurus	2,014	10.07 %
Bovidae	1,755	8.77 %
Pecora	1,041	5.21 %
Unknown::Not assigned	539	2.70 %
Homo sapiens	377	1.88 %
cellular organisms	327	1.64 %
Ovis aries	245	1.23 %
Capra hircus	224	1.12 %
Bovinae	196	0.98 %
Budorcas taxicolor	115	0.57 %
Eutheria	110	0.55 %
Cetartiodactyla	105	0.53 %
Sus scrofa	93	0.46 %
Caprinae	84	0.42 %

Bilan par division

Division	Nb Seq.	%
Eukaryota	6,916	34.58 %
cellular organisms	327	1.64 %
Bacteria	35	0.18 %
Viruses	5	0.03 %
Archaea	0	0.00 %

Bilan par mot-clé

Mot-clé	Nb Seq.	%
BAC	2,362	11.81 %
mitochondri	16	0.08 %
virus	13	0.07 %
transposase	2	0.01 %
chloroplast	0	0.00 %

+ fungi (V 1.9)

** pas encore disponible pour le CNG*

INTERFACE READSET – DÉTAILS (10/14)

Code couleur : valide, non valide, « - »

NOM du lot de séquence → **BAT_AAKOSW_2_C39G9ACXX.IND2** Disponible ← Etat (cf. workflow)

Editer

Général | Avancé

Code	BAT_AAKOSW_2_C39G9ACXX.IND2	Nb Séquences utiles	200,311,306	Run / N° Piste	140108_CARBONE_C39G9ACXX / 2
Etat	Disponible	Nb Bases utiles	39,920,324,496	Type de Run	RHS2000
Valide QC ?	Non	Valide BioInfo ?	Oui	Nb Cycles	209
Comptes Rendus QC	Problème qualité Q30 duplicat > 30	Comptes Rendus BioInfo			
Critères QC	Default	Critères BioInfo	Low		
Evalué par	ngsrg (10/03/2014)	Evalué par	ngsrg (10/03/2014)		
Détails évaluation	commentaires suppl. sur l'évalution				

Onglets traitements [READSET] →

NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | Trimming (vs RAW) | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) * | Read Quality (vs CLEAN) | Duplicates (vs CLEAN) | Mapping (vs CLEAN) | Merging (vs CLEAN)

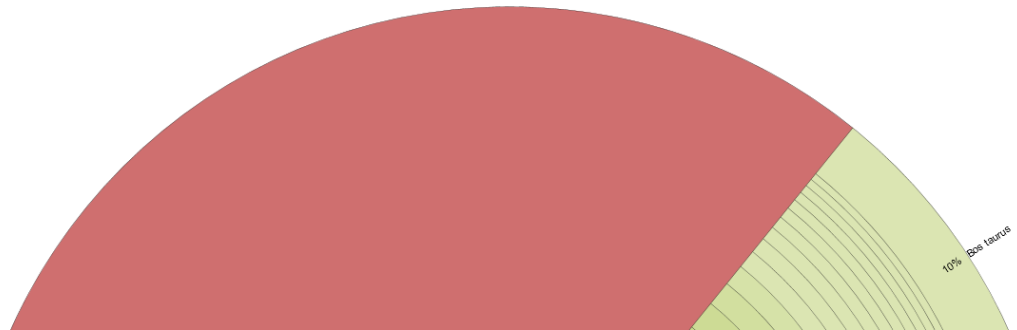
Bilans | Krona | Arbre phylogénétique

Statistiques associées au traitement [READSET]

Krona ← → Search: x

- 32 + Max depth
- 11 + Font size
- + Chart size
- ☑ Collapse
- Snapshot
- Link
- ?

all ↔
Total: 20000



* pas encore disponible pour le CNG

INTERFACE READSET – DÉTAILS (11/14)

Code couleur : valide, non valide, « - »

NOM du lot de séquence → **BAT_AAKOSW_2_C39G9ACXX.IND2** Disponible ← Etat (cf. workflow)

Editer

Général | Avancé

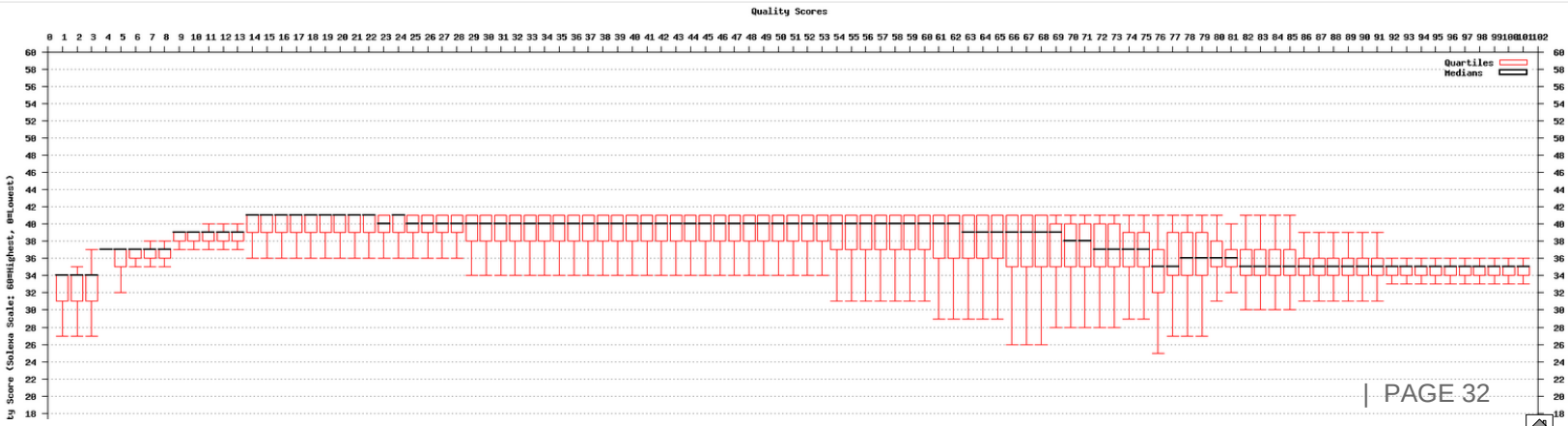
Code	BAT_AAKOSW_2_C39G9ACXX.IND2	Nb Séquences utiles	200,311,306	Run / N° Piste	140108_CARBONE_C39G9ACXX / 2
Etat	Disponible	Nb Bases utiles	39,920,324,496	Type de Run	RHS2000
Valide QC ?	Non	Valide BioInfo ?	Oui	Nb Cycles	209
Comptes Rendus QC	Problème qualité Q30 duplicat > 30	Comptes Rendus BioInfo			
Critères QC	Default	Critères BioInfo	Low		
Evalué par	ngsrg (10/03/2014)	Evalué par	ngsrg (10/03/2014)		
Détails évaluation	commentaires suppl. sur l'évalution				

Onglets traitements [READSET] →

NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | Trimming (vs RAW) | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) | **Read Quality (vs CLEAN)** | Duplicates (vs CLEAN) | Mapping (vs CLEAN) | Merging (vs CLEAN)

Statistiques associées au traitement [READSET]

Read 1 | Read 2



INTERFACE READSET – DÉTAILS (12/14)

Code couleur : valide, non valide, « - »

NOM du lot de séquence → **BAT_AAKOSW_2_C39G9ACXX.IND2** Disponible ← Etat (cf. workflow)

Editer

Général | Avancé

Code	BAT_AAKOSW_2_C39G9ACXX.IND2	Nb Séquences utiles	200,311,306	Run / N° Piste	140108_CARBONE_C39G9ACXX / 2
Etat	Disponible	Nb Bases utiles	39,920,324,496	Type de Run	RHS2000
Valide QC ?	Non	Valide BioInfo ?	Oui	Nb Cycles	209
Comptes Rendus QC	Problème qualité Q30 duplicat > 30	Comptes Rendus BioInfo			
Critères QC	Default	Critères BioInfo	Low		
Evalué par	ngsrg (10/03/2014)	Evalué par	ngsrg (10/03/2014)		
Détails évaluation	commentaires suppl. sur l'évalution				

Onglets traitements [READSET] →

NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | Trimming (vs RAW) | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) | Read Quality (vs CLEAN) | **Duplicates (vs CLEAN)** | Mapping (vs CLEAN) | Merging (vs CLEAN)

Statistiques associées au traitement [READSET]

	Read 1	Read 2	Paired
Estimation de lectures dupliquées	22,675,240 (11.32 %) 83.42 % observé 2 fois 7.78 % observé 3 fois 1.27 % observé 4 fois 0.63 % observé 5 fois 0.40 % observé 6 fois 0.28 % observé 7 fois 0.29 % observé 8 fois 0.26 % observé 9 fois 0.25 % observé 10 fois 5.41 % observé 11 fois ou plus	21,453,341 (10.71 %) 84.28 % observé 2 fois 7.24 % observé 3 fois 1.52 % observé 4 fois 0.56 % observé 5 fois 0.44 % observé 6 fois 0.38 % observé 7 fois 0.28 % observé 8 fois 0.21 % observé 9 fois 0.14 % observé 10 fois 4.95 % observé 11 fois ou plus	4,507,005 (2.25 %) 95.64 % observé 2 fois 4.03 % observé 3 fois 0.17 % observé 4 fois 0.05 % observé 5 fois 0.03 % observé 6 fois 0.01 % observé 7 fois 0.01 % observé 8 fois 0.01 % observé 9 fois 0.05 % observé 11 fois
Estimation de lectures uniques	177,636,066 (88.68 %)	178,857,965 (89.29 %)	195,804,301 (97.75 %)

INTERFACE READSET – DÉTAILS (13/14)

Code couleur : valide, non valide, « - »

NOM du lot de séquence → **BAT_AAKOSW_2_C39G9ACXX.IND2** Disponible ← Etat (cf. workflow)

Editer

Général | Avancé

Code	BAT_AAKOSW_2_C39G9ACXX.IND2	Nb Séquences utiles	200,311,306	Run / N° Piste	140108_CARBONE_C39G9ACXX / 2
Etat	Disponible	Nb Bases utiles	39,920,324,496	Type de Run	RHS2000
Valide QC ?	Non	Valide BioInfo ?	Oui	Nb Cycles	209
Comptes Rendus QC	Problème qualité Q30 duplicat > 30	Comptes Rendus BioInfo			
Critères QC	Default	Critères BioInfo	Low		
Evalué par	ngsrg (10/03/2014)	Evalué par	ngsrg (10/03/2014)		
Détails évaluation	commentaires suppl. sur l'évalution				

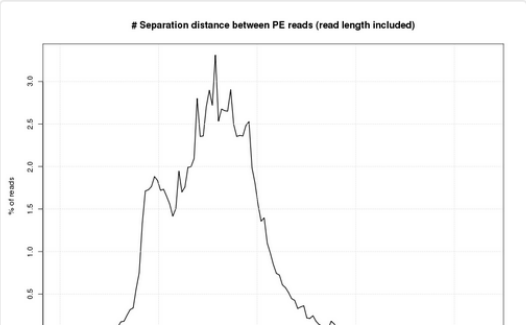
Onglets traitements [READSET] →

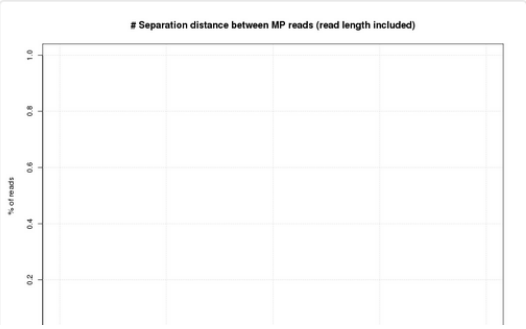
NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | Trimming (vs RAW) | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) | Read Quality (vs CLEAN) | Duplicates (vs CLEAN) | **Mapping (vs CLEAN)** | Merging (vs CLEAN)

Statistiques associées au traitement [READSET]

Données | Référence

	Sample input	Non-chimeric aligned reads	FR (PE) aligned reads	RF (MP) aligned reads	FF aligned reads	RR aligned reads	Trans aligned reads	Single aligned reads
Number	40,000	36,426	30,938	30	32	24	4,044	1,358
% total reads	100 %	91.10 %	77.30 %	0.10 %	0.10 %	0.10 %	10.10 %	3.40 %





INTERFACE READSET – DÉTAILS (14/14)

Code couleur : valide, non valide, « - »

NOM du lot de séquence → **BAT_AAKOSW_2_C39G9ACXX.IND2** Disponible ← Etat (cf. workflow)

Editer

Général | Avancé

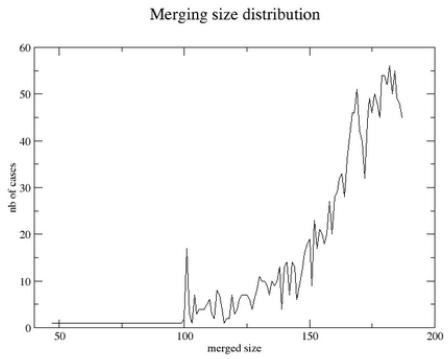
Code	BAT_AAKOSW_2_C39G9ACXX.IND2	Nb Séquences utiles	200,311,306	Run / N° Piste	140108_CARBONE_C39G9ACXX / 2
Etat	Disponible	Nb Bases utiles	39,920,324,496	Type de Run	RHS2000
Valide QC ?	Non	Valide BioInfo ?	Oui	Nb Cycles	209
Comptes Rendus QC	Problème qualité Q30 duplicat > 30	Comptes Rendus BioInfo			
Critères QC	Default	Critères BioInfo	Low		
Evalué par	ngsrg (10/03/2014)	Evalué par	ngsrg (10/03/2014)		
Détails évaluation	commentaires suppl. sur l'évalution				

Onglets traitements [READSET] →

NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | Trimming (vs RAW) | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) | Read Quality (vs CLEAN) | Duplicates (vs CLEAN) | Mapping (vs CLEAN) | Merging (vs CLEAN) *

Statistiques associées au traitement [READSET]

Sample input	Merged reads	Mediane size (bases)	Average size (bases)	Min size (bases)	Max size (bases)
20,000	1,798 (8.99 %)	170	162	47	187

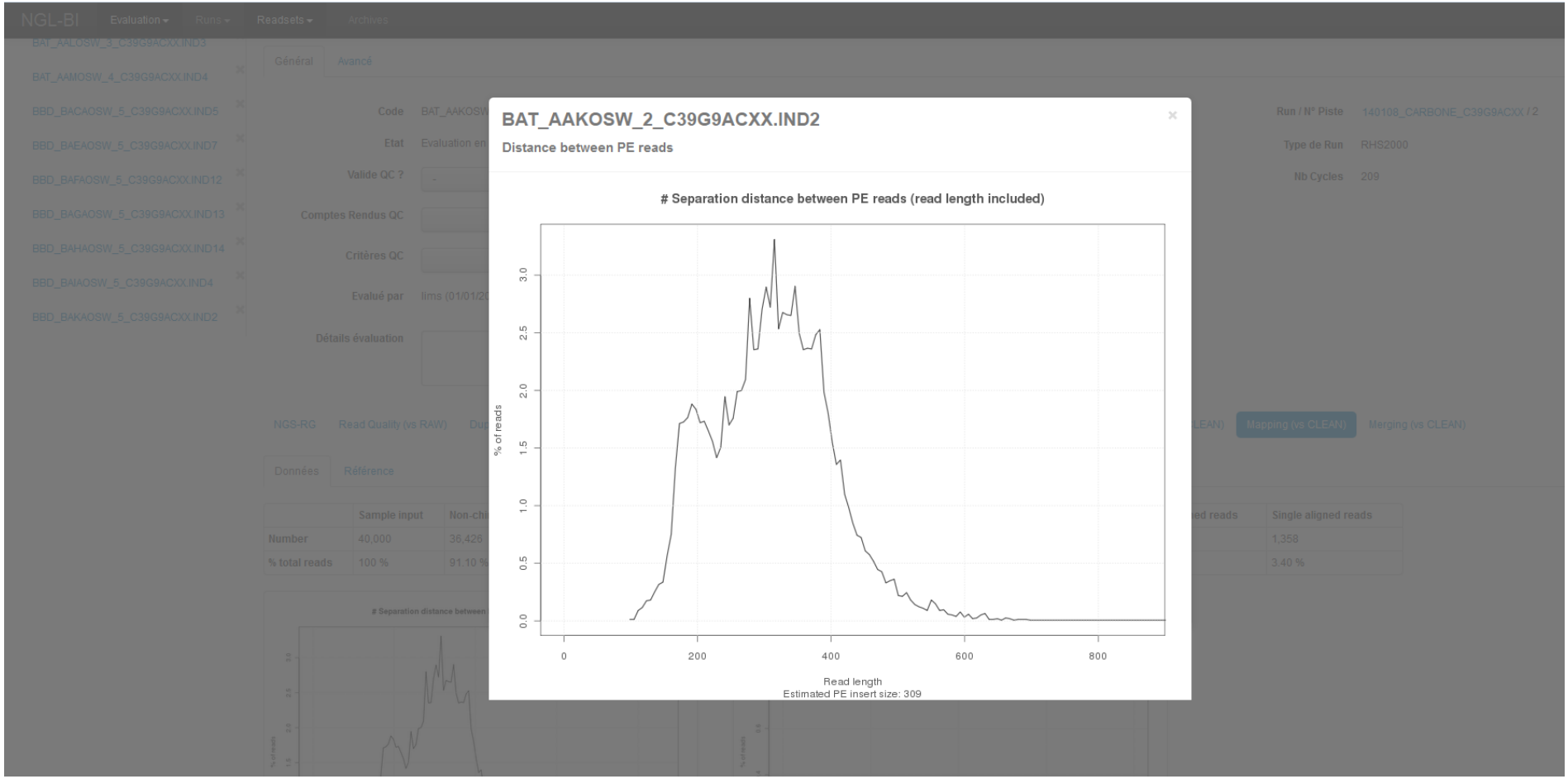


Overlap distribution

** pas encore disponible pour le CNG*

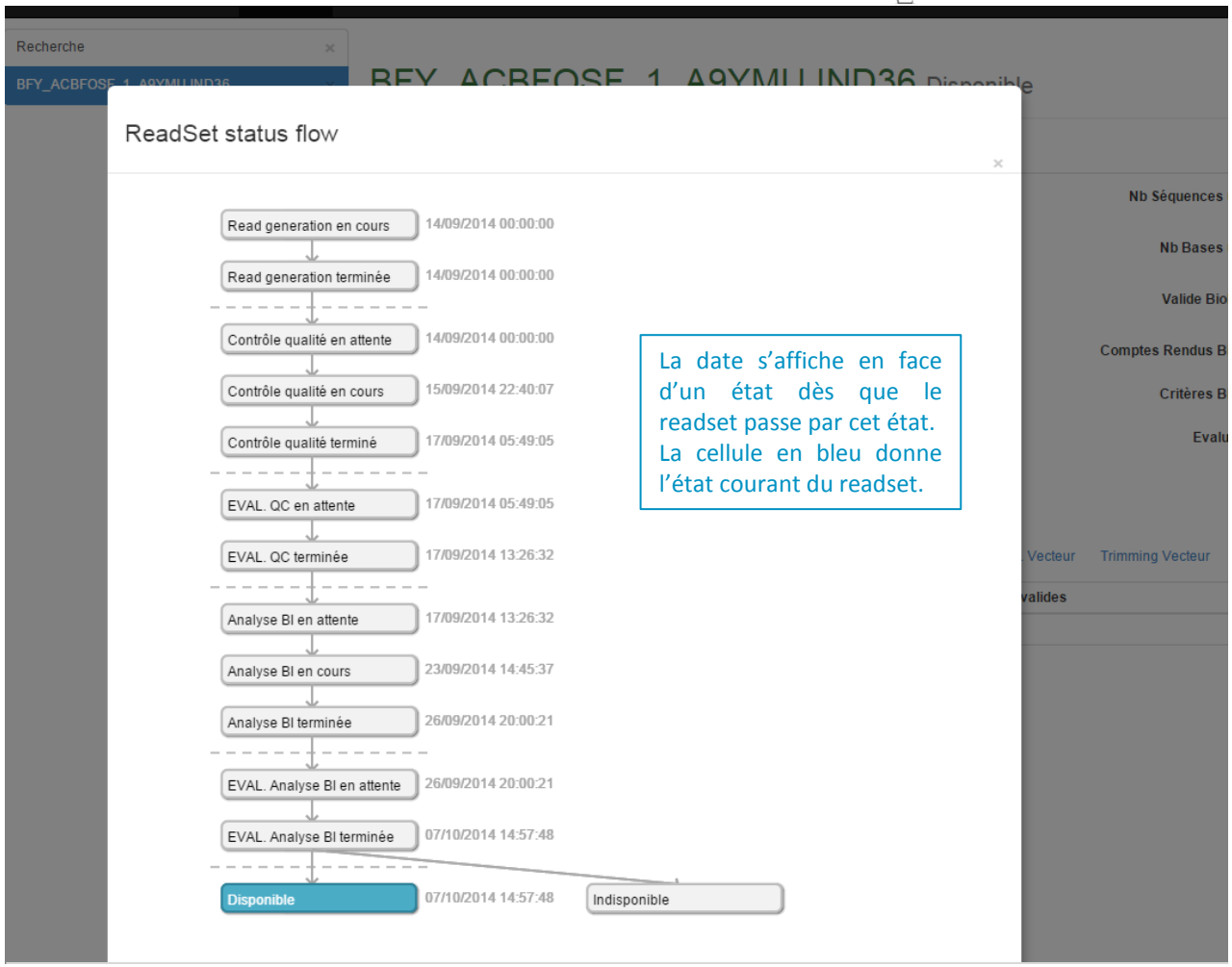
INTERFACE READSET – ZOOM IMAGES

Cliquer sur une image pour la faire apparaître en plein écran.



VISUALISER UN READSET DANS SON WORKFLOW

Cliquer sur l'état du readset pour le visualiser dans son workflow.

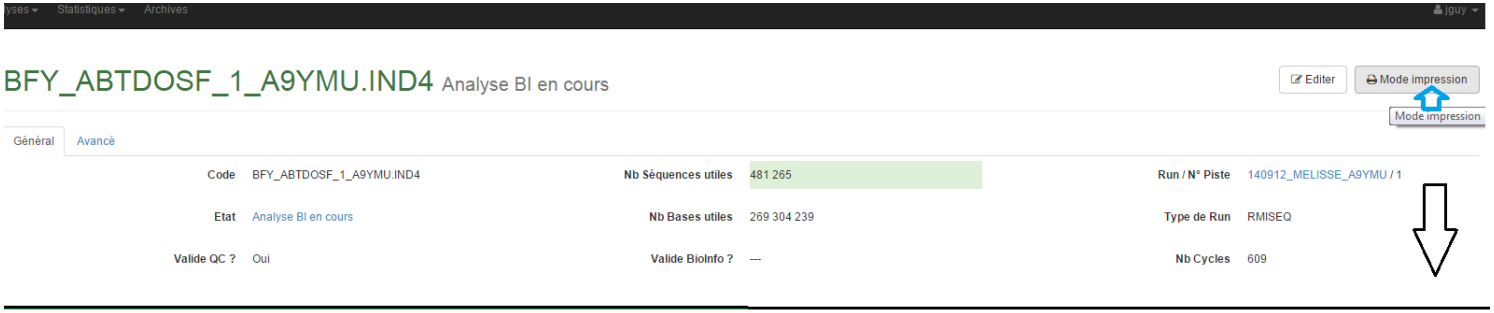


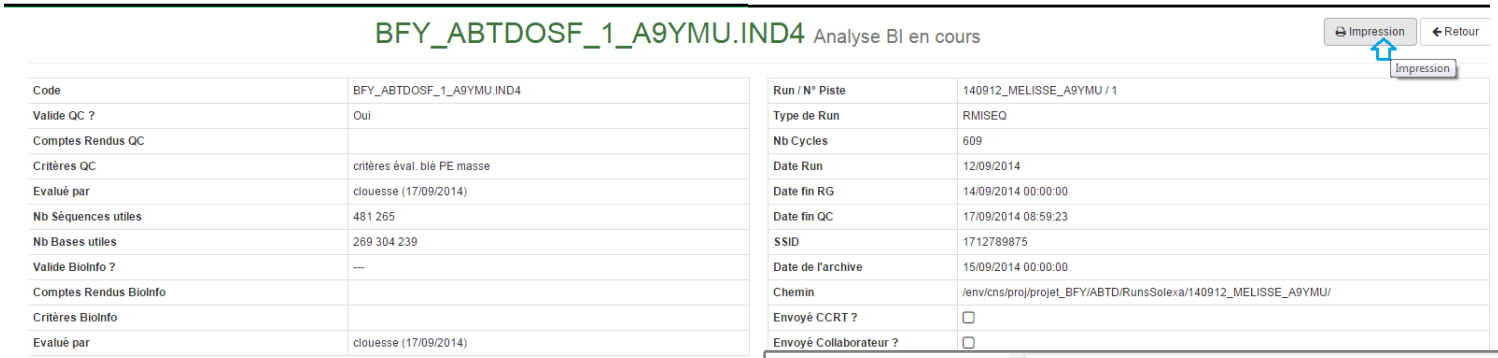
La date s'affiche en face d'un état dès que le readset passe par cet état. La cellule en bleu donne l'état courant du readset.

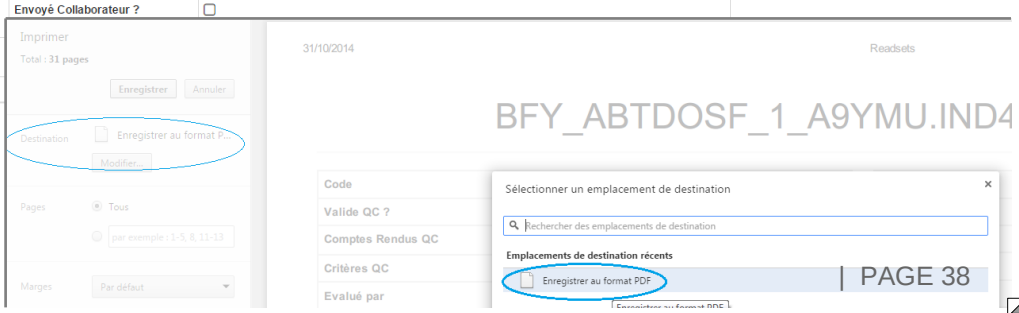
EXPORT PDF DES RÉSULTATS QUALITÉ D'UN READSET

Il est possible d'imprimer ou d'exporter en PDF les statistiques et graphiques qualité d'un readset. Pour cela :

- Se rendre sur l'interface d'un readset,
- Cliquer sur « Mode impression » en haut à droite => l'aperçu d'impression s'affiche ;
- Cliquer sur « Impression » puis choisir Destination : imprimante ou « enregistrer au format PDF »







Visualiser les détails et les traitements d’une analyse

INTERFACE ANALYSE – DÉTAILS

Code couleur : valide, non valide, « - »

NOM de l'analyse *(BA_code readset maître)* → BA_BFY_ACBAOSF_1_A9YMU.IND31 Evaluation terminée ← Etat (cf. workflow)

Permet d'éditer une analyse *(renseigner les champs : valide ?, Comptes rendus, critères, détails évaluation, etc.)* en dehors du mode « évaluation »

Editer

Général | Avancé

Code	BA_BFY_ACBAOSF_1_A9YMU.IND31	Type	BAC pool assembly
Etat	Evaluation terminée	Code ReadSets Maitres	BFY_ACBAOSF_1_A9YMU.IND31
Valid ?	Oui	Code ReadSets	BFY_ACBAOSF_1_A9YMU.IND31 BFY_ACBOSN_1_HAJK2ADXX.IND32
Comptes Rendus			
Critères éval.	critères éval. Analyse blé		
Evalué par	mangenot (07/10/2014)		
Commentaire			

Readset qui suit le workflow « avec analyses biologiques »

Readsets qui suivent le workflow « classique »

Critère utilisés pour l'évaluation ; permet de générer des highlights

Lien vers les interfaces des readsets concernés par l'analyse

Traitements effectués sur l'analyse (tableaux de statistiques, graphiques…) → Merging | Assembly | Contig Filter | Scaffolding | Gap Closing

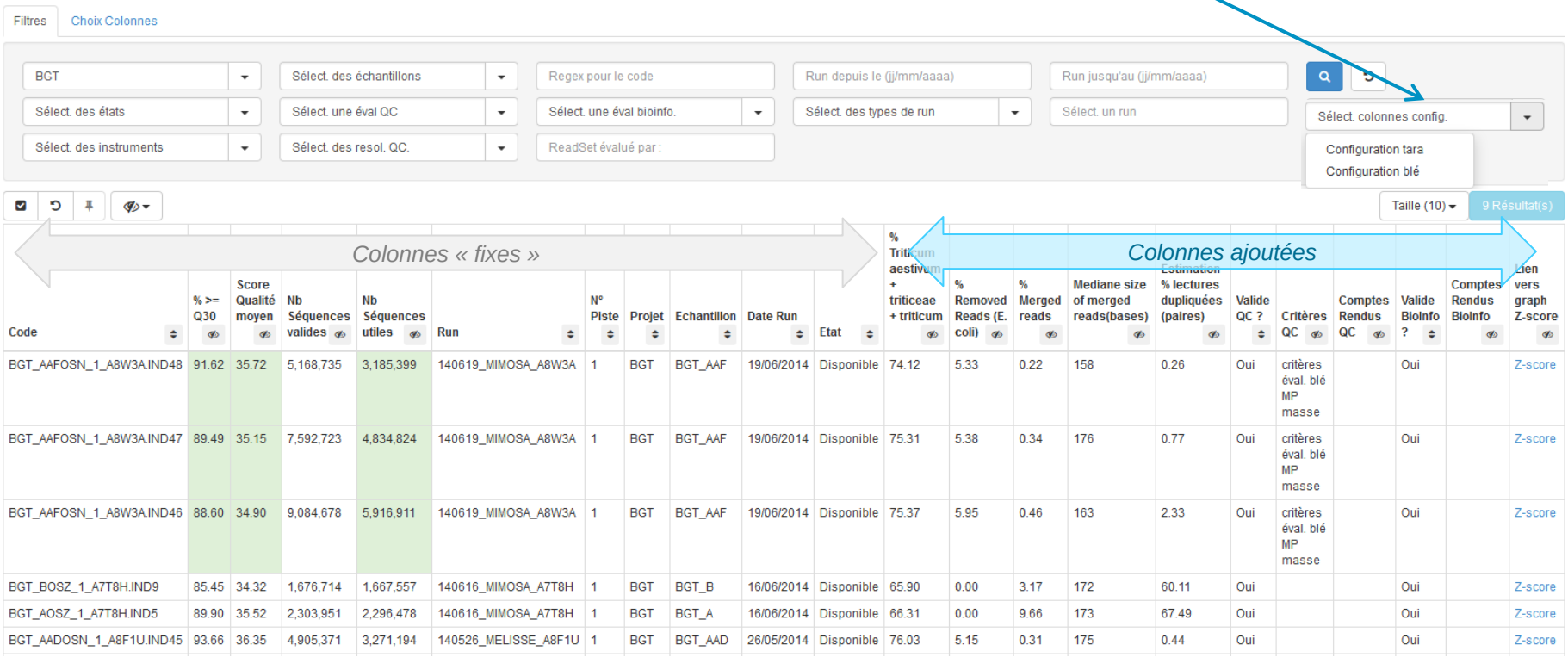
Reads input (pairs)	Merged reads	% merged reads	Mediane size (bases)	Average size (bases)	Min size (bases)	Max size (bases)
604 596	411 044	67,99	533	524,00	74	587

Visualiser des readsets ou analyses selon une configuration de tableau spécifique

CONFIGURATION SPÉCIFIQUE

Dans le cadre de certains projets où il peut être nécessaire de visualiser en masse, dans le tableau résultat des readsets, certaines colonnes supplémentaires par rapport à la configuration initiale du tableau, il est possible de visualiser des configurations spécifiques (*à partir du moment où elles ont été configurées – après demande utilisateur*).

Le choix de la configuration se fait via le bouton « Sélect. colonnes config », à droite des filtres de recherche.



Code	% >= Q30	Score Qualité moyen	Nb Séquences valides	Nb Séquences utiles	Run	N° Piste	Projet	Echantillon	Date Run	Etat	% Triticum aestivum + triticeae + triticum	% Removed Reads (E. coli)	% Merged reads	Mediane size of merged reads(bases)	Estimation % lectures dupliquées (paires)	Valide QC ?	Critères QC	Comptes Rendus QC	Valide BioInfo ?	Comptes Rendus BioInfo	Lien vers graph Z-score
BGT_AAFOSN_1_A8W3A.IND48	91.62	35.72	5,168,735	3,185,399	140619_MIMOSA_A8W3A	1	BGT	BGT_AAF	19/06/2014	Disponible	74.12	5.33	0.22	158	0.26	Oui	critères éval. blé MP masse		Oui		Z-score
BGT_AAFOSN_1_A8W3A.IND47	89.49	35.15	7,592,723	4,834,824	140619_MIMOSA_A8W3A	1	BGT	BGT_AAF	19/06/2014	Disponible	75.31	5.38	0.34	176	0.77	Oui	critères éval. blé MP masse		Oui		Z-score
BGT_AAFOSN_1_A8W3A.IND46	88.60	34.90	9,084,678	5,916,911	140619_MIMOSA_A8W3A	1	BGT	BGT_AAF	19/06/2014	Disponible	75.37	5.95	0.46	163	2.33	Oui	critères éval. blé MP masse		Oui		Z-score
BGT_BOSZ_1_A7T8H.IND9	85.45	34.32	1,676,714	1,667,557	140616_MIMOSA_A7T8H	1	BGT	BGT_B	16/06/2014	Disponible	65.90	0.00	3.17	172	60.11	Oui			Oui		Z-score
BGT_AOSZ_1_A7T8H.IND5	89.90	35.52	2,303,951	2,296,478	140616_MIMOSA_A7T8H	1	BGT	BGT_A	16/06/2014	Disponible	66.31	0.00	9.66	173	67.49	Oui			Oui		Z-score
BGT_AADOSN_1_A8F1U.IND45	93.66	36.35	4,905,371	3,271,194	140526_MELISSE_A8F1U	1	BGT	BGT_AAD	26/05/2014	Disponible	76.03	5.15	0.31	175	0.44	Oui			Oui		Z-score

Attention, les config. sont indépendantes des highlights.

Une « config » est un format de tableau donné (avec des colonnes supplémentaires par rapport au tableau par défaut).

Une « highlight » apparaît à partir du moment où un **critère d'éval. QC** est enregistré. Elle reste visible à tout moment.

Effectuer du reporting sur les readsets

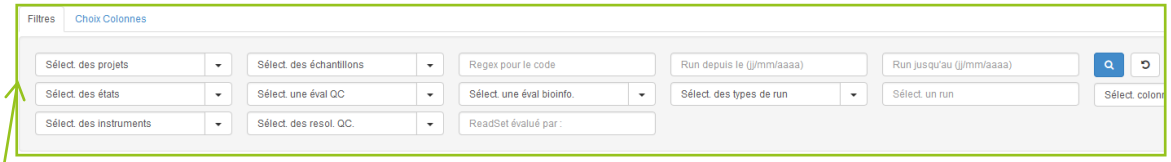
REPORTING READSETS – CHOIX COLONNES

Il existe des *configurations spécifiques* pour visualiser des readsets en masse (cf. p31).

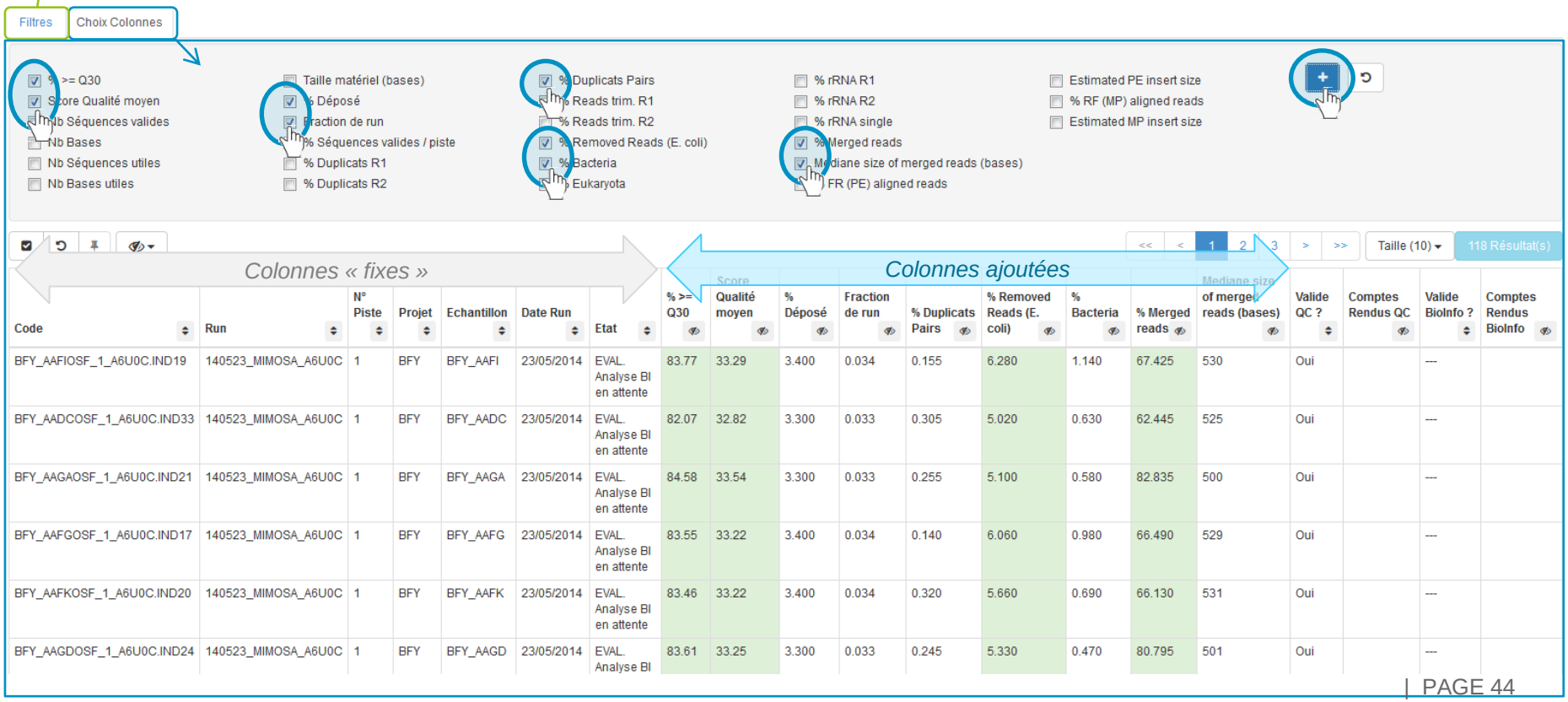
Ici, l'utilisateur est libre de configurer son tableau résultat avec les colonnes de son choix, parmi celles disponibles dans « Choix colonnes ».

1. Effectuer la recherche de readsets via l'onglet Filtres

2. Personnaliser le tableau résultat via l'onglet « choix colonnes ».



Nouveauté 1.9 : Mode « hybride » : il est possible de cumuler config + ajout de colonnes supplémentaires (dans cet ordre). *L'ajout de colonne suivi de la config réinitialise tout et donne uniquement la config.*



Code	Run	N° Piste	Projet	Echantillon	Date Run	Etat	% >= Q30	Score Qualité moyen	% Déposé	Fraction de run	% Duplicats Pairs	% Removed Reads (E. coli)	% Bacteria	% Merged reads	Mediane size of merged reads (bases)	Valide QC ?	Comptes Rendus QC	Valide BioInfo ?	Comptes Rendus BioInfo
BFY_AAFIOSF_1_A6U0C.IND19	140523_MIMOSA_A6U0C	1	BFY	BFY_AAFI	23/05/2014	EVAL. Analyse BI en attente	83.77	33.29	3.400	0.034	0.155	6.280	1.140	67.425	530	Oui		---	
BFY_AADCOSF_1_A6U0C.IND33	140523_MIMOSA_A6U0C	1	BFY	BFY_AADC	23/05/2014	EVAL. Analyse BI en attente	82.07	32.82	3.300	0.033	0.305	5.020	0.630	62.445	525	Oui		---	
BFY_AAGAOSF_1_A6U0C.IND21	140523_MIMOSA_A6U0C	1	BFY	BFY_AAGA	23/05/2014	EVAL. Analyse BI en attente	84.58	33.54	3.300	0.033	0.255	5.100	0.580	82.835	500	Oui		---	
BFY_AAFGOSF_1_A6U0C.IND17	140523_MIMOSA_A6U0C	1	BFY	BFY_AAFG	23/05/2014	EVAL. Analyse BI en attente	83.55	33.22	3.400	0.034	0.140	6.060	0.980	66.490	529	Oui		---	
BFY_AAFKOSF_1_A6U0C.IND20	140523_MIMOSA_A6U0C	1	BFY	BFY_AAFK	23/05/2014	EVAL. Analyse BI en attente	83.46	33.22	3.400	0.034	0.320	5.660	0.690	66.130	531	Oui		---	
BFY_AAGDOSF_1_A6U0C.IND24	140523_MIMOSA_A6U0C	1	BFY	BFY_AAGD	23/05/2014	EVAL. Analyse BI	83.61	33.25	3.300	0.033	0.245	5.330	0.470	80.795	501	Oui		---	

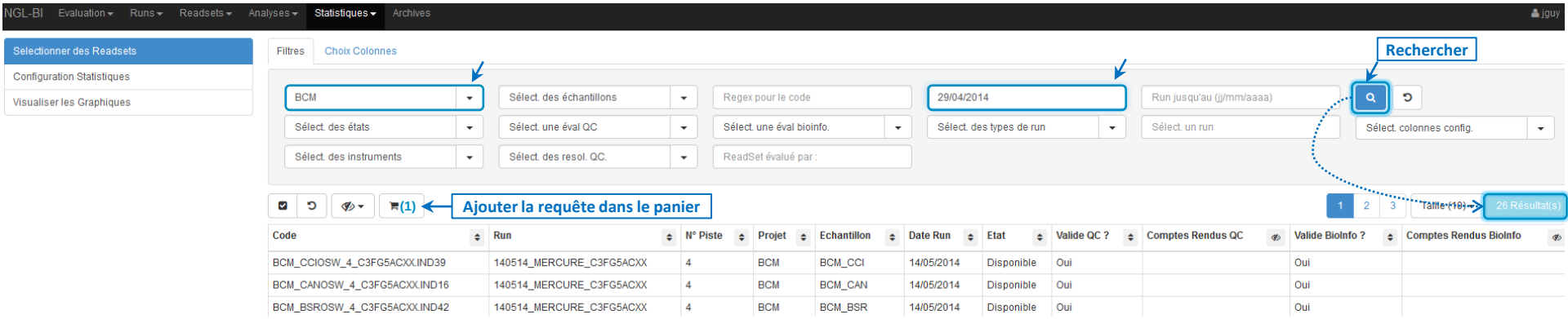
Générer des représentations graphiques de statistiques sur les readsets

SÉLECTIONNER DES READSETS
POUR GÉNÉRER DES REPRÉSENTATIONS GRAPHIQUES DE STATISTIQUES

Aller dans le menu Statistiques > Readsets.

➢ **Sélectionner des readsets**

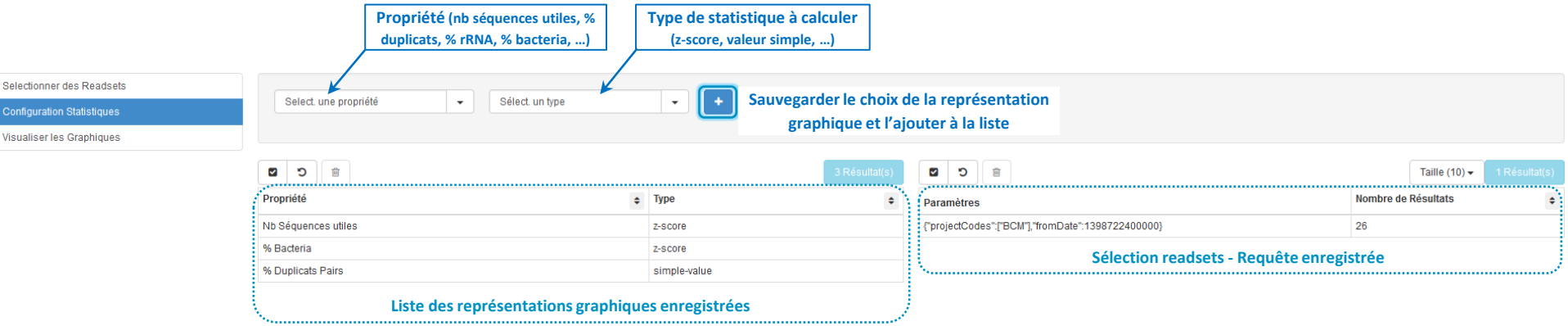
- Dans l'onglet « Filtres », renseigner les filtres pour obtenir les readsets qui nous intéresse.
- Pour enregistrer cette sélection comme celle devant être utilisée pour réaliser les graphiques, cliquez sur le *chariot* (« ajouter la requête dans le panier »).
- Si vous voulez modifier votre sélection, faites une nouvelle recherche et recliquer sur le chariot pour enregistrer cette dernière.



CONFIGURATION DES STATISTIQUES

➢ **Configuration Statistiques**

- Une fois la sélection de readsets effectuée, et la requête ajoutée au panier, cliquer sur l'onglet « Configuration statistiques » dans la barre à gauche.
- La requête enregistrée précédemment est rappelée dans un tableau à droite avec le nombre de résultats qu'elle comprend.
- Choisir une propriété et le type de statistique à afficher *(z-score, valeur simple, …)* via les boutons « select. une propriété » et « select. un type ». Cliquer sur le bouton + (ajouter) pour l'ajouter à la liste des représentations graphiques à générer.
- Pour ajouter une nouvelle représentation graphique, renseigner « select. une propriété » et « select. un type » et cliquer sur + .
- Pour supprimer une représentation graphique existante, la sélectionner et cliquer sur supprimer.



VISUALISER LES GRAPHIQUES

➢ **Visualiser les graphiques**

- Cliquer ensuite sur l'onglet « Visualiser les graphiques » dans la barre à gauche.
- Les représentations graphiques apparaissent les unes en dessous des autres, ainsi que le tableau résultat (en haut).

Selectionner des Readsets
Configuration Statistiques
Visualiser les Graphiques

Tableau résultats

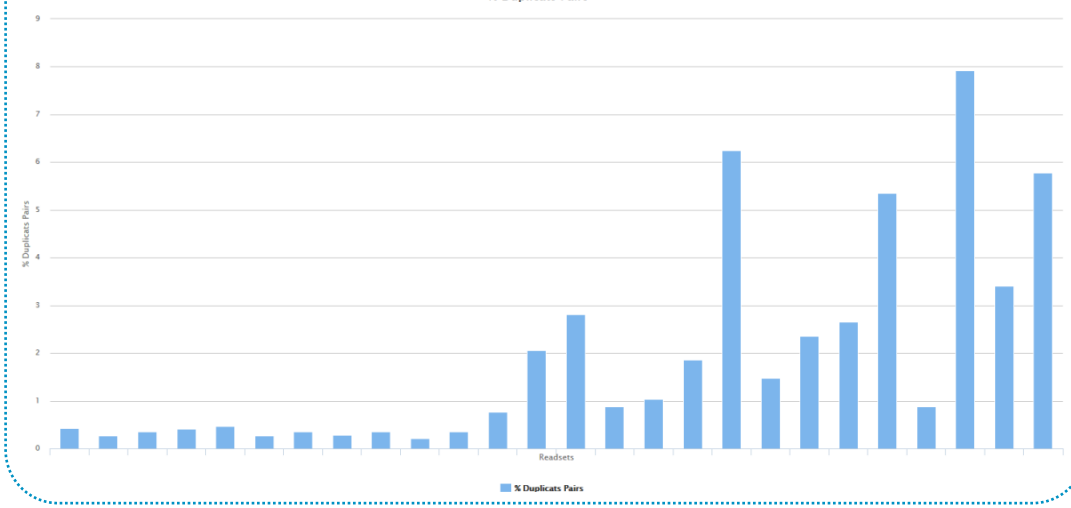
1 2 3 Taille (10) 26 Résultat(s)

Code	Run	N° Piste	Projet	Echantillon	Date Run	Etat	Nb Séquences utiles	% Duplicats Pairs	% Bacteria	Valide QC ?	Comptes Rendus QC	Valide BioInfo ?	Comptes Rendus BioInfo
BCM_AKKOSW_1_H89E9ADXX.IND20	140429_FLUOR_H89E9ADXX	1	BCM	BCM_AKK	29/04/2014	Disponible	2,123,847	0.420	0.070	Oui		Oui	
BCM_BFFOSW_1_H89E9ADXX.IND16	140429_FLUOR_H89E9ADXX	1	BCM	BCM_BFF	29/04/2014	Disponible	2,145,972	0.275	0.100	Oui		Oui	
BCM_APBOSW_1_H89E9ADXX.IND24	140429_FLUOR_H89E9ADXX	1	BCM	BCM_APB	29/04/2014	Disponible	2,480,531	0.350	0.040	Oui		Oui	
BCM_ALFOSW_1_H89E9ADXX.IND22	140429_FLUOR_H89E9ADXX	1	BCM	BCM_ALF	29/04/2014	Disponible	2,136,364	0.415	0.040	Oui		Oui	
BCM_ARDOSW_1_H89E9ADXX.IND21	140429_FLUOR_H89E9ADXX	1	BCM	BCM_ARD	29/04/2014	Disponible	2,007,483	0.470	0.030	Oui		Oui	

z-score : Nb Séquences utiles

z-score : % Bacteria

% Duplicats Pairs



Représentations graphiques choisies dans « configuration des statistiques »

VISUALISER LES GRAPHIQUES

➢ **Visualiser les graphiques**

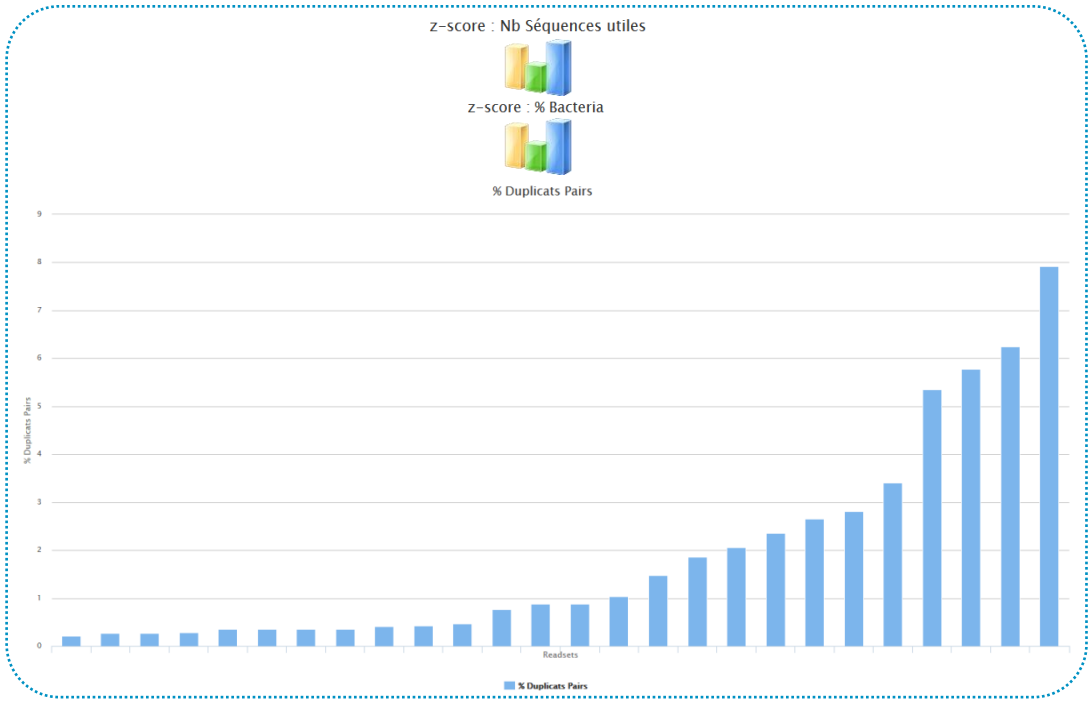
- L'abscisse (code readset) peut être triée de différentes façons : utiliser les tris sur les colonnes du tableau pour ordonner l'abscisse (code readset) par ordre alphabétique, chronologique, selon la valeur croissante ou décroissante d'une propriété, etc.

Selectionner des Readsets
Configuration Statistiques
Visualiser les Graphiques

Tableau résultats

1 2 3 Taille (10) 26 Résultat(s)

Code	Run	N° Piste	Projet	Echantillon	Date Run	Etat	Nb Séquences utiles	% Duplicats Pairs	% Bacteria	Valide QC ?	Comptes Rendus QC	Valide BioInfo ?	Comptes Rendus BioInfo
BCM_AKKOSW_1_H89E9ADXX.IND20	140429_FLUOR_H89E9ADXX	1	BCM	BCM_AKK	29/04/2014	Disponible	2,123,847	0.420	0.070	Oui		Oui	
BCM_BFFOSW_1_H89E9ADXX.IND16	140429_FLUOR_H89E9ADXX	1	BCM	BCM_BFF	29/04/2014	Disponible	2,145,972	0.275	0.100	Oui		Oui	
BCM_APBOSW_1_H89E9ADXX.IND24	140429_FLUOR_H89E9ADXX	1	BCM	BCM_APB	29/04/2014	Disponible	2,480,531	0.350	0.040	Oui		Oui	
BCM_ALFOSW_1_H89E9ADXX.IND22	140429_FLUOR_H89E9ADXX	1	BCM	BCM_ALF	29/04/2014	Disponible	2,136,364	0.415	0.040	Oui		Oui	
BCM_ARDOSW_1_H89E9ADXX.IND21	140429_FLUOR_H89E9ADXX	1	BCM	BCM_ARD	29/04/2014	Disponible	2,007,483	0.470	0.030	Oui		Oui	



Représentations graphiques choisies dans « configuration des statistiques »

TRI sur % duplicats pairs croissant *(exemple)*

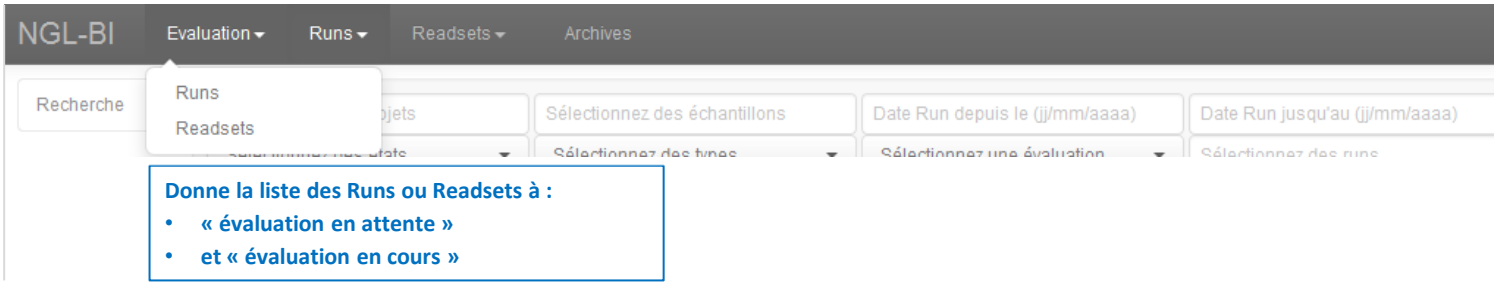
Évaluer un run

et générer des highlights via des critères d’évaluation

ÉVALUER UN RUN

Pour aller sur la liste des runs à évaluer (i.e. à « évaluation en attente » ou « évaluation en cours »), aller sur :

- Evaluation > Runs ;
- OU Runs > Evaluation.



Sélectionnez les runs sur lesquels vous voulez travailler et faire « afficher les détails ».

Le run est alors en **mode édition** *(vous pouvez renseigner les champs : valide ?, Comptes rendus, critères, détails évaluation, etc.)*

RUN - ÉVALUATION

Code couleur : Run valide, Run non valide, Run à « - »

NOM du RUN → 121203_HISEQ7_D1DAGACXX Evaluation en cours ← Etat (cf. workflow)

Ne pas oublier de sauvegarder: Sauvegarder | Annuler

Partie évaluation RUN

* liste des résolutions (triées par catégorie) fournie par les utilisateurs.

Code	121203_HISEQ7_D1DAGACXX	Type	RHS2000	Nb Cycles	101	Version RTA	1.13.48
Etat	Evaluation en cours	Nb Clusters (total)	1,217,697,789	Ligne Contrôle	1	Version Flowcell	HiSeq Flow Cell v3
Valide ?	Oui, Non, -	% Clusters filt. (moyenne)	93.93	Code Flowcell	D1DAGACXX	Code Instrument	HISEQ7
Comptes Rendus *Non obligatoire*	reso1,reso2 *CR multi valeurs	Nb Clusters filt. (total)	1,143,757,649	Position Flowcell	A	Type d'Instrument	HISEQ2000
Critères *Non obligatoire*	Critères d'évaluation utilisés	Nb Bases (total)	239,045,348,641				
Evalué par	ngsrg (20/02/2014) User et Date						
Détails évaluation *Non obligatoire*	détails éval en commentaire						

Données ngsrg [RUN]

(critère éval. Miseq PE300 par exemple)

→ **Highlights** en fonction des critères utilisés

Commentaire (saisie libre) pour apporter des détails à l'évaluation

Onglets traitements [LANES] → SAV | NGS-RG

Statistiques associées au traitement [LANES]

#	Nb Cycles	Nb Clusters	% Clusters filtrés (illumina)	Nb Cluster filtrés (illumina)	% Séquences filtrées (interne)	Nb Séquences (filtre illumina + interne)	Nb Bases (filtre illumina + interne)	% Perte
1	101,101	189,116,299	93.4	176,641,897	99.26	175,327,073	35,416,068,746	1.9
2	101,101	153,970,984	93.99	144,714,120	98.99	143,245,849	28,935,661,498	7
3	101,101	147,022,449	94.36	138,725,497	96.36	133,673,081	27,001,962,362	7.07
4	101,101	143,428,210	94.28	135,228,393	98.63	133,376,759	26,942,105,318	6.94
5	101,101	161,139,741	93.65	150,911,677	98.8	149,102,517	30,118,708,434	6.83
6	101,101	172,322,696	93.12	160,466,218	98.95	158,786,308	32,074,834,216	7.41
7	101,101	148,274,740	94.05	139,458,662	98.79	137,768,267	27,829,189,934	6.92
8	101,101	102,422,670	95.3	97,611,185	97.81	95,476,421	19,286,237,042	16.24

Partie évaluation LANES après consultation du (des) onglet(s) traitement(s)

Pistes

#	Valide ?	Comptes Rendus
		Saisie en masse
1	Non	reso1,reso2,reso3 CR multi valeurs
2	-	*Non obligatoire*
3	-	
4	-	
5	-	
6	-	
7	-	
8	-	

Récapitulatif des lots de séquences sur la FC et métriques importantes

ReadSets (10)

Sélectionnez des lanes | Voir ReadSets | Evaluer ReadSets

N° Piste	Code	Etat	% Séquences valides / piste	Nb Séquences valides	Nb Bases	% >= Q30	Score Qualité moyen	Valide QC ?	Valide BioInfo ?
1	E421_CB_B00EUIT_1_D1DAGACXX.IND4	Disponible	38.05	66,712,820	13,475,989,640	92.81	36.2	---	---
1	E421_CB_B00EPPR_1_D1DAGACXX.IND2	Disponible	37.72	66,141,643	13,360,611,886	92.7	36.15	---	---
1	E421_CB_B00EUJ1_1_D1DAGACXX.IND5	Disponible	22.33	39,157,090	7,909,732,180	92.88	36.21	---	---
2	E410_FA_B00EP3N_2_D1DAGACXX.IND5	Disponible	93	133,223,439	26,911,134,678	90.73	35.47	---	---
3	E410_FA_B00FPLS_3_D1DAGACXX.IND5	Disponible	92.93	124,226,220	25,093,696,440	90.58	35.43	---	---

ÉVALUATION RUN : CHOIX DES RÉSOLUTIONS

La liste des résolutions est commune au run et à la lane. Chaque institut (CNS / CNG) a sa propre liste de résolutions.

NGL-BI | Evaluation | Runs | Readsets | Archives | ngsrg

Runs à Evaluer

140122_CARBONE_C37T3ACXX

140122_CARBONE_C37T3ACXX Evaluation terminée

Sauvegarder | Annuler

Code	140122_CARBONE_C37T3ACXX	Type	RHS2000	Nb Cycles	209	Version RTA	1.13.48
Etat	Evaluation terminée	Nb Clusters (total)	1,724,000,233	Ligne Contrôle	0	Version Flowcell	HiSeq Flow Cell v3
Valide ?	Oui	% Clusters filt. (moyenne)	94.41	Code Flowcell	C37T3ACXX	Code Instrument	CARBONE
Comptes Rendus	rehyb primer R1,densité clu	Nb Clusters filt. (total)	1,627,640,831	Position Flowcell	A	Type d'Instrument	HISEQ2000
Critères			76,933,679				
Evalué par							
Détails évaluation							

Catégorie de résolutions

Résolutions

- Problème machine
  - indéterminé
  - chiller
  - pelletier
  - fluidique
  - laser
  - camera
  - lentille
  - pb de vide
  - PE module
  - cBot
- Problème réactifs
  - indéterminé
  - flowcell
  - cBot
  - séquencage
  - indexing
  - PE module
  - rehyb primer R1 ✔
  - rehyb primer R2
  - erreur réactifs
  - rajout réactifs
- Problème qualité : SAV
  - intensité
  - densité clusters trop élevée ✔
  - densité clusters trop faible
  - densité clusters nulle
  - %PF

NGS-RG

#	Nb Cycles	Nb Clusters	%	rés (illumina)	% Séquences filtrées (interne)	Nb Séquences (filtre illumina + interne)	Nb Bases (filtre illumina + interne)	% Perte
1	101,101	216,662,095	9		98.9	202,445,350	40,893,960,700	2.05
2	101,101	245,577,871	9		99.07	226,328,328	45,718,322,256	2.15
3	101,101	263,532,461	9		99.25	239,924,225	48,464,693,450	3.87
4	101,101	212,944,243	9		98.95	199,785,721	40,356,715,642	2.36
5	101,101	198,563,806	9		98.9	187,608,186	37,896,853,572	2.15
6	101,101	197,800,988	9		98.91	187,033,175	37,780,701,350	2.3
7	101,101	196,424,310	9		98.81	185,732,728	37,518,011,056	1.83
8	101,101	192,494,459	9		98.87	182,021,066	36,768,255,332	1.76

Pistes

#	Valide ?	Comptes Rendus
1	Non	fluidique
2	Oui	
3	Oui	
4	Non	densité clusters trop élevée
5	Oui	
6	Oui	
7	Non	phasing,prephasing,error rate
8	Oui	

Readsets (44)

ÉVALUATION LANES : CHOIX DES RÉSOLUTIONS

La liste des résolutions est commune au run et à la lane. Chaque institut (CNS / CNG) a sa propre liste de résolutions.

NGL-BI | Evaluation | Runs | Readsets | Archives | ngsrg

Runs à Evaluer

140122_CARBONE_C37T3ACXX

140122_CARBONE_C37T3ACXX Evaluation terminée

Sauvegarder | Annuler

Code	140122_CARBONE_C37T3ACXX	Type	RHS2000	Nb Cycles	209	Version RTA	1.13.48
Etat	Evaluation terminée	Nb Clusters (total)	1,724,000,233	Ligne Contrôle	0	Version Flowcell	HiSeq Flow Cell v3
Valide ?	Oui	% Clusters filt. (moyenne)	94.41	Code Flowcell	C37T3ACXX	Code Instrument	CARBONE
Comptes Rendus	rehyb primer R1,densité clu	Nb Clusters filt. (total)	1,627,640,831	Position Flowcell	A	Type d'Instrument	HISEQ2000
Critères		Nb Bases (total)	340,176,933,679				
Evalué par	ngsrg (25/02/2014)						
Détails évaluation							

NGS-RG

#	Nb Cycles	Nb Clusters	% Clusters filtrés (illumina)	Nb Cluster filtrés (illumina)	% Séquences filtrées (interne)	Nb Séquences (filtre illumina + interne)	Nb Bases (filtre illumina + interne)	% Perte
1	101,101	216,662,095	94.48	204,693,737	98.9	202,445,350	40,893,960,700	2.05
2	101,101	245,577,871	93.03	228,450,902	99.07	226,328,328	45,718,322,256	2.15
3	101,101	263,532,461	91.73	241,729,139	99.25	239,924,225	48,464,693,450	3.87
4	101,101	212,944,243	94.82	201,908,831	98.95	199,785,721	40,356,715,642	
5	101,101	198,563,806	95.53	189,685,352	98.9	187,608,186	37,896,853,572	2.15
6	101,101	197,800,988	95.6	189,099,113	98.91	187,033,175	37,780,701,350	2.3
7	101,101	196,424,310	95.7	187,972,367	98.81	185,732,728	37,518,011,056	1.83
8	101,101	192,494,459	95.64	184,101,390	98.87	182,021,066	36,768,255,332	1.76

Pistes

#	Valide ?	Comptes Rendus
1	Non	fluidique
3	Oui	
4	Non	
5	Oui	
6	Oui	
7	Non	
8	Oui	

Catégorie de résolutions → Problème machine

Résolutions:
- indéterminé
- chiller
- pelletier
- fluidique ✓
- laser
- camera
- lentille
- pb de vide
- PE module
- cBot

Problème réactifs
- indéterminé

NB : Lors du passage d'un run de «évaluation en attente (ou en cours)» à «évaluation terminée», **toute lane mise à valid=NON** impliquera le fait que **tous les readsets** à valid.QC=«-» de cette lane seront mis à **valid.QC=NON** avec le **Compte-Rendu : lane abandonnée.**

HIGHLIGHT ET CRITÈRES D'ÉVALUATION

Des **highlights** (mise en couleur de la cellule contenant la donnée à tester) peuvent être mises en place si l'on connaît les métriques à tester et les valeurs seuils à leur associer. L'ensemble de ces informations constituent un **critère d'évaluation**.

Exemple [CNS] : évaluation de run de type MISEQ – critère « RMISEQ 2x301 ». Les highlights n'apparaissent que si un critère d'évaluation est sélectionné.

140523_MIMOSA_A6U0C Evaluation en attente

Sauvegarder | Annuler

Code	140523_MIMOSA_A6U0C	Type	RMISEQ	Nb Cycles	609	Version RTA	1.18.42
Etat	Evaluation en attente	Nb Clusters (total)	20,553,440	Ligne Contrôle	0	Version Flowcell	inconnue
Valide ?	-	% Clusters filt. (moyenne)	94.86	Code Flowcell	A6U0C	Code Instrument	MIMOSA
Comptes Rendus		Nb Clusters filt. (total)	19,497,460	Position Flowcell	-	Type d'Instrument	MISEQ
Critères	RMISEQ 2x301	Nb Bases (total)	11,854,455,680				
Evalué par	jguy (11/06/2014)	A conserver ?	☐				
Détails évaluation	depot 8pM nbre de bases un peu faible						

Métriques à tester

NGS-RG

#	Nb Cycles	Nb Clusters	% Clusters filtrés (illumina)	Nb Cluster filtrés (illumina)	% Séquences filtrées (interne)	Nb Séquences (filtre illumina + interne)	Nb Bases (filtre illumina + interne)	% Perte
1	301,301	20,553,440	94.86	19,497,460	99.51	19,401,366	11,679,622,332	1.67

Pistes

#	Valide ?	Comptes Rendus
1	-	

Readsets (30)

Sélect. des lanes | Voir Readsets | Evaluer Readsets

N° Piste	Code	Etat	% déposé	% Séquences valides / piste	Nb Séquences valides	Nb Bases	% >= Q30	Score Qualité moyen	Valide QC ?	Valide BioInfo ?
1	BFY_AAFHOSF_1_A6U0C.IND18	EVAL. Analyse BI en attente	3.40	3.32	643,585	387,438,170	82.01	32.76	Oui	---
1	BFY_AAGHOSF_1_A6U0C.IND28	EVAL. Analyse BI en attente	3.30	1.80	348,375	209,721,750	83.76	33.27	Oui	---
1	BFY_AAGEOSF_1_A6U0C.IND25	EVAL. Analyse BI en attente	3.30	3.45	668,756	402,591,112	84.47	33.50	Oui	---
1	BFY_AADBOSF_1_A6U0C.IND32	EVAL. Analyse BI en attente	3.30	5.40	1,047,739	630,738,878	82.87	33.05	Oui	---
1	BFY_AAFAOSF_1_A6U0C.IND11	EVAL. Analyse BI en attente	3.40	2.93	568,223	342,070,246	81.56	32.66	Oui	---
1	BFY_AADAOSF_1_A6U0C.IND31	EVAL. Analyse BI en attente	3.30	5.33	1,033,591	622,221,782	82.81	33.04	Oui	---
1	BFY_AAGCOSF_1_A6U0C.IND27	EVAL. Analyse BI en attente	3.30	2.78	539,902	325,021,004	84.95	33.66	Oui	---

Pour connaître les règles de comparaison définissant les highligths, se référer au lien Alfresco mis dans les **Annexes**, page Critères d'évaluation définissant les Highlights.

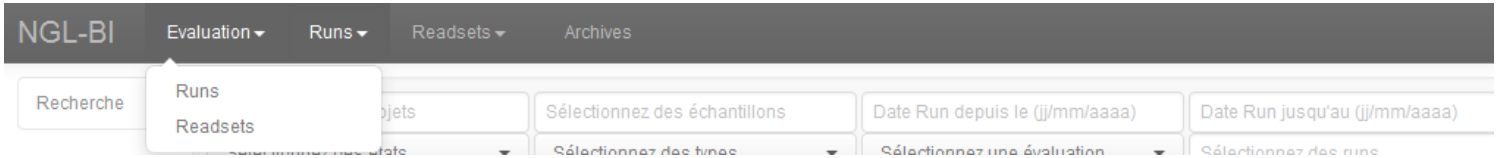
Évaluer un (des) readset(s) (possible en masse)

et générer des highlights via des critères d'évaluation

ÉVALUER UN READSET

Pour aller sur la liste des readsets à évaluer (i.e. à « évaluation en attente » ou « évaluation en cours »), aller sur :

- Evaluation > Readsets ;
- OU Readsets > Evaluation.

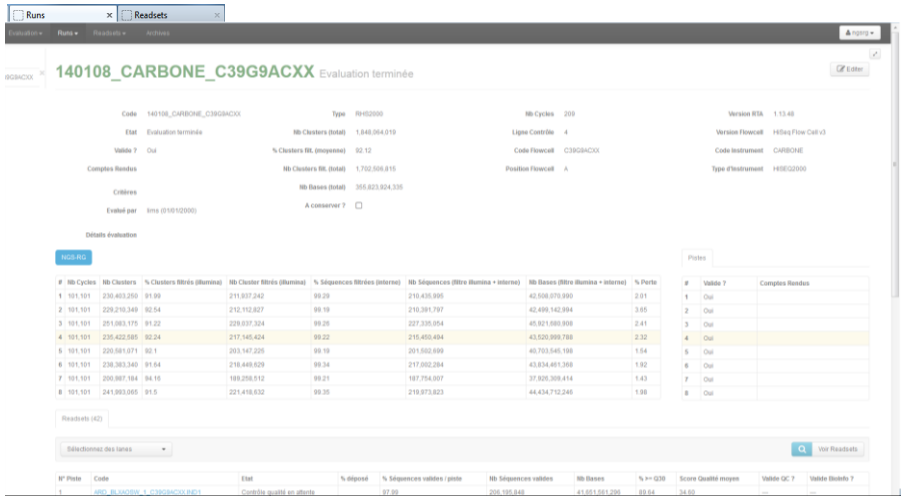


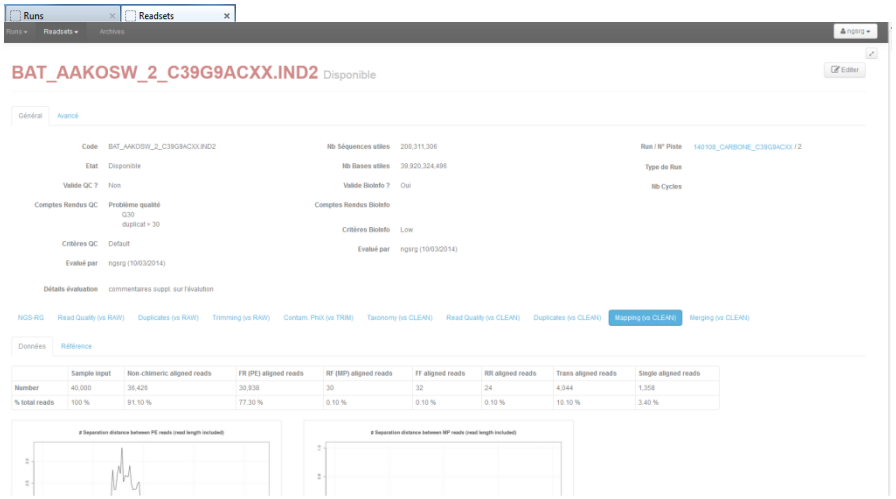
Donne la liste des Runs ou Readsets à :

- **« évaluation en attente »**
- **et « évaluation en cours »**

Sélectionnez les readsets sur lesquels vous voulez travailler et faire « afficher les détails ».

NB : vous pouvez afficher en parallèle la page correspondant au run afin de visualiser ses métriques.





READSET – ÉVALUATION

Code couleur : valide, non valide, « - »

NOM du lot de séquence → BAT_AAKOSW_2_C39G9ACXX.IND2 Evaluation en attente ← Etat (cf. workflow)

Ne pas oublier de sauvegarder

Sauvegarder | Annu

Général | Avancé

Partie évaluation readset

Code	BAT_AAKOSW_2_C39G9ACXX.IND2	Nb Séquences utiles	200,311,306	Run / N° Piste	140108_CARBONE_C39G9ACXX / 2
Etat	Contrôle qualité en attente	Nb Bases utiles	39,920,324,496	Type de Run	RHS2000
Valide QC ?	- **Valide QC ?**	Valide BioInfo ?	- **Notion de valide bioinfo**	Nb Cycles	209
Comptes Rendus QC *Non obligatoire*	**CR multi valeurs**	Comptes Rendus BioInfo *Non obligatoire*			
Critères QC *Non obligatoire*	**Critères d'évaluation utilisés**	Critères BioInfo *Non obligatoire*			
Evalué par	lims (01/01/2000) **User et Date**	Evalué par	lims (28/01/2014)		
Détails évaluation *Non obligatoire*	**Commentaire (saisie libre) pour apporter des détails à l'évaluation**				

Se met automatiquement à oui ou non quand la validation QC (oui ou non) est enregistrée. On peut ensuite y revenir pour changer la validation bioinfo uniquement.

Stat. globales

Raccourci pour aller sur le RUN

Onglets traitements [READSET] →

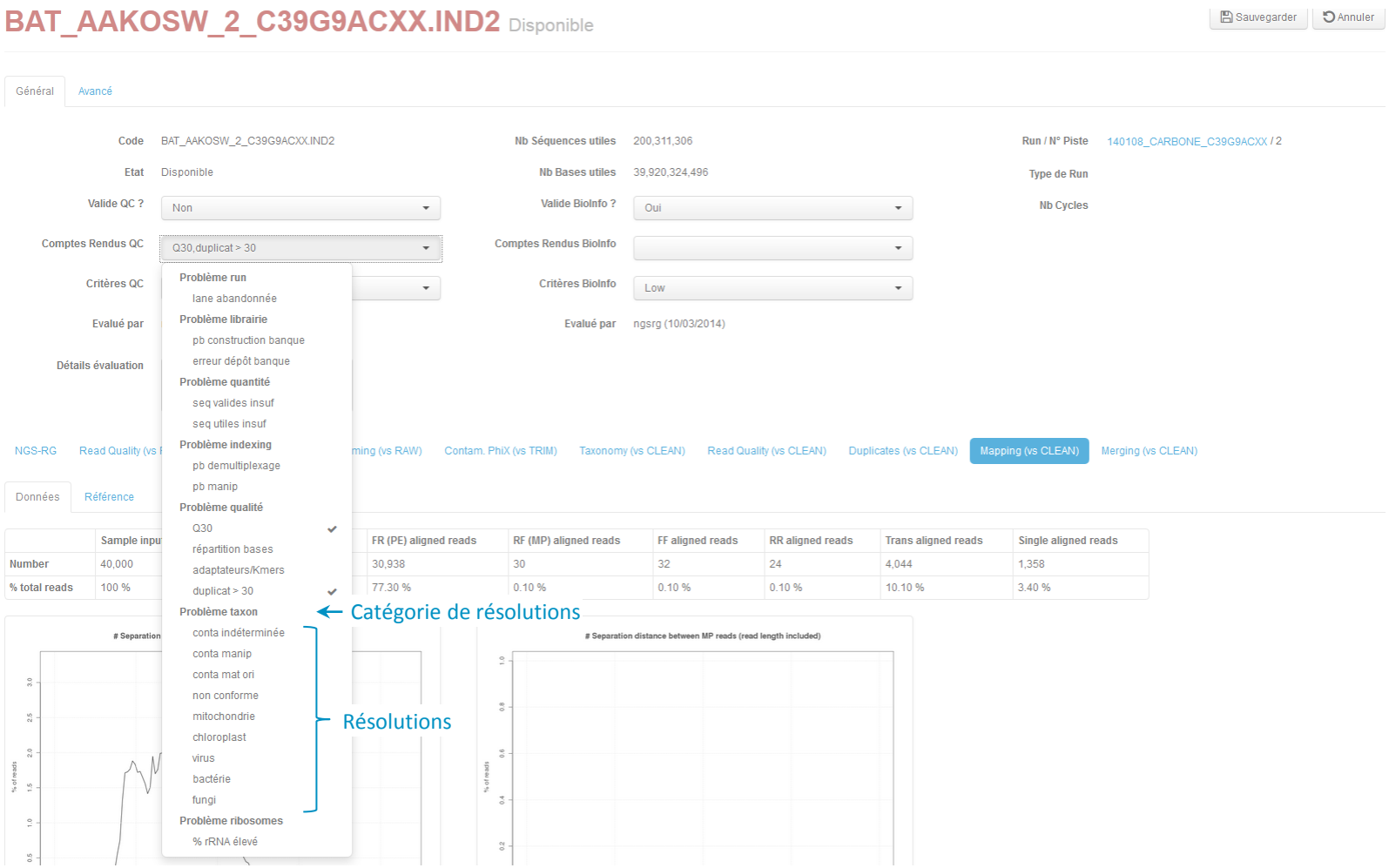
NGS-RG | Read Quality (vs RAW) | Duplicates (vs RAW) | Trimming (vs RAW) | Contam. PhiX (vs TRIM) | Taxonomy (vs CLEAN) | Read Quality (vs CLEAN) | Duplicates (vs CLEAN) | Mapping (vs CLEAN) | Merging (vs CLEAN)

Statistiques associées au traitement [READSET]

% Séquences valides / piste	Nb Séquences valides	Nb Bases	% >= Q30	Score Qualité moyen
96.35	202,710,191	40,947,458,582	92.91	36.29

ÉVALUATION READSET : CHOIX DES RÉSOLUTIONS

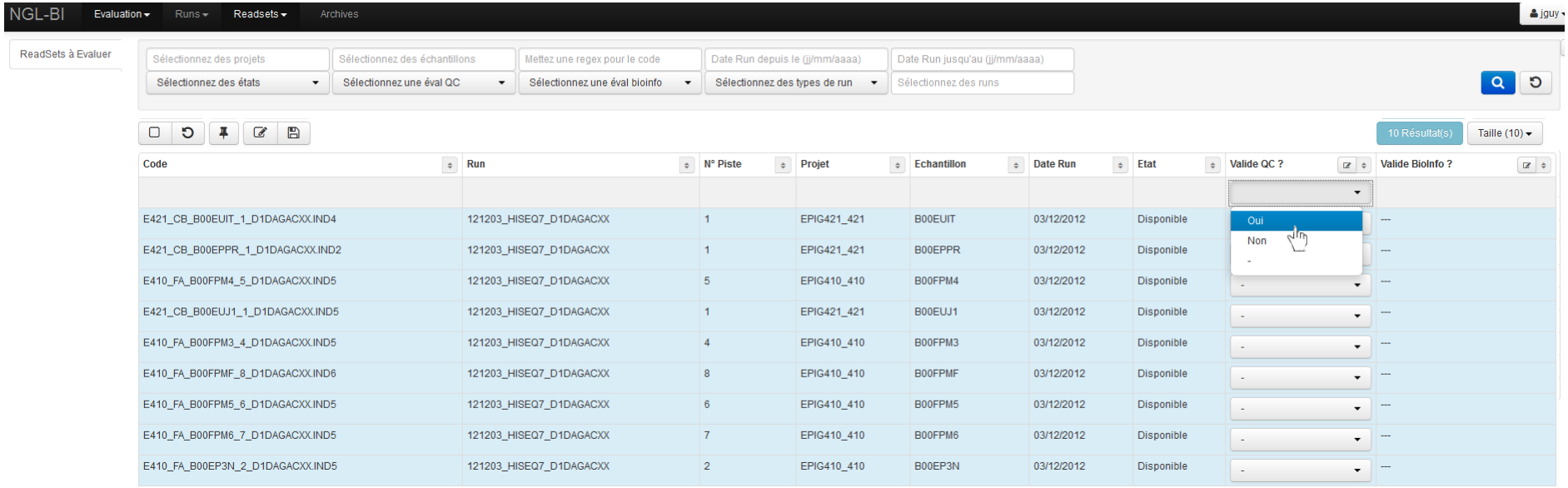
La liste des résolutions est propre aux readsets et différente de celle des runs-lanes. Chaque institut (CNS / CNG) a sa propre liste de résolutions.



ÉVALUATION DES READSETS EN MASSE

Pour valider ou invalider des readsets en masse :

- <u>Si les lots de séquences sont à « évaluation en attente » ou « évaluation en cours » (workflow normal) :</u>
  - Aller sur Readsets > Evaluation ou Evaluation > Readsets. Filtrer les readsets au moyen des filtres de recherche et cliquer sur « Tout sélectionner ».
  - Cliquez sur « Editer » en haut à gauche du tableau ou dans l'en-tête Valide QC ? et choisir la validation à appliquer à tous les readsets sélectionnés.
  - Ne pas oublier de sauvegarder (en haut à gauche du tableau).
- <u>Si les lots de séquences sont dans un état différent :</u>
  - Bloqués à « contrôle qualité en attente/en cours » car le changement automatique ne s'est pas fait : les basculer à « contrôle qualité terminé » ou « évaluation en attente »;
  - Si les lots de séquences appartiennent tous au même run, il est possible de passer par l'interface du run, de faire éditer et « évaluer les readsets » (en sélectionnant ou non au préalable certaines lanes) (exemple : quelques lanes ont été abandonnées et je souhaite directement invalider les readsets qui s'y trouvent).



HIGHLIGHT ET CRITÈRES D'ÉVALUATION

Des **highlights** (mise en couleur de la cellule contenant la donnée à tester) peuvent être mises en place si l'on connaît les métriques à tester et les valeurs seuils à leur associer. L'ensemble de ces informations constituent un **critère d'évaluation** (QC ou bioinfo).

Exemple [CNS] : éval. Readsets, critères « éval. blé MP masse ». Les highlights n'apparaissent que si un critère d'évaluation est sélectionné.

Readsets à Evaluer

Sélect. un projet | Sélect. un échantillon | Regex pour le code | Run depuis le (jj/mm/aaaa) | Run jusqu'au (jj/mm/aaaa)

EVAL. QC en attente | Sélect. une éval QC | Sélect. une éval bioinfo. | Sélect. des types de run | Sélect. un run | Configuration blé

Sélect. des instruments | Sélect. des resol. QC. | ReadSet évalué par :

Métriques à tester

1 2 Taille (10) 17 Résultat(s)

Code	% >= Q30	Score Qualité moyen	Nb Séquences valides	Nb Séquences utiles	Run	N° Piste	Projet	Echantillon	Date Run	Etat	% Triticum aestivum + triticeae + triticum	% Removed Reads (E. coli)	% Merged reads	Mediane size (bases)	Estimation % lectures dupliquées (paires)	Valide QC ?	Critères QC	Comptes Rendus QC	Valide BioInfo ?	Comptes Rendus BioInfo	Lien vers graph Z-score
																	critè				
AWK_DOSW_1_A7R3A.IND14	97.72	37.62	1,413,407	1,408,834	140217_MELISSE_A7R3A	1	AWK	AWK_D	17/02/2014	EVAL. QC en attente			67.13	20			critè		-		Z-score
BFB_AAGAOSW_1_A7R3A.IND8	94.34	36.52	1,673,772	1,657,257	140217_MELISSE_A7R3A	1	BFB	BFB_AAGA	17/02/2014	EVAL. QC en attente			94.23	142	79.01	-	critè		-		Z-score
AWU_AOSL_2_H84U3ADXX.BC7	93.83	35.50	6,532,879	6,427,835	140204_PLATINE_H84U3ADXX	2	AWU	AWU_A	04/02/2014	EVAL. QC en attente			83.31	172	9.70	-	critè		-		Z-score
BEA_PLOSW_1_A8F1N.IND31	90.36	35.10	72,158	71,845	140423_MELISSE_A8F1N	1	BEA	BEA_PL	23/04/2014	EVAL. QC en attente			97.88	191	0.04	-	critè		-		Z-score
BEA_PMOSW_1_A8F1N.IND32	87.68	34.21	227,916	226,972	140423_MELISSE_A8F1N	1	BEA	BEA_PM	23/04/2014	EVAL. QC en attente			97.38	181	0.46	-	critè		-		Z-score
BEA_PQOSW_1_A8F1N.IND35	90.88	35.28	246,847	246,149	140423_MELISSE_A8F1N	1	BEA	BEA_PQ	23/04/2014	EVAL. QC en attente	0.01		97.92	191	0.10	-	critè		-		Z-score
BEA_PSOSW_1_A8F1N.IND37	89.27	34.77	183,119	182,758	140423_MELISSE_A8F1N	1	BEA	BEA_PS	23/04/2014	EVAL. QC en attente			97.92	188	0.10	-	critè		-		Z-score
BEA_QAOSW_1_A8F1N.IND40	85.16	33.45	208,099	207,401	140423_MELISSE_A8F1N	1	BEA	BEA_QA	23/04/2014	EVAL. QC en attente	0.01		97.14	177		-	critè		-		Z-score
BEA_QBOSW_1_A8F1N.IND41b	88.54	34.45	217,259	216,830	140423_MELISSE_A8F1N	1	BEA	BEA_QB	23/04/2014	EVAL. QC en attente	0.01		97.97	180	0.14	-	critè		-		Z-score
BEA_QCOSW_1_A8F1N.IND42	89.46	34.87	195,611	194,514	140423_MELISSE_A8F1N	1	BEA	BEA_QC	23/04/2014	EVAL. QC en attente	0.01		97.54	192	0.65	-	critè		-		Z-score

critères éval. blé PE masse
critères éval. blé MP masse ✔

Pour connaître les règles de comparaison définissant les highligths, se référer au lien Alfresco mis dans les **Annexes**, page Critères d'évaluation définissant les Highlights.

Évaluer des analyses

et générer des highlights via des critères d'évaluation

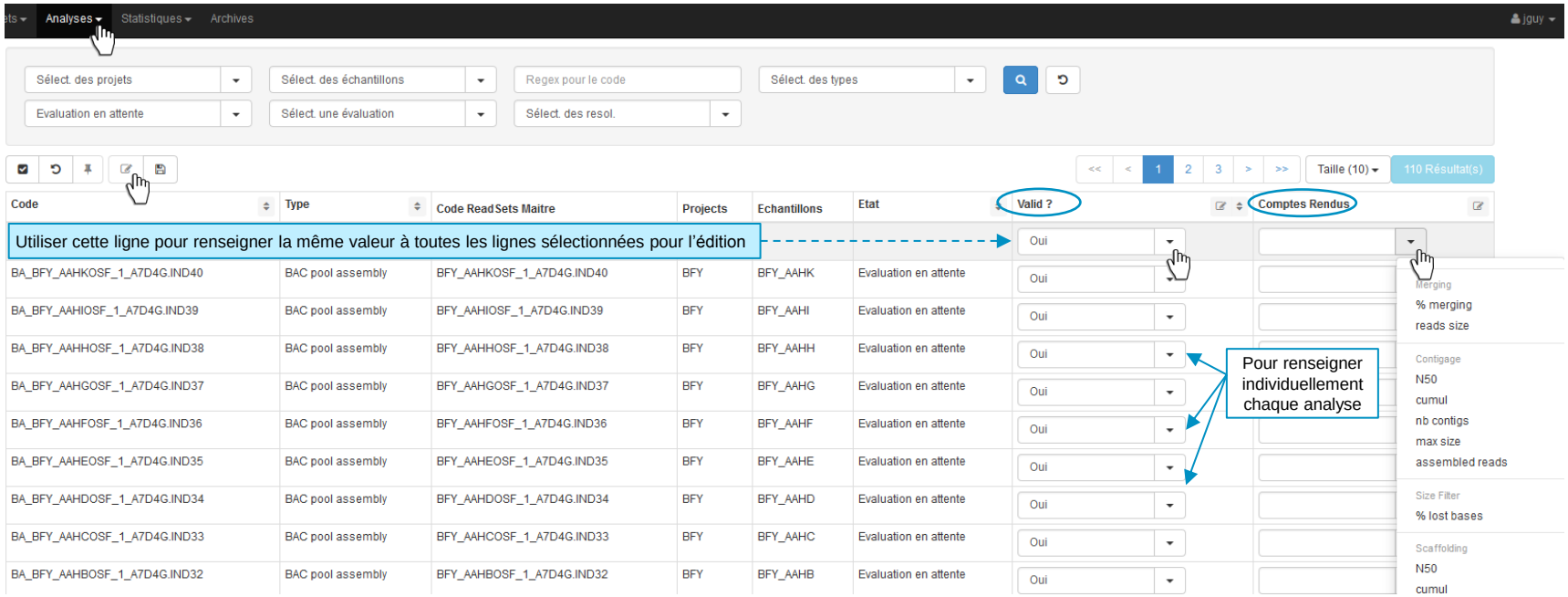
ÉVALUER DES ANALYSES

Les analyses doivent être évaluées dans NGL-BI comme un run ou un readset :

- Via le menu **Analyses > Evaluation**
- *(même principe de fonctionnement que l'évaluation en masse de readsets)*
- Colonne « Valid ? » à renseigner et « Comptes Rendus » si besoin (non obligatoire)
- Une configuration peut être utilisée pour aider à l'évaluation *(exemple : config. Analyse Blé)*
- Des critères d'évaluation peuvent être utilisés pour générer des highlights *(exemple : critère éval analyse blé)*

Workflow Analyses

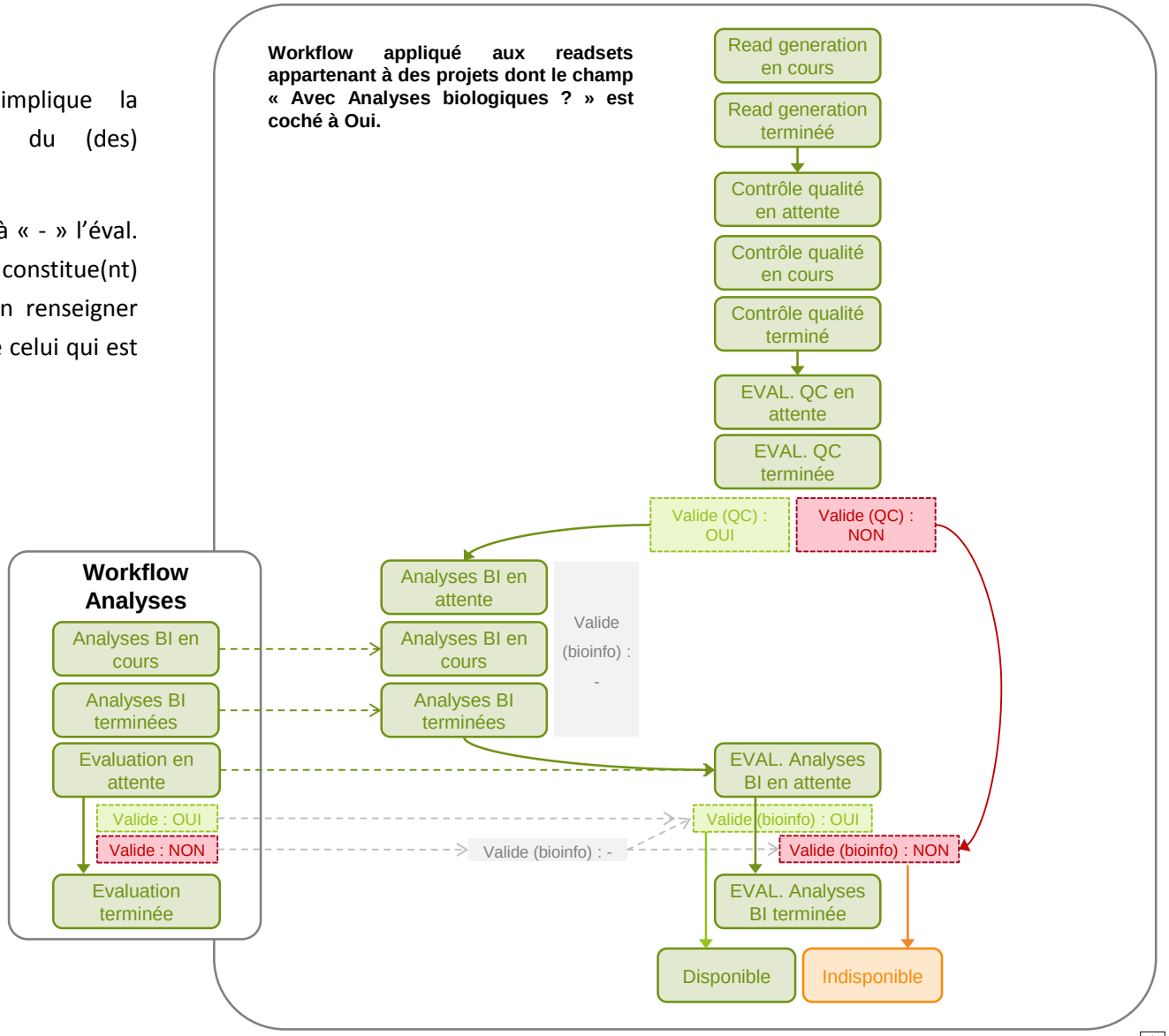
- Analyses BI en cours
- Analyses BI terminées
- Evaluation en attente
- Evaluation terminée



ÉVALUATION ANALYSES

ET INTERACTION AVEC LES READSETS QUI LA CONSTITUENT

- Règles mises en place :
  - La validation d'une analyse implique la validation bioinfo automatique du (des) readset(s) qui la constitue(nt).
  - L'invalidation d'une analyse laisse à « - » l'éval. bioinfo du (des) readset(s) qui la constitue(nt) afin de laisser le bio informaticien renseigner quel readset est invalide bioinfo de celui qui est valide bioinfo.



Retour au Workflow Readsets avec Analyses Biologiques

HIGHLIGHTS ET CRITÈRES D'ÉVALUATION

Des **highlights** (mise en couleur de la cellule contenant la donnée à tester) peuvent être mises en place si l'on connaît les métriques à tester et les valeurs seuils à leur associer. L'ensemble de ces informations constituent un **critère d'évaluation**.

Exemple [CNS] : évaluation d'analyses – critère « Analyses Blé ». Les highlights n'apparaissent que si un critère d'évaluation est sélectionné.

Sélect. des projets | Sélect. des échantillons | Regex pour le code | Sélect. des types
Evaluation en attente | Sélect. une évaluation | Analyse évaluée par : | Sélect. des résol. | Config. Analyse blé

Propriétés testées

<< < 1 2 3 > >> Taille (10) 227 Résultat(s)

Code	Type	Code ReadSets Maître	Projects	Echantillons	Etat	% merged reads	Mediane size (bases)	N50 contig size (bases)	% assembled reads	Assembly : % expected pool size	% bases perdues après Contig Filter	Number of scaffolds	N50 scaffold size (bases)	Scaffolding : % expected pool size	Gap closing : % N	Taille pool (bases)	Couverture utile (readset maître)	Couverture (Nb lect merg * mediane size / taille pool)	% Removed Reads (E. coli) (readset maître)	% Duplicats Pairs (readset maître)	Valid ?	Critères d'évaluation
																						cr
BA_BFY_ACBKOSF_1_A9YMU.IND40	BAC pool assembly	BFY_ACBKOSF_1_A9YMU.IND40	BFY	BFY_ACBK	Evaluation en attente	64,45	533	18 097	97,25	173,93	20,62	31	132 922	141,06	1,10	583 000	506,35	311,77	5,89			cr
BA_BFY_ACBIOSF_1_A9YMU.IND39	BAC pool assembly	BFY_ACBIOSF_1_A9YMU.IND39	BFY	BFY_ACBI	Evaluation en attente	67,12	534	31 589	98,27	124,31	6,71	18	538 839	117,43	0,02	735 000	433,11	276,34	6,49	0,25		cr
BA_BFY_ACBHOSF_1_A9YMU.IND38	BAC pool assembly	BFY_ACBHOSF_1_A9YMU.IND38	BFY	BFY_ACBH	Evaluation en attente	65,59	533	22 381	97,58	144,97	10,78	18	601 348	132,82	1,32	762 000	472,52	295,69	5,72	0,26		cr
BA_BFY_ACBEOSF_1_A9YMU.IND35	BAC pool assembly	BFY_ACBEOSF_1_A9YMU.IND35	BFY	BFY_ACBE	Evaluation en attente	61,23	532	31 094	94,09	163,04	10,11	19	260 322	151,63	1,54	650 000	481,90	286,03	7,33	0,26		cr
BA_BFY_ACBBOSF_1_A9YMU.IND32	BAC pool assembly	BFY_ACBBOSF_1_A9YMU.IND32	BFY	BFY_ACBB	Evaluation en attente	66,62	534	17 491	97,47	137,75	11,16	17	677 550	126,71	1,63	835 000	398,28	252,48	7,30	0,25		cr
BA_BFY_ACAKOSF_1_A9YMU.IND30	BAC pool assembly	BFY_ACAKOSF_1_A9YMU.IND30	BFY	BFY_ACAK	Evaluation en attente	68,99	539	12 746	97,44	121,04	14,74	11	309 280	109,38	4,38	375 000	897,81	592,63	9,47	1,24		cr
BA_BFY_ACAIOSF_1_A9YMU.IND29	BAC pool assembly	BFY_ACAIOSF_1_A9YMU.IND29	BFY	BFY_ACAI	Evaluation en attente	67,30	540	7 263	97,19	93,09	27,02	28	122 671	71,18	1,66	713 000	329,03	211,91	8,49	0,50		cr
BA_BFY_ACAHOSF_1_A9YMU.IND28	BAC pool assembly	BFY_ACAHOSF_1_A9YMU.IND28	BFY	BFY_ACAH	Evaluation en attente	69,06	540	5 293	97,33	82,74	22,05	21	111 383	71,26	5,67	671 000	243,47	160,57	10,03	0,81		cr
BA_BFY_ACAGOSF_1_A9YMU.IND27	BAC pool assembly	BFY_ACAGOSF_1_A9YMU.IND27	BFY	BFY_ACAG	Evaluation en attente	69,78	539	11 553	96,68	111,29	15,02	19	127 094	98,25	2,54	778 000	294,67	196,43	10,13	0,37		cr
BA_BFY_ACAFOSF_1_A9YMU.IND26	BAC pool assembly	BFY_ACAFOSF_1_A9YMU.IND26	BFY	BFY_ACAF	Evaluation en attente	71,10	539	16 983	97,35	123,17	16,07	13	359 026	104,39	0,00	603 000	400,11	270,70	9,70	0,32		cr

critères éval. Analyse blé

Pour connaître les règles de comparaison définissant les highligths, se référer au lien Alfresco mis dans les **Annexes**, page Critères d'évaluation définissant les Highlights.

Fonctionnalités du tableau de résultats

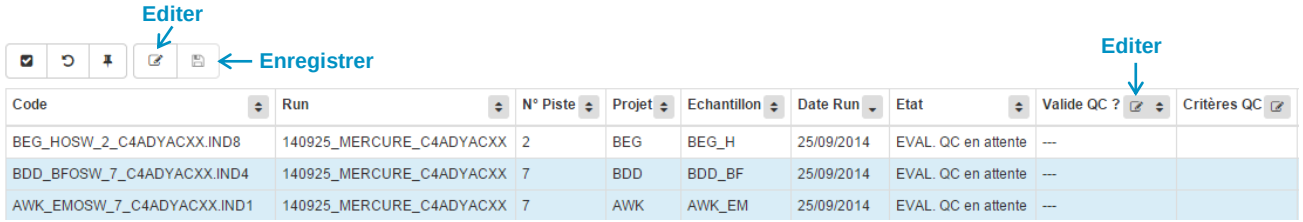
FONCTIONNALITÉS DU TABLEAU RÉSULTATS

Tableau général :

- **Tout Sélectionner** : sélectionne toutes les lignes du tableau de la page courante ;
- **Annuler** : annule la sélection effectuée ou les valeurs renseignées au cours de l'édition (si la sauvegarde n'a pas encore eu lieu) ;
- **Afficher Détails** : « épingle » les objets sélectionnés dans le volet de gauche ; en cliquant sur ceux-ci on consulte leur interface individuelle ;
- **Export CSV** : permet d'exporter tous les résultats de la recherche dans un fichier .csv ;
- **Cacher** : permet de sélectionner les colonnes du tableau à cacher / à faire réapparaître ;
- **Editer** *(quand menu autre que recherche) :* permet d'éditer les lignes sélectionnées sur toutes les colonnes éditables ;
- **Enregistrer** *(quand menu autre que recherche) :* enregistre les valeurs renseignées en édition.

Colonnes :

- **Tri** *(croissant / décroissant)* : s'effectue sur tous les résultats de la recherche, pas seulement sur les lignes du tableau de la page courante ;
- **Grouper** : permet de grouper toutes les lignes d'un même run, projet ou échantillon (selon la colonne qu'on a choisi de grouper) et de faire apparaître sur une ligne la somme, moyenne (ou autre) des valeurs de chaque colonne.
- **Cacher** : permet de cacher la colonne / de la faire réapparaître ;
- **Editer** : permet d'éditer les lignes sélectionnées sur cette colonne uniquement.

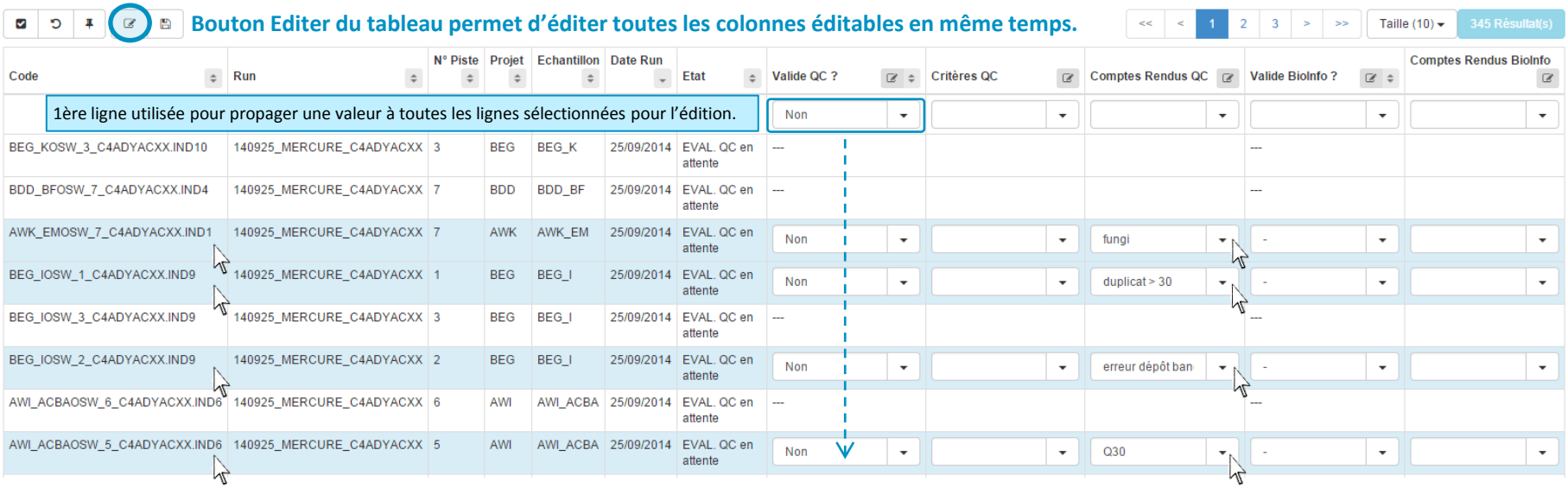


Fonction éditer dans un tableau

FONCTION ÉDITER DANS UN TABLEAU (1/2)

- Pour éditer des cellules dans un tableau, commencer par sélectionner les lignes à éditer, via le bouton « Sélectionner tout » si l'on souhaite sélectionner toutes les lignes du tableau affiché (NB : ne sélectionne pas tous les résultats de la recherche), ou en cliquant une par une sur chaque ligne concernée.
- Cliquer ensuite sur le bouton éditer.
- La 1ère ligne qui s'affiche sert à renseigner une valeur 1 seule fois et de la propager à toutes les cellules en mode édition
- *Exemple :*

Bouton Editer du tableau permet d'éditer toutes les colonnes éditables en même temps.

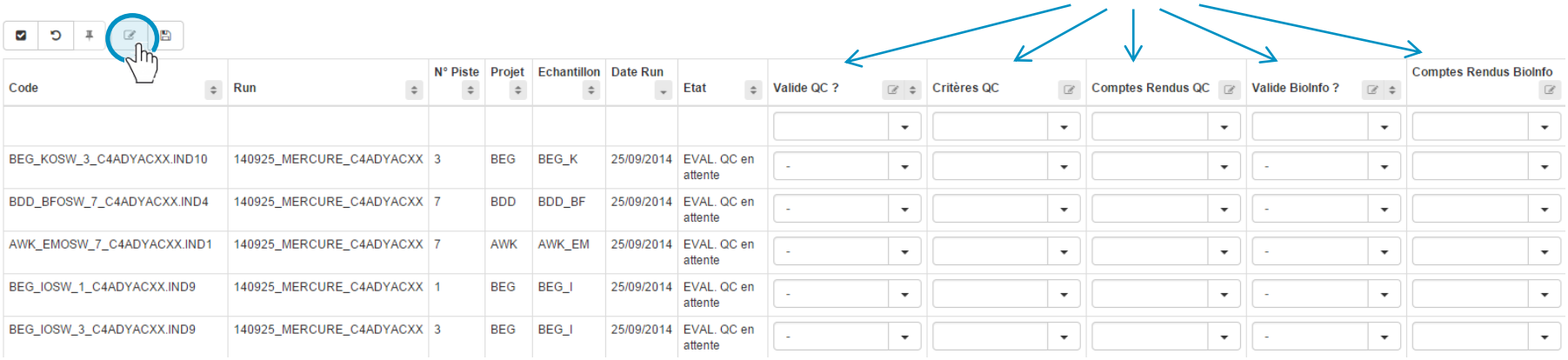


Lignes sélectionnées individuellement, en cliquant dessus.

Chaque ligne renseignée indépendamment pour CR QC

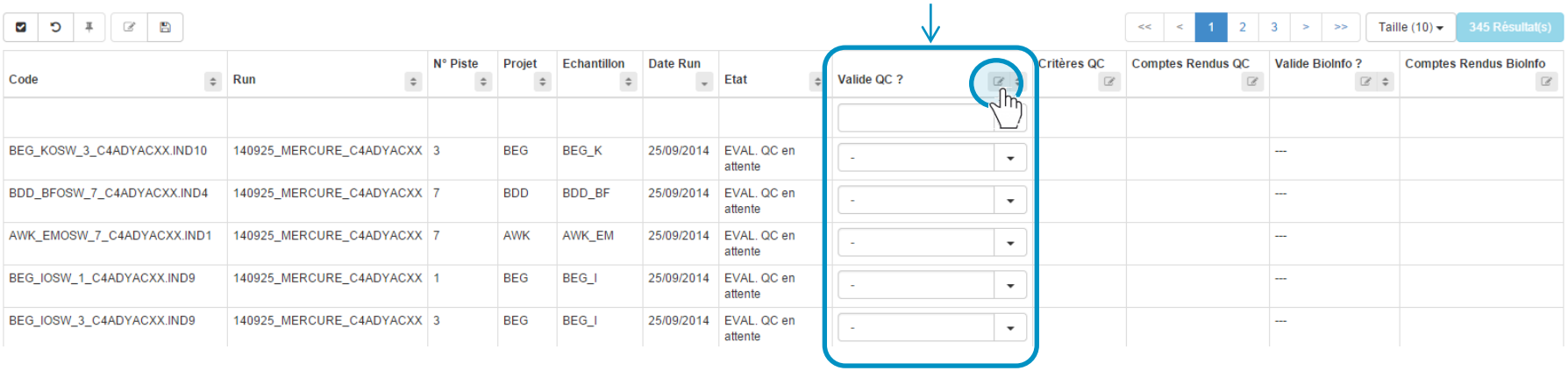
FONCTION ÉDITER DANS UN TABLEAU (2/2)

- **Le bouton « Editer » en haut à gauche du tableau permet d'éditer toutes les colonnes éditables en même temps.**



Code	Run	N° Piste	Projet	Echantillon	Date Run	Etat	Valide QC ?	Critères QC	Comptes Rendus QC	Valide BioInfo ?	Comptes Rendus BioInfo
BEG_KOSW_3_C4ADYACXX.IND10	140925_MERCURE_C4ADYACXX	3	BEG	BEG_K	25/09/2014	EVAL. QC en attente	-			-	
BDD_BFOSW_7_C4ADYACXX.IND4	140925_MERCURE_C4ADYACXX	7	BDD	BDD_BF	25/09/2014	EVAL. QC en attente	-			-	
AWK_EMOSW_7_C4ADYACXX.IND1	140925_MERCURE_C4ADYACXX	7	AWK	AWK_EM	25/09/2014	EVAL. QC en attente	-			-	
BEG_IOSW_1_C4ADYACXX.IND9	140925_MERCURE_C4ADYACXX	1	BEG	BEG_I	25/09/2014	EVAL. QC en attente	-			-	
BEG_IOSW_3_C4ADYACXX.IND9	140925_MERCURE_C4ADYACXX	3	BEG	BEG_I	25/09/2014	EVAL. QC en attente	-			-	

- **Le bouton « Editer » en haut d'un colonne permet d'éditer uniquement cette colonne.**



Code	Run	N° Piste	Projet	Echantillon	Date Run	Etat	Valide QC ?	Critères QC	Comptes Rendus QC	Valide BioInfo ?	Comptes Rendus BioInfo
BEG_KOSW_3_C4ADYACXX.IND10	140925_MERCURE_C4ADYACXX	3	BEG	BEG_K	25/09/2014	EVAL. QC en attente	-			---	
BDD_BFOSW_7_C4ADYACXX.IND4	140925_MERCURE_C4ADYACXX	7	BDD	BDD_BF	25/09/2014	EVAL. QC en attente	-			---	
AWK_EMOSW_7_C4ADYACXX.IND1	140925_MERCURE_C4ADYACXX	7	AWK	AWK_EM	25/09/2014	EVAL. QC en attente	-			---	
BEG_IOSW_1_C4ADYACXX.IND9	140925_MERCURE_C4ADYACXX	1	BEG	BEG_I	25/09/2014	EVAL. QC en attente	-			---	
BEG_IOSW_3_C4ADYACXX.IND9	140925_MERCURE_C4ADYACXX	3	BEG	BEG_I	25/09/2014	EVAL. QC en attente	-			---	

Exporter les résultats de tableaux en CSV

EXPORT DES RÉSULTATS EN CSV

Le bouton « **Export CSV** » situé au dessus du tableau permet d'exporter en fichier CSV l'ensemble des résultats affichés dans le tableau (disponible pour les menus run, readset, analyses). Si la fonction GROUPER est utilisée, l'export des lignes groupées (uniquement) est possible.

Readsets | Analyses | Statistiques | Archives | jguy

Filtres | Filtres supplémentaires | Colonnes supplémentaires

- [x] Type Run
- [x] Type d'échantillon
- [x] Taille échantillon (bases)
- [x] % >= Q30
- [x] Score Qualité moyen
- [x] Nb Séquences valides
- [ ] Nb Bases
- [x] Couverture (Nb Bases/Taille échantillon)
- [x] Nb Séquences utiles
- [ ] Nb Bases utiles
- [x] Couverture utile (Nb Bases utiles/Taille échantillon)
- [x] % Déposé
- [x] Fraction de run
- [x] % Séquences valides / piste
- [ ] % Duplicats R1
- [ ] % Duplicats R2
- [ ] % Duplicats Pairs
- [ ] % Reads trim. R1
- [ ] % Reads trim. R2
- [ ] Stored Pairs ap. trimming
- [ ] % Removed Reads (E. coli)
- [ ] Remaining Reads ap. decontamPhiX
- [ ] % Bacteria
- [ ] % Eukaryota
- [ ] % Unknown::No hits
- [ ] % rRNA R1
- [ ] % rRNA R2
- [ ] % rRNA single
- [ ] Useful Sequences ap. sortingRibo
- [ ] % Merged reads
- [ ] Mediane size of merged reads (bases)
- [ ] % FR (PE) aligned reads
- [ ] Estimated PE insert size
- [ ] % RF (MP) aligned reads
- [ ] Estimated MP insert size

Export CSV

<< < 1 2 3 > >> Taille (10) 360 Résultat(s)

Code	Run	Type Run	N° Piste	Projet	Echantillon	Type d'échantillon	Taille échantillon (bases)	Date Run	Etat	% >= Q30	Score Qualité moyen	Nb Séquences valides	Couverture (Nb Bases/Taille échantillon)	Nb Séquences utiles	Couverture utile (Nb Bases utiles/Taille échantillon)	% Déposé	Fraction de run	% Séquences valides / piste
AMZ_AOSF_1_A9Y40.IND16	141006_MELISSE_A9Y40	RMISEQ	1	AMZ	AMZ_A	ADN Génomique	800 000 000	06/10/2014	Disponible	72,35	29,86	30 901 122	23,25	30 587 278	19,54	100,00	1	97,24
BHD_APOSW_5_C4ACVACXX.IND4	141001_CARBONE_C4ACVACXX	RHS2000	5	BHD	BHD_AP	ARN total	1 000	01/10/2014	Disponible	95,71	37,03	25 025 283	5 055 107,17	24 530 217	4 901 393,34	16,67	0,021	14,32
BHD_AQOSW_5_C4ACVACXX.IND5	141001_CARBONE_C4ACVACXX	RHS2000	5	BHD	BHD_AQ	ARN total	1 000	01/10/2014	Disponible	95,71	37,02	27 293 430	5 513 272,86	26 863 274	5 374 444,80	16,67	0,021	15,62
BHD_ATOSW_5_C4ACVACXX.IND12	141001_CARBONE_C4ACVACXX	RHS2000	5	BHD	BHD_AT	ARN total	1 000	01/10/2014	Disponible	95,28	36,86	30 250 029	6 110 505,86	29 560 783	5 925 684,99	16,67	0,021	17,31
BGN_ABLOSW_3_C4ACVACXX.IND14	141001_CARBONE_C4ACVACXX	RHS2000	3	BGN	BGN_ABL	ADN Génomique	1 242 060 000	01/10/2014	Disponible	90,38	35,48	2 480 506	0,40	2 426 818	0,39	2,08	0,003	1,19
BGN_SDOSW_3_C4ACVACXX.IND43	141001_CARBONE_C4ACVACXX	RHS2000	3	BGN	BGN_SD	ADN Génomique	2 000 000 000	01/10/2014	Disponible	92,12	36,00	4 136 673	0,42	4 066 089	0,41	2,08	0,003	1,98
BGN_MROSW_1_C4ACVACXX.IND20	141001_CARBONE_C4ACVACXX	RHS2000	1	BGN	BGN_MR	ADN Génomique	2 000 000 000	01/10/2014	Disponible	87,67	34,74	2 651 895	0,27	2 574 550	0,26	2,08	0,003	1,12
BGN_NEOSW_1_C4ACVACXX.IND28	141001_CARBONE_C4ACVACXX	RHS2000	1	BGN	BGN_NE	ADN Génomique	2 000 000 000	01/10/2014	Disponible	87,83	34,64	4 764 418	0,48	4 630 037	0,46	2,08	0,003	2,01
BGN_NSOSW_1_C4ACVACXX.IND40	141001_CARBONE_C4ACVACXX	RHS2000	1	BGN	BGN_NS	ADN Génomique	1 711 500 000	01/10/2014	Disponible	89,45	35,22	4 709 354	0,56	4 605 320	0,53	2,08	0,003	1,99
BGN_MSOSW_1_C4ACVACXX.IND21	141001_CARBONE_C4ACVACXX	RHS2000	1	BGN	BGN_MS	ADN Génomique	2 000 000 000	01/10/2014	Disponible	90,87	35,62	6 707 806	0,68	6 614 693	0,66	2,08	0,003	2,84

Export des 360 résultats

Readsets_20141020_1....csv

Fonctionnalité « GROUPER » dans un tableau

FONCTIONNALITÉ « GROUPER » DANS TABLEAU

- **Grouper** : permet de grouper toutes les lignes d'un même run, projet ou échantillon (selon la colonne qu'on a choisi de grouper) et de faire apparaître sur une ligne la somme, moyenne (ou autre) des valeurs de chaque colonne.

Code	Run	N° Piste	Projet	Echantillon	Type d'échantillon	Taille échantillon (bases)	Date Run	Etat	% >= Q30	Nb Séquences valides	Couverture (Nb Bases/Taille échantillon)	% Duplicats R1	% Unknown::No hits	% Merged reads	Mediane size of merged reads (bases)	Valide QC ?	Comptes Rendus QC	Valide BioInfo ?	Comptes Rendus BioInfo
AEY_HAOSW_1_AAD7G.IND3	140922_MELISSE_AAD7G	1	AEY	AEY_HA	ADN Génomique	3 300 000	22/09/2014	Disponible	77,82	1 523 759	231,80	0,89	35,25	85,28	231	Oui	**Problème qualité** adaptateurs/Kmers	Oui	
AEY_GVOSW_1_AAD7G.IND2	140922_MELISSE_AAD7G	1	AEY	AEY_GV	ADN Génomique	3 300 000	22/09/2014	Disponible	81,97	2 013 268	306,26	2,66	12,29	87,47	231	Oui	**Problème qualité** adaptateurs/Kmers	Oui	
AEY_MIOSW_1_AAD7G.IND8	140922_MELISSE_AAD7G	1	AEY	AEY_MI	ADN Génomique	3 300 000	22/09/2014	Disponible	82,59	3 180 098	483,76	51,94	48,85	87,84	233	Oui	**Problème qualité** adaptateurs/Kmers	Oui	
AEY_MKOSW_1_AAD7G.IND4	140922_MELISSE_AAD7G	1	AEY	AEY_MK	ADN Génomique	3 300 000	22/09/2014	Disponible	74,93	2 012 598	306,16	0,92	44,60	84,42	224	Oui	**Problème qualité** adaptateurs/Kmers	Oui	
AEY_HFOSW_1_AAD7G.IND5	140922_MELISSE_AAD7G	1	AEY	AEY_HF	ADN Génomique	3 300 000	22/09/2014	Disponible	86,99	3 415 900	519,63	13,18	0,55	87,58	241	Oui	**Problème qualité** adaptateurs/Kmers	Oui	
AEY_GTOSW_1_AAD7G.IND1	140922_MELISSE_AAD7G	1	AEY	AEY_GT	ADN Génomique	3 300 000	22/09/2014	Disponible	84,66	2 121 420	322,71	4,63	2,51	89,58	232	Oui	**Problème qualité** adaptateurs/Kmers	Oui	
AEY_HGOSW_1_AAD7G.IND6	140922_MELISSE_AAD7G	1	AEY	AEY_HG	ADN Génomique	3 300 000	22/09/2014	Disponible	87,57	3 161 718	480,96	29,27	0,64	87,82	244	Oui	**Problème qualité** adaptateurs/Kmers	Oui	
AEY_HHOSW_1_AAD7G.IND7	140922_MELISSE_AAD7G	1	AEY	AEY_HH	ADN Génomique	3 300 000	22/09/2014	Disponible	86,73	3 209 257	488,20	45,49	14,41	87,51	243	Oui	**Problème qualité** adaptateurs/Kmers	Oui	
			AEY			3 300 000			82,91	20 638 018	3 139,48	18,62	19,89	87,19	234,875				
AHX_BQAIOSF_1_HA4CYADXX.IND8	140922_FLUOR_HA4CYADXX	1	AHX	AHX_BQAI	ADN Métagénomique	10 000 000	22/09/2014	Disponible	90,17	145 151 902	2 932,07	10,92	97,84	87,80	171	Oui		Oui	
AHX_BPTIOSF_4_C3YNEACXX.IND6	140910_PHOSPHORE_C3YNEACXX	4	AHX	AHX_BPTI	ADN Métagénomique	10 000 000	10/09/2014	Disponible	84,10	227 220 665	4 589,86	9,24	97,00	89,73	164	Oui		Oui	
AHX_BPSIOSF_4_C3YEPACXX.IND5	140910_SOUFRE_C3YEPACXX	4	AHX	AHX_BPSI	ADN Métagénomique	10 000 000	10/09/2014	Disponible	88,26	204 396 875	4 128,82	8,97	97,44	93,01	163	Oui		Oui	
			AHX			10 000 000			87,51	576 769 442	11 650,74	9,71	97,43	90,18	166				
ALP_KHOSW_8_C3YNEACXX.IND1	140910_PHOSPHORE_C3YNEACXX	8	ALP	ALP_KH	ADN Métagénomique	10 000 000	10/09/2014	Disponible	86,25	64 569 420	1 304,30	79,58	99,53	0,25	137	Oui	**Problème quantité** seq valides insuf	Oui	
			ALP			10 000 000			86,25	64 569 420	1 304,30	79,58	99,53	0,25	137				
BDY_AACOSW_1_A9YNN.IND18	140915_MELISSE_A9YNN	1	BDY	BDY_AAC	Amplicon	429	15/09/2014	Disponible	78,54	2 170 472	3 045 743,93	67,11	0,09	95,11	430	Oui		Oui	
BDY_AADOSW_1_A9YNN.IND19	140915_MELISSE_A9YNN	1	BDY	BDY_AAD	Amplicon	370	15/09/2014	Disponible	81,36	2 157 439	3 510 211,56	68,31	0,95	96,52	372	Oui		Oui	
BDY_AAAOSW_1_A9YNN.IND5	140915_MELISSE_A9YNN	1	BDY	BDY_AAA	Amplicon	370	15/09/2014	Disponible	80,07	1 962 135	3 192 446,68	71,97	0,69	97,50	372	Oui		Oui	
BDY_AABOSW_1_A9YNN.IND22	140915_MELISSE_A9YNN	1	BDY	BDY_AAB	Amplicon	429	15/09/2014	Disponible	78,31	1 697 052	2 381 410,97	69,74	0,14	95,10	429	Oui		Oui	
			BDY			#MULTI			79,57	7 987 098	12 129 813,13	69,28	0,47	96,06	400,75				
									Moyenne	*Somme*	*Somme*	*Moy.*	*Moy.*	*Moy.*	*Moy.*				

Grouper

Pour connaître les règles de calculs utilisés lorsque la fonction GROUPE est utilisée, se référer au lien Alfresco mis dans les **Annexes**, page colonnes supplémentaires et Comportement si la Fonctionnalité « Groupe » du tableau est utilisée.

GROUPER SUR TOUTE LA SÉLECTION

- **La fonction Grouper** peut être utilisée sur une colonne du tableau OU sur toute la sélection.

Filtres | Filtres supplémentaires | Colonnes supplémentaires

Sélect. des projets | Sélect. des échantillons | Regex pour le code | 01/09/2014 | Run jusqu'au (jj/mm/aaaa)
Sélect. des états | Sélect. une éval QC | Sélect. une éval bioinfo. | Sélect. des types de run | Sélect. un run
Sélect. des instruments | Sélect. des résol. QC. | Sélect. des résol. bioinfo. | ReadSet évalué par :

<< < 1 2 3 > >> Taille (10) 1228 Résultat(s)

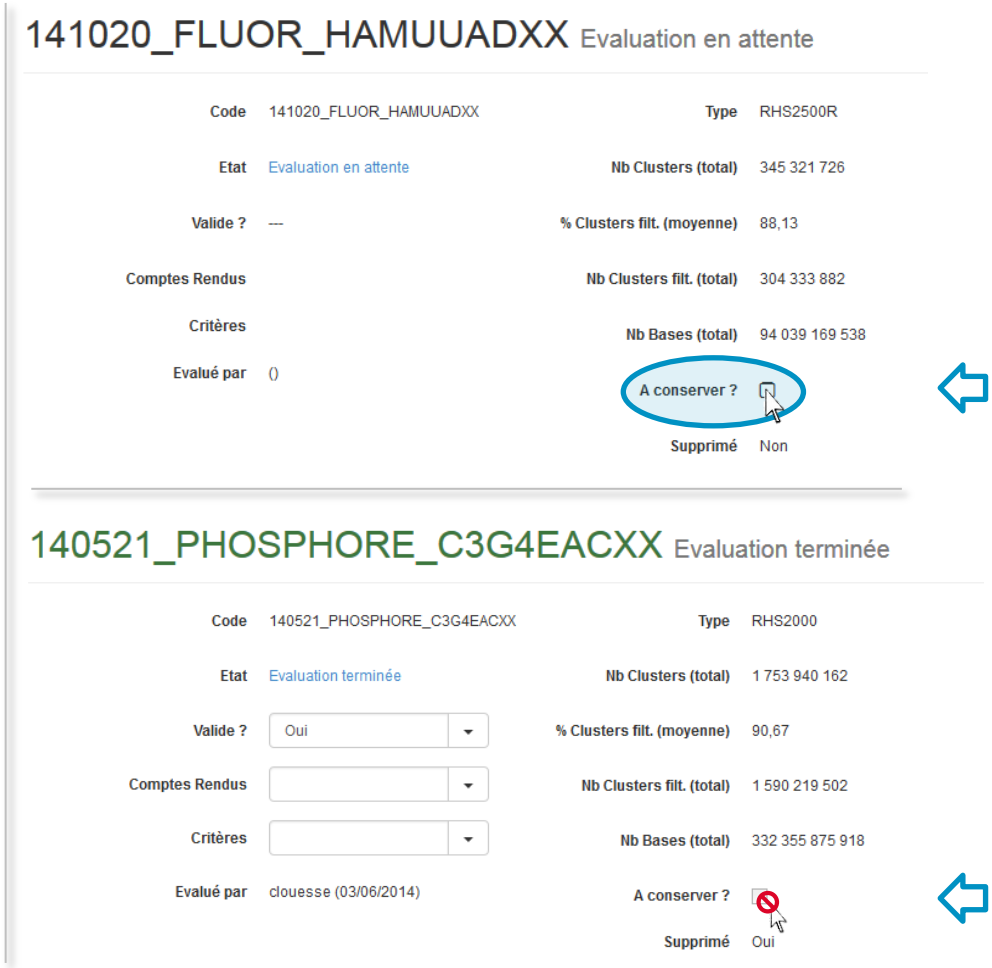
Code		N° Piste	Projet	Echantillon	Date Run	Etat	Score Qualité moyen	Nb Séquences valides	Nb Bases	% Duplicats Pairs	Valide QC ?	Comptes Rendus QC	Valide BioInfo ?	Comptes Rendus BioInfo
A361_DA_D000P2X_8_C5J	C5JK2ACXX	8	AGING_361	D000P2X	27/11/2014	EVAL. Analyse BI en attente	36,65	145 252 241	29 340 952 682	0,58	Oui		---	
S464_RD_D000P3J_4_C5JK2ACXX.IND5	141127_HISEQ6_C5JK2ACXX	4	SPLICOTER_464	D000P3J	27/11/2014	Analyse BI en attente	36,1	62 846 149	12 694 922 098	33,01	Oui		---	
S464_RD_D000P3K_5_C5JK2ACXX.IND5	141127_HISEQ6_C5JK2ACXX	5	SPLICOTER_464	D000P3K	27/11/2014	Analyse BI en attente	35,76	44 644 698	9 018 228 996	39,87	Oui		---	
S464_RD_D000P2A_1_C5JK2ACXX.IND2	141127_HISEQ6_C5JK2ACXX	1	SPLICOTER_464	D000P2A	27/11/2014	Analyse BI en attente	35,86	45 751 991	9 241 902 182	40,67	Oui		---	
A361_DA_D000P2X_7_C5JK2ACXX.IND5	141127_HISEQ6_C5JK2ACXX	7	AGING_361	D000P2X	27/11/2014	EVAL. Analyse BI en attente	36,59	149 288 890	30 156 355 780	0,65	Oui		---	
S464_RD_D000P3S_3_C5JK2ACXX.IND15	141127_HISEQ6_C5JK2ACXX	3	SPLICOTER_464	D000P3S	27/11/2014	Analyse BI en attente	35,84	51 194 826	10 341 354 852	33,92	Oui		---	
S464_RD_D000P2F_2_C5JK2ACXX.IND7	141127_HISEQ6_C5JK2ACXX	2	SPLICOTER_464	D000P2F	27/11/2014	Analyse BI en attente	35,54	46 316 180	9 355 868 360	24,16	Oui		---	
S464_RD_D000P3Q_5_C5JK2ACXX.IND6	141127_HISEQ6_C5JK2ACXX	5	SPLICOTER_464	D000P3Q	27/11/2014	Analyse BI en attente	35,61	52 081 452	10 520 453 304	38,20	Oui		---	
S464_RD_D000P39_2_C5JK2ACXX.IND19	141127_HISEQ6_C5JK2ACXX	2	SPLICOTER_464	D000P39	27/11/2014	Analyse BI en attente	35,38	48 823 082	9 862 262 564	33,68	Oui		---	
S464_RD_D000P41_6_C5JK2ACXX.IND21	141127_HISEQ6_C5JK2ACXX	6	SPLICOTER_464	D000P41	27/11/2014	Analyse BI en attente	36,12	33 584 764	6 784 122 328	19,23	Oui		---	
LIGNE GROUPEE sur toute la sélection							34,660	84 902 834 267	17 096 785 211 434	#ERROR				

Pour connaître les règles de calculs utilisés lorsque la fonction GROUPE est utilisée, se référer au lien Alfresco mis dans les **Annexes**, page colonnes supplémentaires et Comportement si la Fonctionnalité « Groupe » du tableau est utilisée.

Renseigner des runs
Renseigner des readsets

RENSEIGNER DES RUNS : A CONSERVER

Le champ **« A conserver »** est utilisé par la bioinfo pour exclure du pipeline automatique les runs à ne pas effacer (actif au CNG).
Ce champ peut être coché **à condition que le run ne soit pas déjà supprimé.**



Le run peut être coché comme « à conserver »

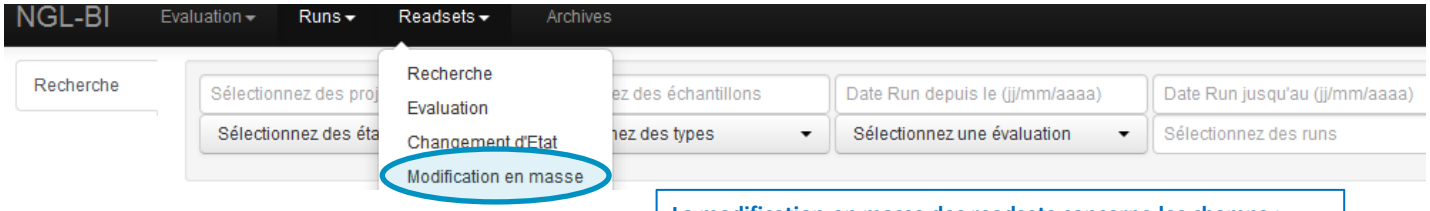
Le run ne peut pas être coché comme « à conserver » ; il a déjà été supprimé.

RENSEIGNER DES READSETS

Il est possible d'indiquer si un readset a été envoyé au CCRT et / ou au collaborateur.

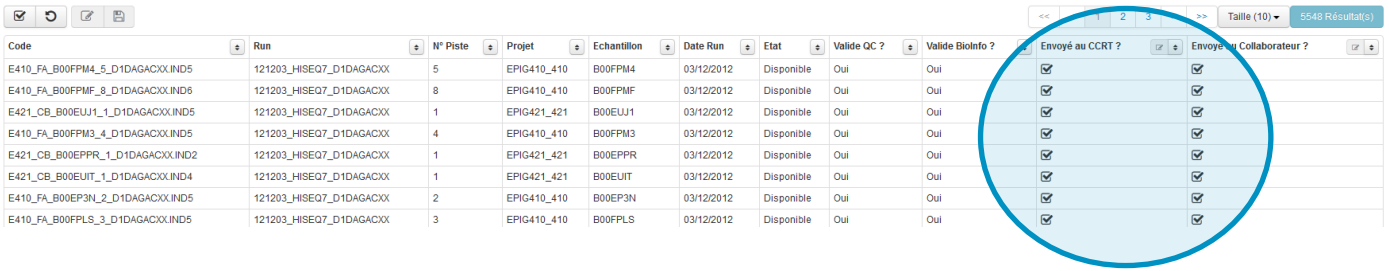
- Via l'interface individuelle du readset, onglet avancé :

- Via le menu Readsets > Modification en masse



La modification en masse des readsets concerne les champs :

- **Envoyé au CCRT ?**
- **Envoyé au collaborateur ?**



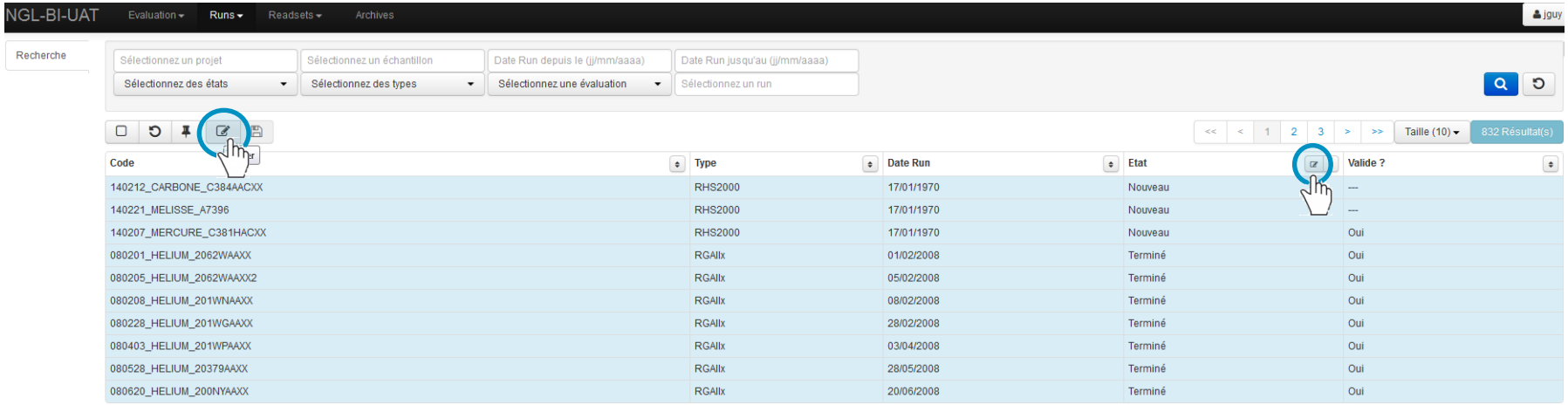
Changer l’état d’un run, readset ou analyse

(se référer au workflow)

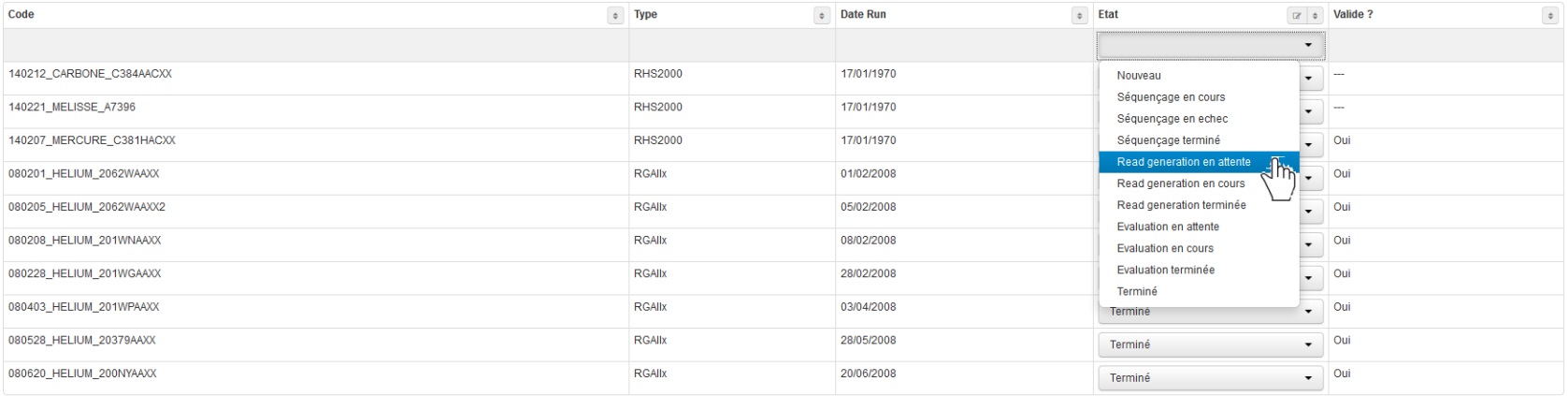
CHANGER L' ÉTAT D'UN RUN
(POSSIBLE EN MASSE)

Pour changer l'état d'un run, aller dans le menu **Runs > Changement d'état.**

- Filtrer le(s) run(s) à changer d'état et les sélectionner. Cliquer sur éditer (2 endroits possibles).



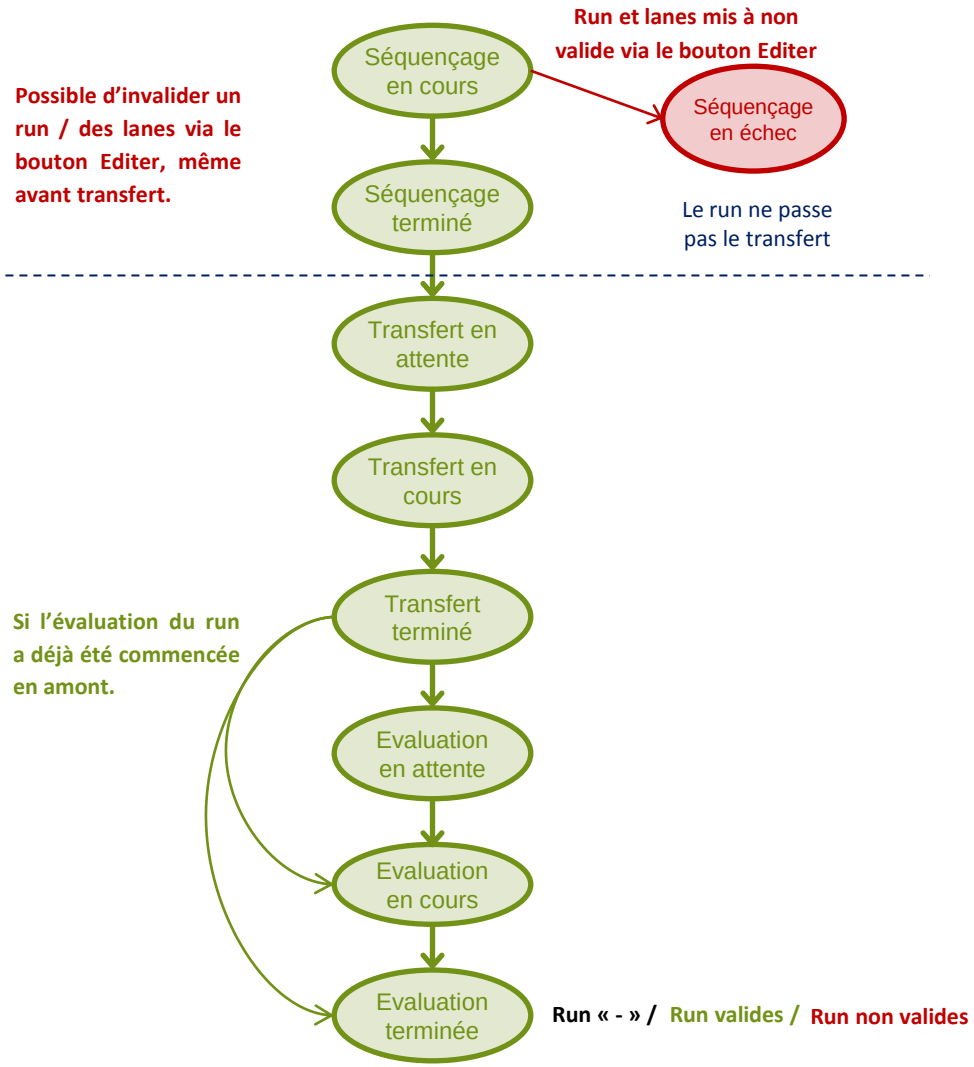
- Choisir le nouvel état.



- Enregistrer.

CAS PARTICULIER DU CHANGEMENT D'ÉTAT RUN À :
SÉQUENÇAGE EN ÉCHEC

Si un run plante en cours de séquençage et qu'il ne passe pas le transfert, son état restera à « séquençage en cours » ; il est alors recommandé d'aller changer manuellement son état à **« séquençage en échec ».**



CAS PARTICULIER DU CHANGEMENT D'ÉTAT RUN À :
SÉQUENÇAGE EN ÉCHEC

NGL-BI-UAT | Evaluation | Runs | Readsets | Archives | jguy

Recherche

140225_HISEQ3_C3FFJACXX

140225_HISEQ3_C3FFJACXX Séquençage en echec

Editer

Code	140225_HISEQ3_C3FFJACXX	Type	RHS2000	Nb Cycles		Version RTA	
Etat	Séquençage en echec	Nb Clusters (total)		Ligne Contrôle		Version Flowcell	
Valide ?	Non	% Clusters filt. (moyenne)		Code Flowcell	C3FFJACXX	Code Instrument	HISEQ3
Comptes Rendus	Problème machine laser camera	Nb Clusters filt. (total)		Position Flowcell		Type d'Instrument	HISEQ2000
		Nb Bases (total)					
Critères	Default	A conserver ?	☐				
Evalué par	jguy (05/03/2014)						
Détails évaluation	idéal : importer SAV au moment où on arrête le run						

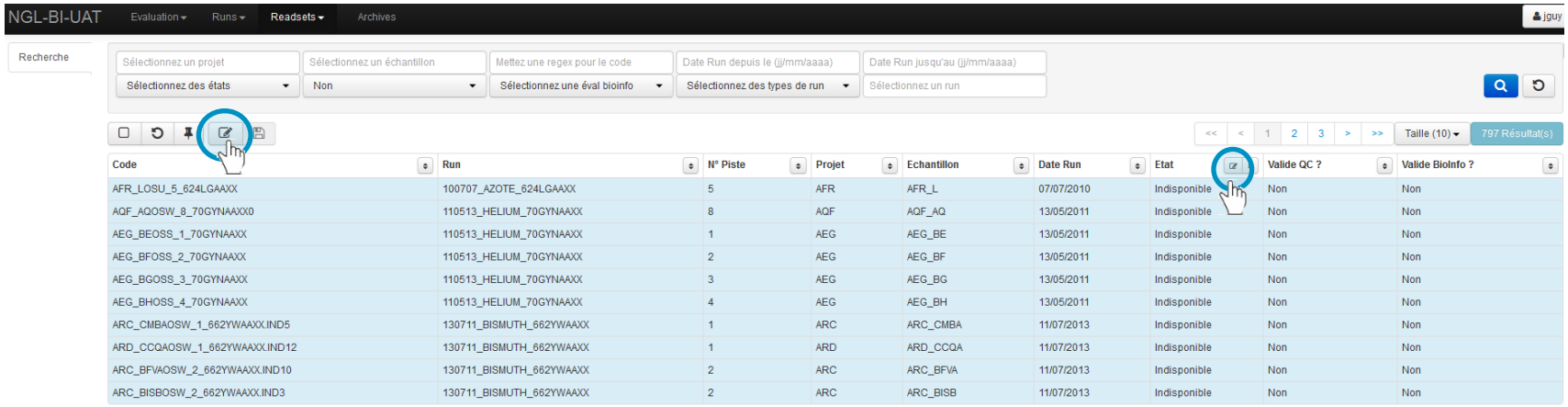
Pistes

#	Valide ?	Comptes Rendus
1	Non	Problème machine laser camera
2	Non	Problème machine laser camera
3	Non	Problème machine laser camera
4	Non	Problème machine laser camera
5	Non	Problème machine laser camera
6	Non	Problème machine laser camera
7	Non	Problème machine laser camera
8	Non	Problème machine

CHANGER L'ÉTAT D'UN READSET
(POSSIBLE EN MASSE)

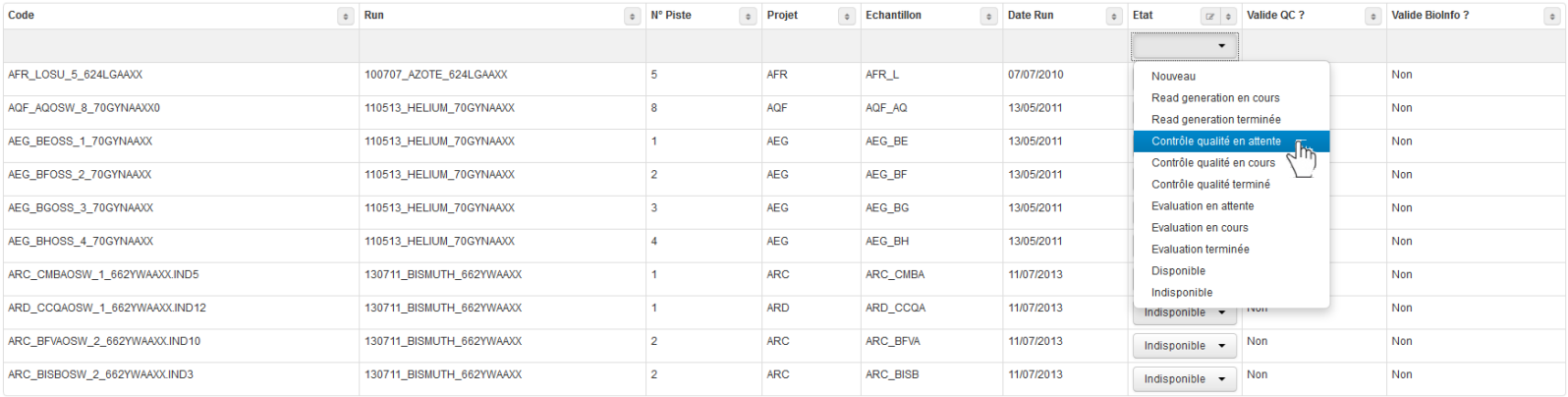
Pour changer l'état d'un readset, aller dans le menu **Readsets > Changement d'état.**

- Filtrer le(s) readset(s) à changer d'état et les sélectionner. Cliquer sur éditer (2 endroits possibles).



Code	Run	N° Piste	Projet	Echantillon	Date Run	Etat	Valide QC ?	Valide BioInfo ?
AFR_LOSU_5_624LGAAXX	100707_AZOTE_624LGAAXX	5	AFR	AFR_L	07/07/2010	Indisponible	Non	Non
AQF_AQOSW_8_70GYNAAXX0	110513_HELIUM_70GYNAAXX	8	AQF	AQF_AQ	13/05/2011	Indisponible	Non	Non
AEG_BEOSS_1_70GYNAAXX	110513_HELIUM_70GYNAAXX	1	AEG	AEG_BE	13/05/2011	Indisponible	Non	Non
AEG_BFOSS_2_70GYNAAXX	110513_HELIUM_70GYNAAXX	2	AEG	AEG_BF	13/05/2011	Indisponible	Non	Non
AEG_BGOSS_3_70GYNAAXX	110513_HELIUM_70GYNAAXX	3	AEG	AEG_BG	13/05/2011	Indisponible	Non	Non
AEG_BHOSS_4_70GYNAAXX	110513_HELIUM_70GYNAAXX	4	AEG	AEG_BH	13/05/2011	Indisponible	Non	Non
ARC_CMBAOSW_1_662YWAAXX.IND5	130711_BISMUTH_662YWAAXX	1	ARC	ARC_CMBA	11/07/2013	Indisponible	Non	Non
ARD_CCQAOSW_1_662YWAAXX.IND12	130711_BISMUTH_662YWAAXX	1	ARD	ARD_CCQA	11/07/2013	Indisponible	Non	Non
ARC_BFVAOSW_2_662YWAAXX.IND10	130711_BISMUTH_662YWAAXX	2	ARC	ARC_BFVA	11/07/2013	Indisponible	Non	Non
ARC_BISBOSW_2_662YWAAXX.IND3	130711_BISMUTH_662YWAAXX	2	ARC	ARC_BISB	11/07/2013	Indisponible	Non	Non

- Choisir le nouvel état.

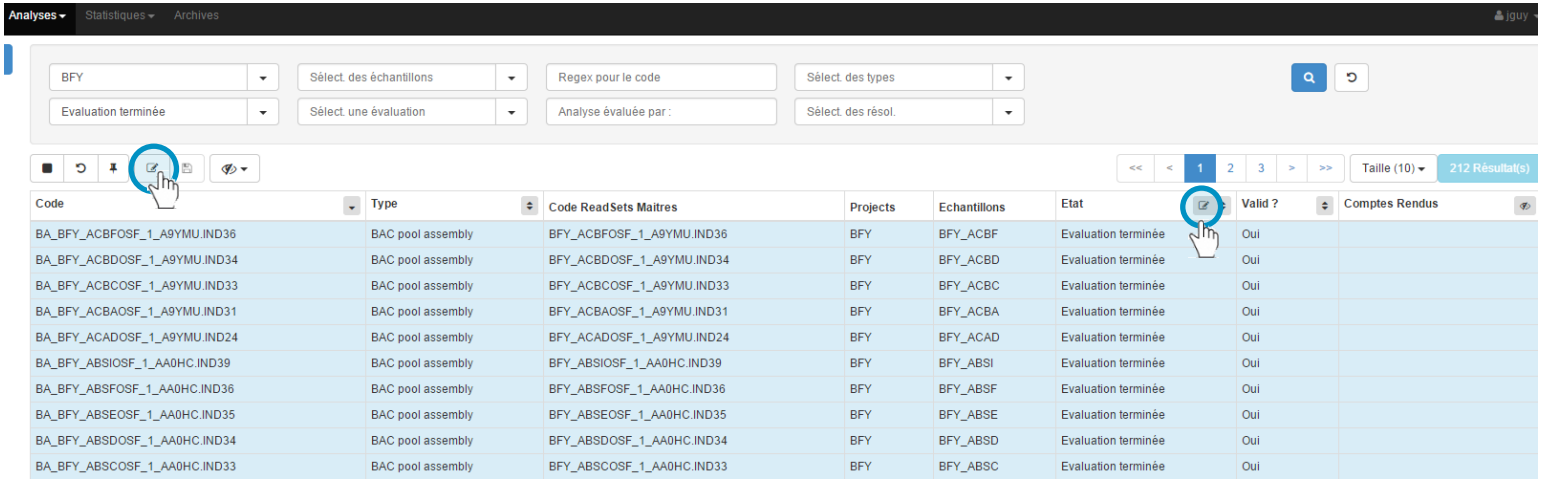


- Enregistrer.

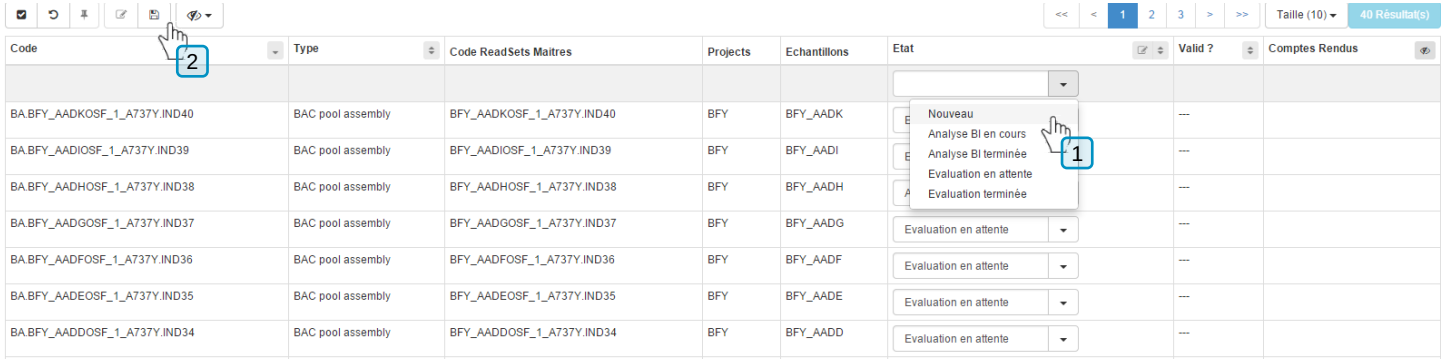
CHANGER L'ÉTAT D'UNE ANALYSE
(POSSIBLE EN MASSE)

Pour changer l'état d'une analyse, aller dans le menu **Analyse > Changement d'état.**

- Filtrer la ou les analyses à changer d'état et les sélectionner. Cliquer sur éditer (2 endroits possibles).



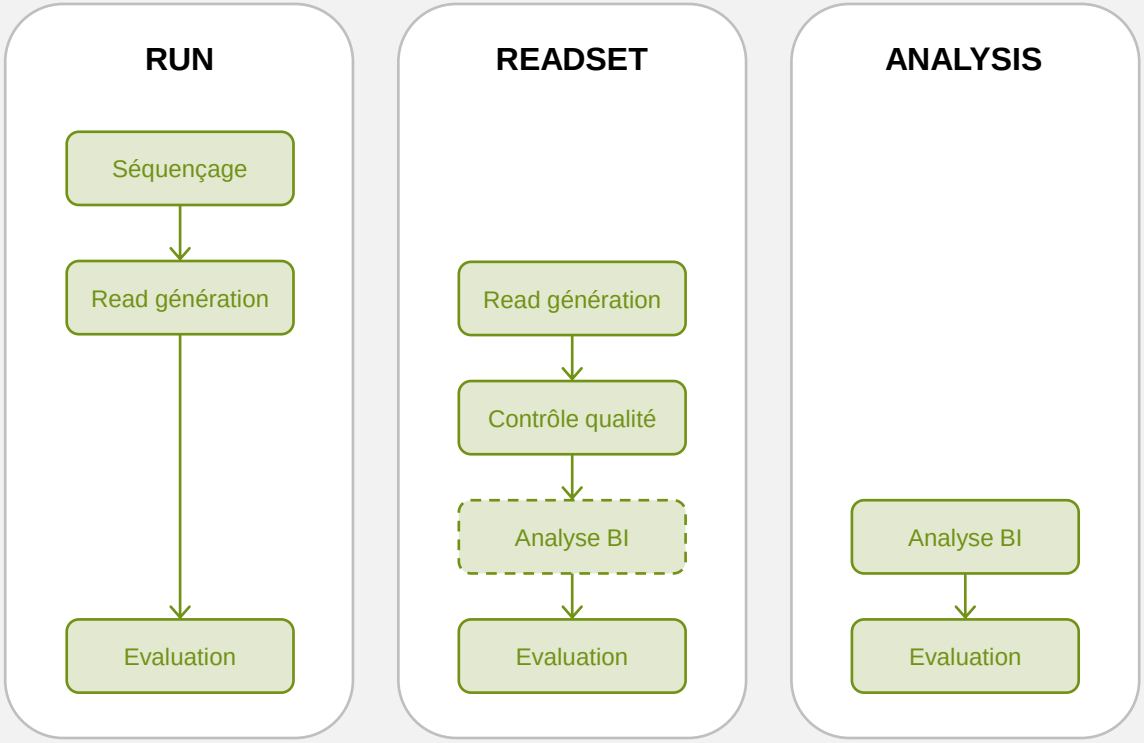
- Choisir le nouvel état (bandeau du haut pour propager l'état à toutes les analyses sélectionnées).



- Enregistrer.

Workflow RUN, READSET et ANALYSES

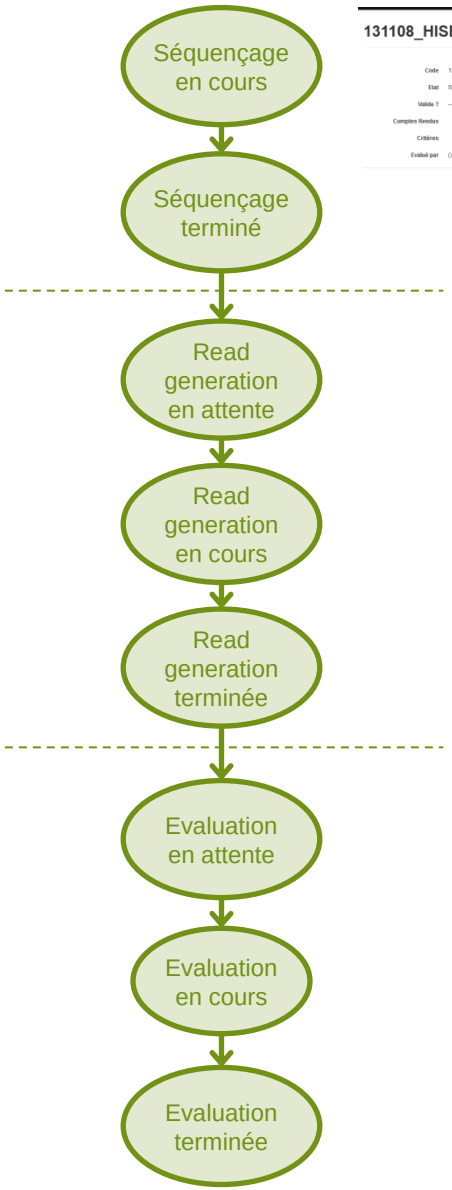
WORKFLOWS SIMPLIFIÉS

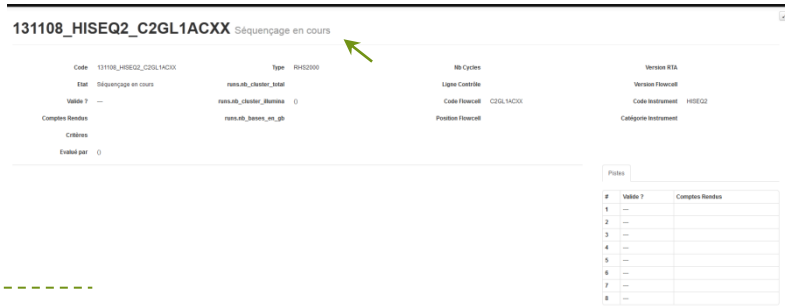


Workflow RUN et vues détaillées des interfaces

WORKFLOW RUN

1. **Séquençage en cours : création automatique du run**
   → interface du run visible dans NGL_BI.





1. RUN EN COURS DE SÉQUENÇAGE

NGL-BI | Evaluation | Runs | Readsets | Archives

Rechercher

131108_HISEQ2_C2GL1ACXX

131108_HISEQ2_C2GL1ACXX Séquençage en cours

Code	131108_HISEQ2_C2GL1ACXX	Type	RHS2000	Nb Cycles		Version RTA	
Etat	Séquençage en cours	Nb clusters (total)		Ligne Contrôle		Version Flowcell	
Valide ?	---	Nb clusters filtrés (total)		Code Flowcell	C2GL1ACXX	Code Instrument	HISEQ2
Comptes Rendus		Nb bases (total)		Position Flowcell		Type instrument	HISEQ2000
Critères							
Evalué par	()						

Pistes

#	Valide ?	Comptes Rendus
1	---	
2	---	
3	---	
4	---	
5	---	
6	---	
7	---	
8	---	

ReadSets (0)

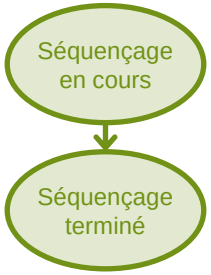
N° Piste	Code	% Séquences valides / piste	Nb Séquences Valides	Nb Bases	Q30	Score Qualité Moyen	Valide Prod ?	Valide BioInfo ?

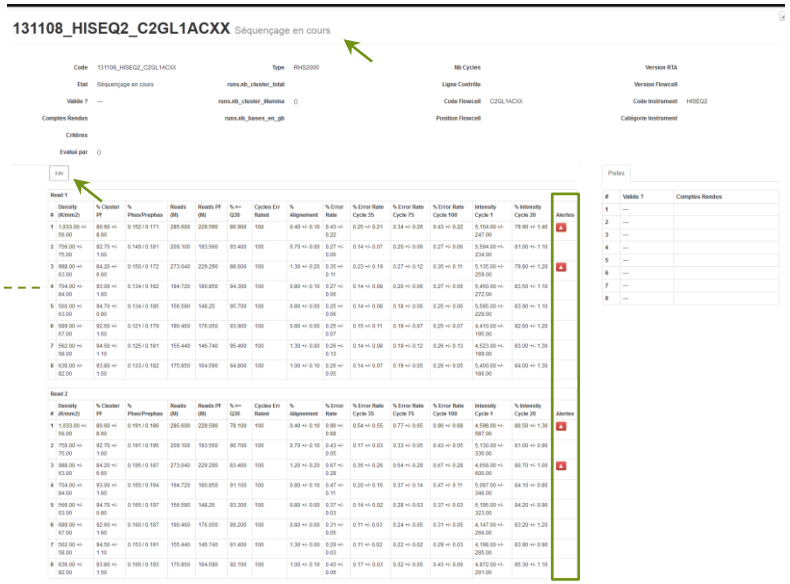
WORKFLOW RUN

1. **Séquençage en cours : création automatique du run**
   → interface du run visible dans NGL_BI.

2. **Import(s) SAV Summary Report en cours ou à la fin du run**
   → données visualisables dans l'interface du run.

   → Paramétrable : alerting SAV





2. DONNÉES SAV ET ALERTING

121203_HISEQ7_D1DAGACXX Séquençage en cours

Code	121203_HISEQ7_D1DAGACXX	Type	RHS2000	Nb Cycles	101	Version RTA	1.13.48
Etat	Séquençage en cours	Nb Clusters (total)	1,217,697,789	Ligne Contrôle	1	Version Flowcell	HiSeq Flow Cell v3
Valide ?	---	% Clusters filt. (moyenne)	93.93	Code Flowcell	D1DAGACXX	Code Instrument	HISEQ7
Comptes Rendus		Nb Bases (total)	1,143,757,649	Position Flowcell	A	Catégorie Instrument	
Critères		runs.nbBase	239,045,348,641				
Evalué par	()						

sav

Read 1

#	Density (K/mm2)	% Cluster PF	% Phas/Prephas	Reads (M)	Reads PF (M)	% >= Q30	Cycles Err Rated	% Alignement	% Error Rate	% Error Rate Cycle 35	% Error Rate Cycle 75	% Error Rate Cycle 100	Intensity Cycle 1	% Intensity Cycle 20	Alertes
1	1,033.00 +/- 56.00	80.60 +/- 8.60	0.152 / 0.171	285.600	229.590	86.900	100	0.40 +/- 0.10	0.43 +/- 0.22	0.25 +/- 0.21	0.34 +/- 0.26	0.43 +/- 0.22	5,154.00 +/- 247.00	79.90 +/- 1.40	
2	756.00 +/- 75.00	92.70 +/- 1.60	0.149 / 0.181	209.100	193.560	93.400	100	0.70 +/- 0.00	0.27 +/- 0.06	0.14 +/- 0.07	0.20 +/- 0.06	0.27 +/- 0.06	5,594.00 +/- 234.00	81.00 +/- 1.10	
3	988.00 +/- 63.00	84.20 +/- 6.60	0.150 / 0.172	273.040	229.290	88.600	100	1.30 +/- 0.20	0.35 +/- 0.11	0.23 +/- 0.19	0.27 +/- 0.12	0.35 +/- 0.11	5,135.00 +/- 259.00	79.80 +/- 1.20	
4	704.00 +/- 84.00	93.00 +/- 1.60	0.134 / 0.182	194.720	180.850	94.300	100	0.80 +/- 0.10	0.27 +/- 0.06	0.14 +/- 0.08	0.20 +/- 0.06	0.27 +/- 0.06	5,450.00 +/- 272.00	83.50 +/- 1.10	
5	566.00 +/- 63.00	94.70 +/- 0.80	0.134 / 0.185	156.590	148.25	95.700	100	0.80 +/- 0.00	0.25 +/- 0.06	0.14 +/- 0.08	0.18 +/- 0.06	0.25 +/- 0.06	5,585.00 +/- 229.00	83.90 +/- 1.10	
6	689.00 +/- 67.00	92.60 +/- 1.60	0.121 / 0.179	190.460	176.050	93.900	100	0.80 +/- 0.00	0.25 +/- 0.07	0.15 +/- 0.11	0.19 +/- 0.07	0.25 +/- 0.07	4,410.00 +/- 195.00	82.60 +/- 1.20	
7	562.00 +/- 58.00	94.50 +/- 1.10	0.125 / 0.181	155.440	146.740	95.400	100	1.30 +/- 0.00	0.26 +/- 0.13	0.14 +/- 0.08	0.19 +/- 0.12	0.26 +/- 0.13	4,523.00 +/- 189.00	83.00 +/- 1.30	
8	636.00 +/- 82.00	93.80 +/- 1.50	0.133 / 0.182	175.850	164.590	94.800	100	1.00 +/- 0.10	0.26 +/- 0.05	0.14 +/- 0.07	0.19 +/- 0.05	0.26 +/- 0.05	5,400.00 +/- 186.00	84.00 +/- 1.30	

Read 2

#	Density (K/mm2)	% Cluster PF	% Phas/Prephas	Reads (M)	Reads PF (M)	% >= Q30	Cycles Err Rated	% Alignement	% Error Rate	% Error Rate Cycle 35	% Error Rate Cycle 75	% Error Rate Cycle 100	Intensity Cycle 1	% Intensity Cycle 20	Alertes
1	1,033.00 +/- 56.00	80.60 +/- 8.60	0.191 / 0.186	285.600	229.590	78.100	100	0.40 +/- 0.10	0.90 +/- 0.68	0.54 +/- 0.55	0.77 +/- 0.65	0.90 +/- 0.68	4,598.00 +/- 587.00	80.50 +/- 1.30	
2	756.00 +/- 75.00	92.70 +/- 1.60	0.191 / 0.195	209.100	193.560	90.700	100	0.70 +/- 0.10	0.43 +/- 0.05	0.17 +/- 0.03	0.33 +/- 0.05	0.43 +/- 0.05	5,130.00 +/- 330.00	81.00 +/- 0.90	
3	988.00 +/- 63.00	84.20 +/- 6.60	0.195 / 0.187	273.040	229.290	83.400	100	1.20 +/- 0.20	0.67 +/- 0.28	0.35 +/- 0.26	0.54 +/- 0.29	0.67 +/- 0.28	4,658.00 +/- 600.00	80.70 +/- 1.00	

Alertes

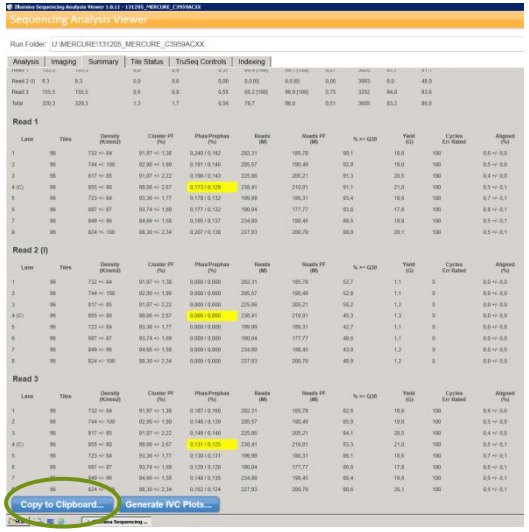
FLAG : Density (K/mm2),

Pistes

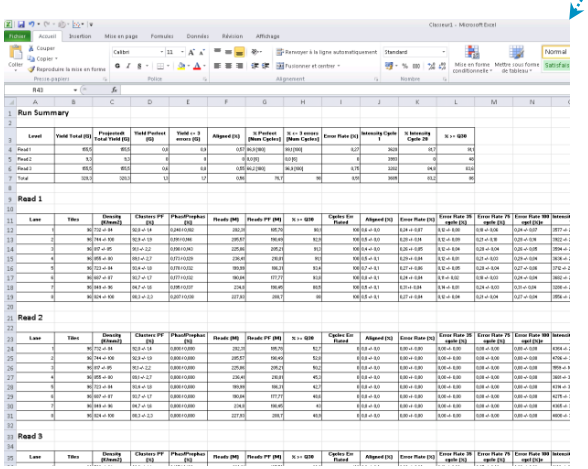
#	Valide ?	Comptes Rendus
1	---	
2	---	
3	---	
4	---	
8	---	

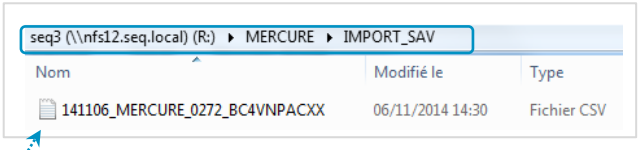
PROCÉDURE D'IMPORT DES DONNÉES SAV

- **Sur séquenceur, SAV (Illumina Sequencing Analysis Viewer)** > Copy to clipboard

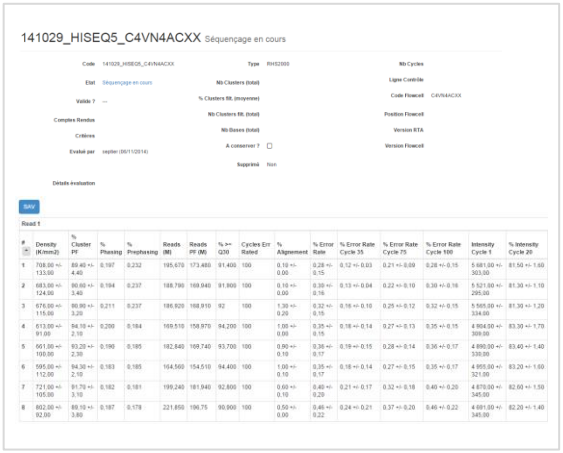


- **Aller dans le répertoire du séquenceur,** dans le dossier **IMPORT_SAV** : on y trouve un fichier CSV vide du nom du run. L'ouvrir.
- **Copier/coller dans ce fichier le summary report et enregistrer sous > CSV** (séparateur point virgule).





- **L'import se fait automatiquement dans NGL-BI.**



- **[disponible au CNG] Génération automatique d'alertes sur les données importées en cours de séquençage :**
  - Alertes envoyées par mail ;
  - Sur la base de seuils (BAD/FLAG) définis au préalable ;
  - Règles vérifiées à chaque import de summary report :
  - Notifications visibles dans l'interface du run, au niveau des lanes.

STRUCTURE DU MAIL D'ALERTE SAV

ALERTING SAV - Run 121023_HISEQ10_C177JACXX - Message (HTML)

De :
À : garnier@cng.fr ; ...
Cc :
Objet : ALERTING SAV - Run 121023_HISEQ10_C177JACXX
Date : mer. 18/12/2013 14:22

Tri par RUN → **Run 121023_HISEQ10_C177JACXX**

Visualisation des données SAV [BAD] et [FLAG]

Tri par READ → • **read1**

N° des lanes et cyclesErrRated →

Lane	BAD	FLAG
Number=2 cyclesErrRated=99	○ clusterDensity=1100 ○ clusterPFPerc=61.0	○ phasing=0.55 ○ prephasing=0.57 ○ errorRatePerc=0.73
Number=7 cyclesErrRated=99	○ greaterQ30Perc=70.0	
Number=8 cyclesErrRated=99		○ % aligned=0.3

• **read2**

Lane	BAD	FLAG
Number=1 cyclesErrRated=99	○ errorRatePercCycle75 =1.1	○ errorRatePercCycle35= 0.9
Number=2 cyclesErrRated=99	○ clusterDensity=1100 ○ clusterPFPerc=61.0	○ phasing=0.33 ○ prephasing=0.42 ○ errorRatePerc=0.91
Number=7 cyclesErrRated=99	○ greaterQ30Perc=55.0	
Number=8 cyclesErrRated=99		○ % aligned=0.3

WORKFLOW RUN

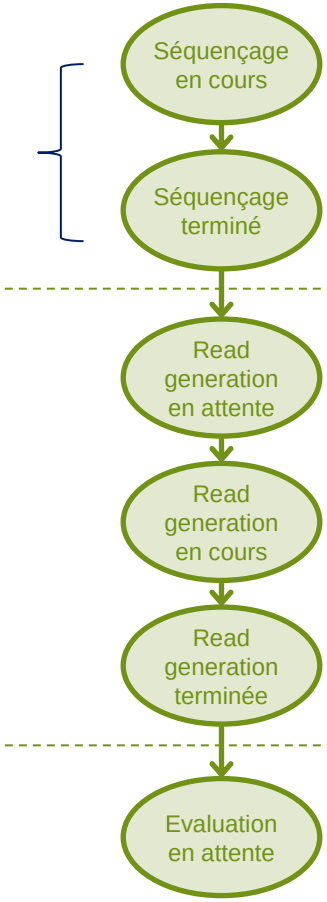
(écrans pris sur différents runs)

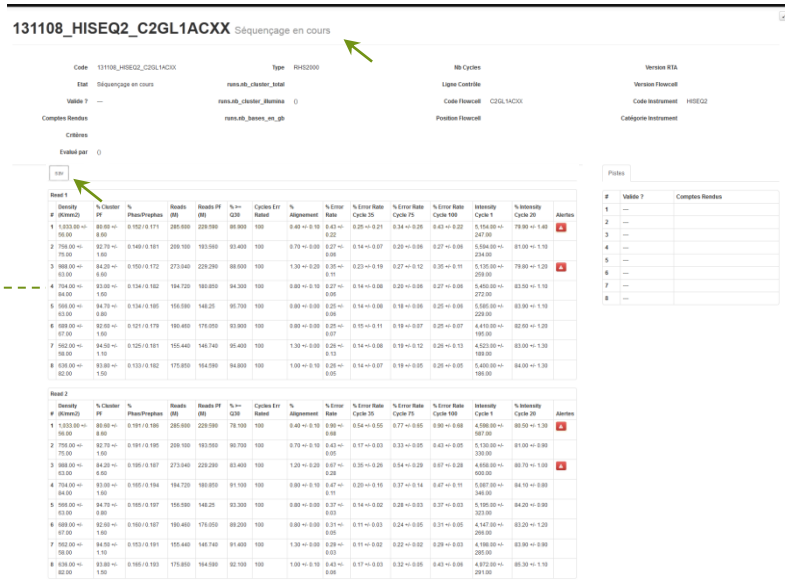
1. **Séquençage en cours : création automatique du run**
   → interface du run visible dans NGL_BI.

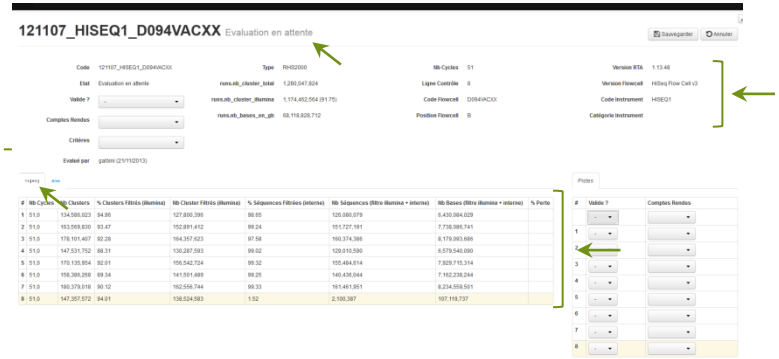
2. **Import(s) SAV Summary Report en cours ou à la fin du run**
   → données visualisables dans l'interface du run.

3. **Transfert terminé → données d'NGS_RG importées dans NGL_BI et visibles dans l'interface du run.**

4. **Passage automatique à « Evaluation en attente ».**







3. TRANSFERT TERMINÉ | 4. ÉVALUATION EN ATTENTE

NGL-BI | Evaluation | Runs | Readsets | Archives

Runs à Evaluer

121107_HISEQ1_D094VACXX

121107_HISEQ1_D094VACXX Evaluation en attente

Sauvegarder | Annuler

Code	121107_HISEQ1_D094VACXX	Type	RHS2000	Nb Cycles	51	Version RTA	1.13.48
Etat	Evaluation en attente	Nb clusters (total)	1,280,047,824	Ligne Contrôle	8	Version Flowcell	HiSeq Flow Cell v3
Valide ?	-	Nb clusters filtrés (total)	1,174,462,564 (91.75)	Code Flowcell	D094VACXX	Code Instrument	HISEQ1
Comptes Rendus		Nb bases (total)	68,118,828,712	Position Flowcell	B	Catégorie Instrument	
Critères							
Evalué par	galbini (21/11/2013)						

ngsrg | sav

#	Nb Cycles	Nb Clusters	% Clusters Filtrés (illumina)	Nb Cluster Filtrés (illumina)	% Séquences Filtrées (interne)	Nb Séquences (filtre illumina + interne)	Nb Bases (filtre illumina + interne)	% Perte
1	51,0	134,586,023	94.96	127,800,396	98.65	126,080,079	6,430,084,029	
2	51,0	163,569,830	93.47	152,891,412	99.24	151,727,191	7,738,086,741	
3	51,0	178,101,407	92.28	164,357,623	97.58	160,374,386	8,179,093,686	
4	51,0	147,531,752	88.31	130,287,593	99.02	129,010,590	6,579,540,090	
5	51,0	170,135,954	92.01	156,542,724	99.32	155,484,614	7,929,715,314	
6	51,0	158,386,268	89.34	141,501,489	99.25	140,436,044	7,162,238,244	
7	51,0	180,379,018	90.12	162,556,744	99.33	161,461,951	8,234,559,501	
8	51,0	147,357,572	94.01	138,524,583	1.52	2,100,387	107,119,737	

Pistes

#	Valide ?	Comptes Rendus
	-	
1	-	
	-	
3	-	
4	-	
5	-	
6	-	
7	-	
8	-	

ReadSets (10)

N° Piste	Code	% Séquences valides / piste	Nb Séquences Valides	Nb Bases	Q30	Score Qualité Moyen	Valide Prod ?	Valide BioInfo ?
1	E421_CB_B00EPPR_1_D1DAGACXX.IND2	37	66,141,643	13,360,611,886	92.7	36.15	---	---
1	E421_CB_B00EUIT_1_D1DAGACXX.IND4	38	66,712,820	13,475,989,640	92.81	36.2	---	---
1	E421_CB_B00EUJ1_1_D1DAGACXX.IND5	22	39,157,090	7,909,732,180	92.88	36.21	---	---
2	E410_FA_B00EP3N_2_D1DAGACXX.IND5	93	133,223,439	26,911,134,678	90.73	35.47	---	---
3	E410_FA_B00FPLS_3_D1DAGACXX.IND5	92	124,226,220	25,093,696,440	90.58	35.43	---	---
4	E410_FA_B00FPM3_4_D1DAGACXX.IND5	93	124,118,098	25,071,855,796	90.22	35.31	---	---
5	E410_FA_B00FPM4_5_D1DAGACXX.IND5	93	138,916,918	28,061,217,436	89.66	35.18	---	---
6	E410_FA_B00FPM5_6_D1DAGACXX.IND5	92	147,023,056	29,698,657,312	89.51	35.1	---	---
7	E410_FA_B00FPM6_7_D1DAGACXX.IND5	93	128,230,516	25,902,564,232	90.72	35.45	---	---
8	E410_FA_B00FPMF_8_D1DAGACXX.IND6	83	79,975,130	16,154,976,260	92.6	36.03	---	---

WORKFLOW RUN

(écrans pris sur différents runs)

1. **Séquençage en cours : création automatique du run**
   → interface du run visible dans NGL_BI.

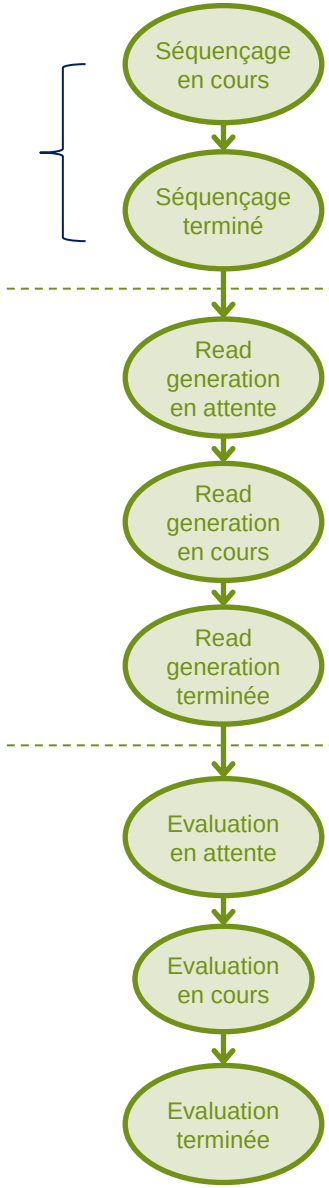
2. **Import(s) SAV Summary Report en cours ou à la fin du run**
   → données visualisables dans l'interface du run.

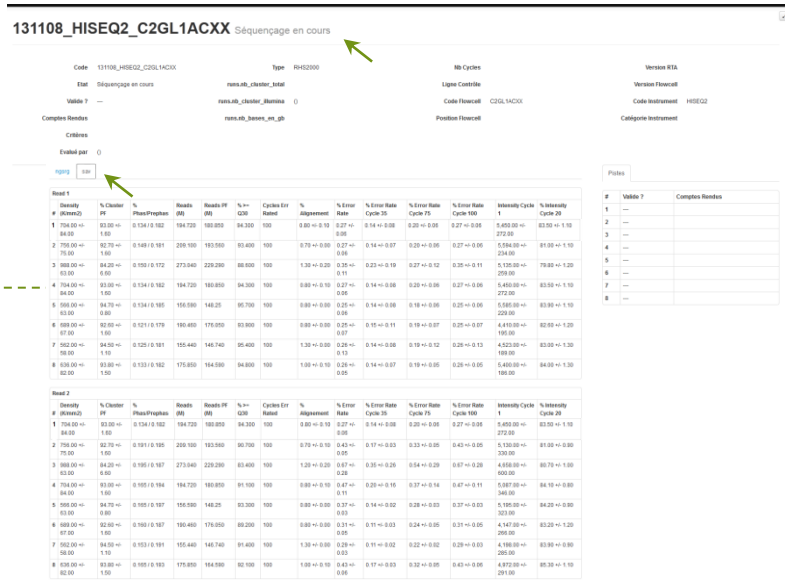
3. **Transfert terminé → données d'NGS_RG importées dans NGL_BI et visibles dans l'interface du run.**

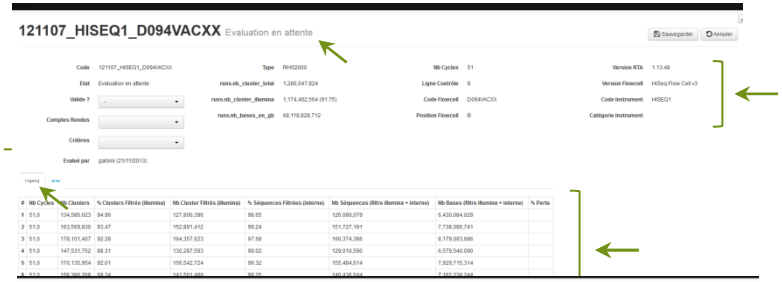
4. **Passage automatique à « Evaluation en attente ».**

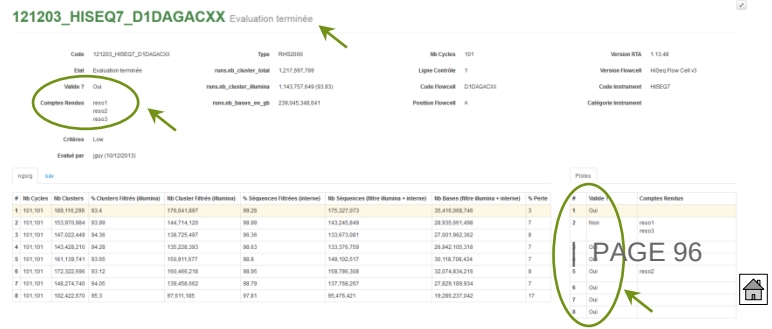
5. **Evaluation du run (et lanes)**

6. **Un run entièrement évalué (i.e. RUN + toutes les LANES) passe automatiquement à « évaluation terminée ».**









6. EVALUATION TERMINÉE

NGL-BI | Evaluation | Runs | Readsets | Archives

Runs à Evaluer

121203_HISEQ7_D1DAGACXX

121203_HISEQ7_D1DAGACXX Evaluation terminée

Sauvegarder | Annuler

Sauvegarde réussie.

Code	121203_HISEQ7_D1DAGACXX	Type	RHS2000	Nb Cycles	101	Version RTA	1.13.48
Etat	Evaluation terminée	runs.nb_cluster_total	1,217,697,789	Ligne Contrôle	1	Version Flowcell	HiSeq Flow Cell v3
Valide ?	Oui	runs.nb_cluster_illumina	1,143,757,649 (93.93)	Code Flowcell	D1DAGACXX	Code Instrument	HISEQ7
Comptes Rendus	reso1, reso2, reso3	runs.nb_bases_en_gb	239,045,348,641	Position Flowcell	A	Catégorie Instrument	
Critères	Low						
Evalué par	jguy (10/12/2013)						

ngsrg | sav

#	Nb Cycles	Nb Clusters	% Clusters Filtrés (illumina)	Nb Cluster Filtrés (illumina)	% Séquences Filtrées (interne)	Nb Séquences (filtre illumina + interne)	Nb Bases (filtre illumina + interne)	% Perte
1	101,101	189,116,299	93.4	176,641,897	99.26	175,327,073	35,416,068,746	3
2	101,101	153,970,984	93.99	144,714,120	98.99	143,245,849	28,935,661,498	7
3	101,101	147,022,449	94.36	138,725,497	96.36	133,673,081	27,001,962,362	8
4	101,101	143,428,210	94.28	135,228,393	98.63	133,376,759	26,942,105,318	7
5	101,101	161,139,741	93.65	150,911,677	98.8	149,102,517	30,118,708,434	7
6	101,101	172,322,696	93.12	160,466,218	98.95	158,786,308	32,074,834,216	8
7	101,101	148,274,740	94.05	139,458,662	98.79	137,768,267	27,829,189,934	7
8	101,101	102,422,670	95.3	97,611,185	97.81	95,476,421	19,286,237,042	17

Pistes

#	Valide ?	Comptes Rendus
	Oui	
1	Oui	
2	Non	reso1, reso3
3	Oui	
4	Oui	
5	Oui	reso2
6	Oui	
7	Oui	
8	Oui	

N° Piste	Code	% Séquences valides / piste	Nb Séquences Valides	Nb Bases	Q30	Score Qualité Moyen	Valide Prod ?	Valide BioInfo ?
1	E421_CB_B00EPPR_1_D1DAGACXX.IND2	37	66,141,643	13,360,611,886	92.7	36.15	---	---
1	E421_CB_B00EUIT_1_D1DAGACXX.IND4	38	66,712,820	13,475,989,640	92.81	36.2	---	---
1	E421_CB_B00EUJ1_1_D1DAGACXX.IND5	22	39,157,090	7,909,732,180	92.88	36.21	---	---
2	E410_FA_B00EP3N_2_D1DAGACXX.IND5	93	133,223,439	26,911,134,678	90.73	35.47	---	---
3	E410_FA_B00FPLS_3_D1DAGACXX.IND5	92	124,226,220	25,093,696,440	90.58	35.43	---	---
4	E410_FA_B00FPM3_4_D1DAGACXX.IND5	93	124,118,098	25,071,855,796	90.22	35.31	---	---
5	E410_FA_B00FPM4_5_D1DAGACXX.IND5	93	138,916,918	28,061,217,436	89.66	35.18	---	---
6	E410_FA_B00FPM5_6_D1DAGACXX.IND5	92	147,023,056	29,698,657,312	89.51	35.1	---	---
7	E410_FA_B00FPM6_7_D1DAGACXX.IND5	93	128,230,516	25,902,564,232	90.72	35.45	---	---
8	E410_FA_B00FPMF_8_D1DAGACXX.IND6	83	79,975,130	16,154,976,260	92.6	36.03	---	---

WORKFLOW RUN

(écrans pris sur différents runs)

1. **Séquençage en cours : création automatique du run**
   → interface du run visible dans NGL_BI.

2. **Import(s) SAV Summary Report en cours ou à la fin du run**
   → données visualisables dans l'interface du run.

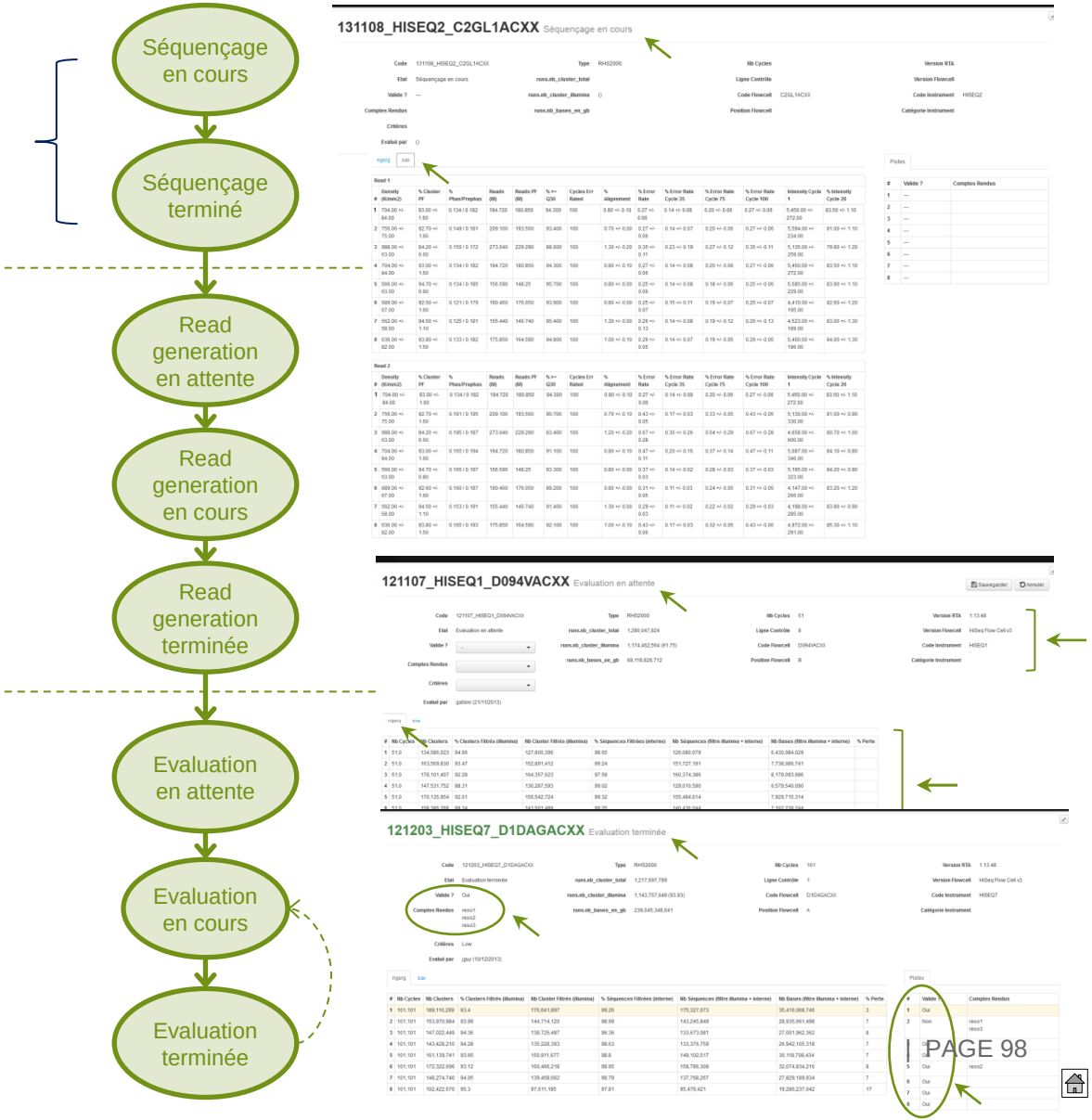
3. **Transfert terminé → données d'NGS_RG importées dans NGL_BI et visibles dans l'interface du run.**

4. **Passage automatique à « Evaluation en attente ».**

5. **Evaluation du run (et lanes)**

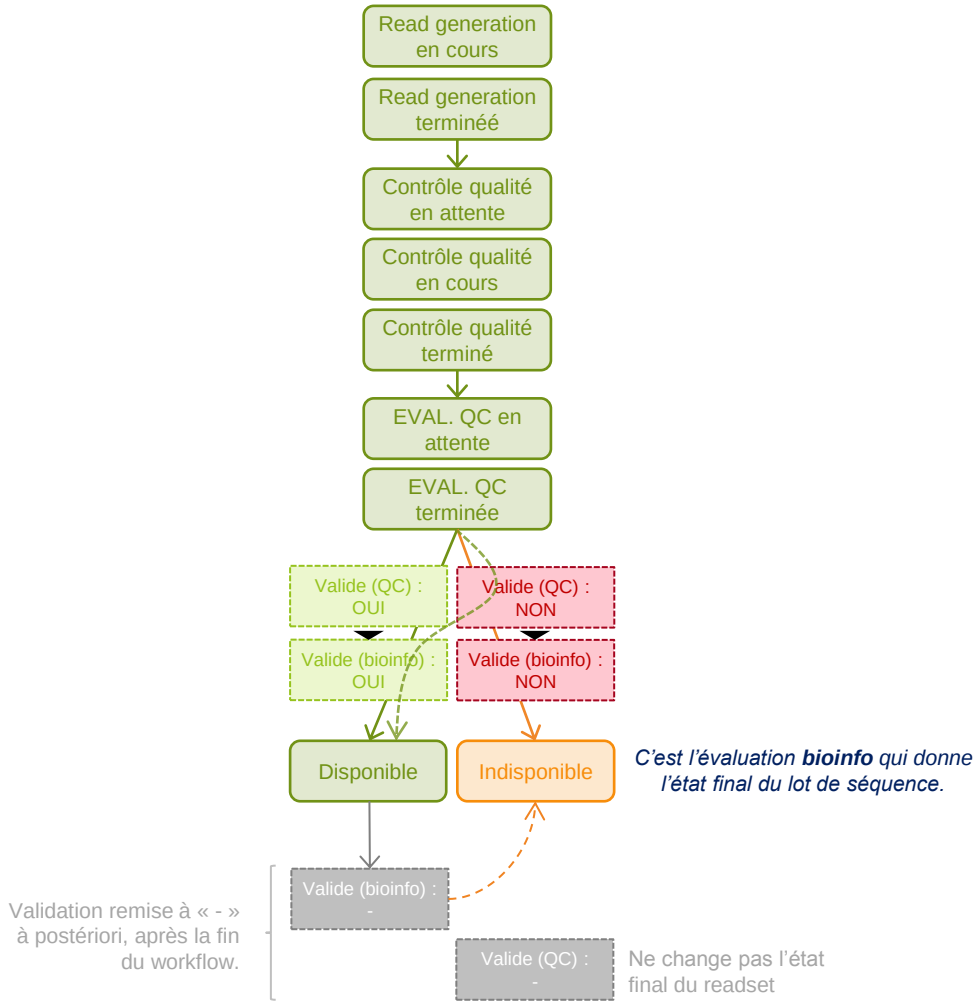
6. **Un run entièrement évalué (i.e. RUN + toutes les LANES) passe automatiquement à « évaluation terminée ».**

Rq : Un run à « évaluation terminée » repasse automatiquement à « évaluation en cours » si l'on remet à « - » une des évaluations (run ou lanes).



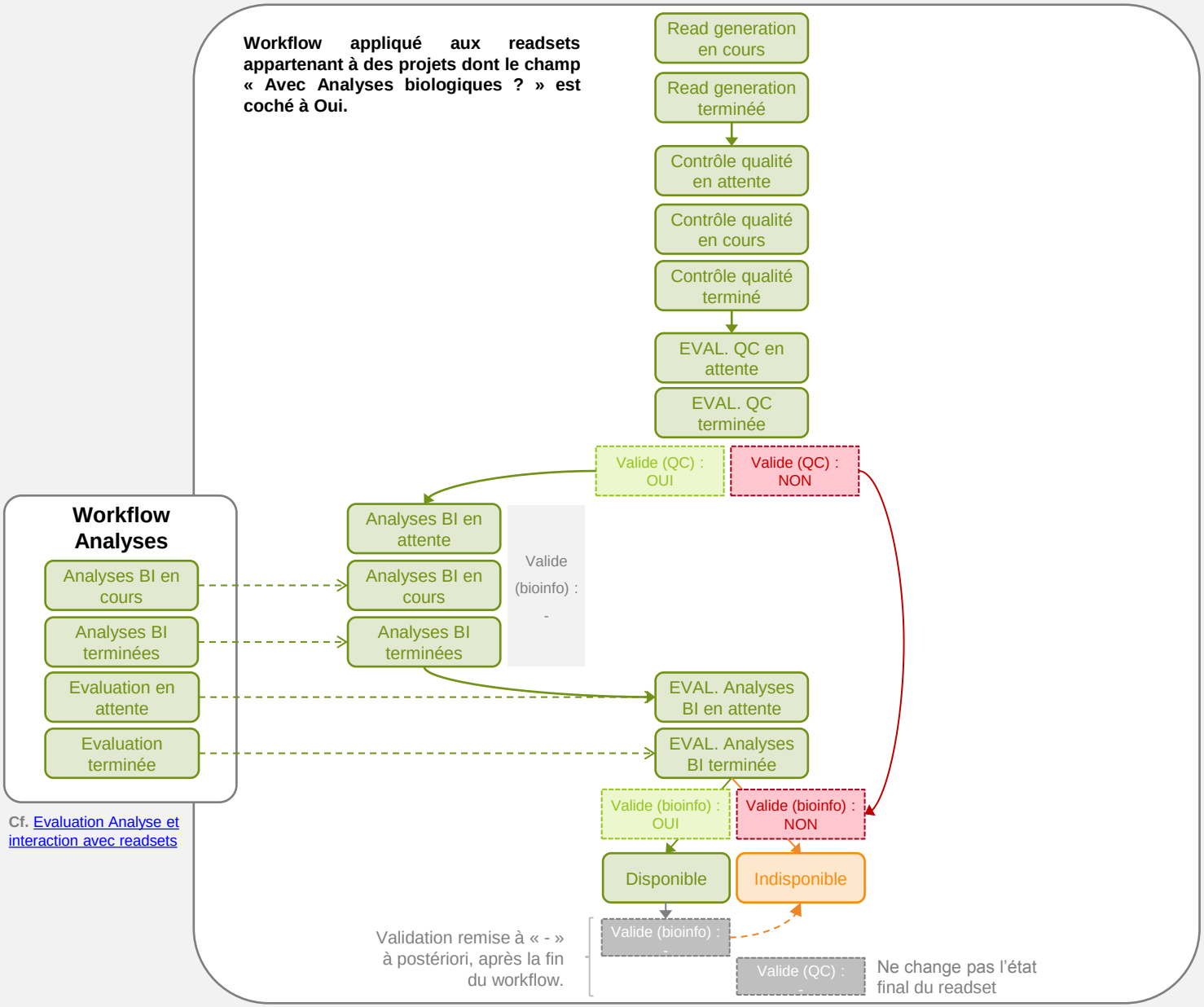
Workflow READSET

WORKFLOW READSET « CLASSIQUE »



WORKFLOW READSET AVEC ANALYSES BIOLOGIQUES

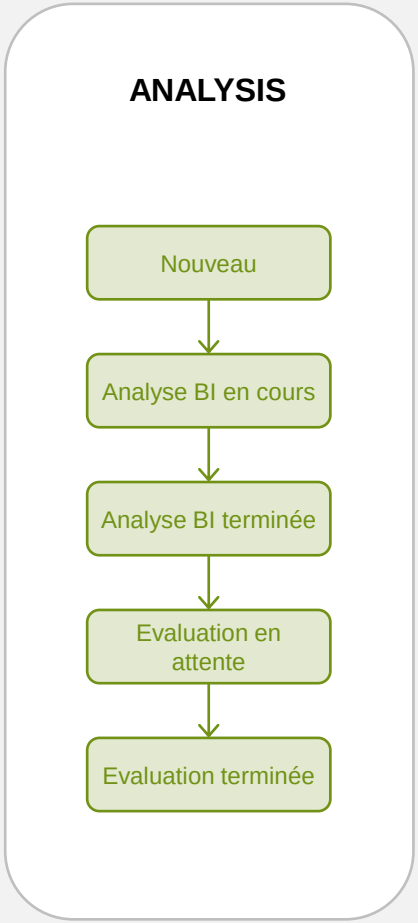
Workflow appliqué aux readsets appartenant à des projets dont le champ « Avec Analyses biologiques ? » est coché à Oui.



Cf. Evaluation Analyse et interaction avec readsets

Workflow ANALYSE

WORKFLOW ANALYSE



ANNEXES

LEXIQUE

- **Analyse :** objet créé à partir d'un ou plusieurs readsets, sur lequel des traitements peuvent être effectués, exemple : analyses blé.
  Pour certains projets, le contrôle qualité « classique » n'est pas suffisant pour évaluer (éval. bioinfo) les readsets. Sur ces projets, tagués au préalable comme devant comporter des analyses biologiques, vont être générées des « analyses », effectuées à partir d'un ou plusieurs readsets.
- **Configuration (de reporting) :** format de tableau donné, avec des colonnes supplémentaires par rapport au tableau par défaut, *exemples : configuration blé, configuration tara, configuration Analyse Blé.*
- **Critères d'évaluation :** « protocole » utilisé pour l'évaluation d'un run, d'un readset ou d'une analyse, *exemples : critères run RHS2000 2x100, critère readsets blé PE masse…* Les documents ValuationCriteria enregistrés dans la base listent l'ensemble des statistiques à tester avec leurs seuils limites et permettent d'afficher des highlights.
- **FC** : abréviation de Flowcell.
- **Highlight :** mise en couleur de la cellule d'une statistique au moment où un critère d'évaluation est sélectionné, *exemple : vert (OK) si < X%, orange (warning) si entre X et Y%, rouge (BAD) si >Y%,* à condition que la statistique soit listée dans un document ValuationCriteria avec une règle de comparaison par rapport à un ou plusieurs seuils limites. Une fois le critère d'évaluation enregistré, l'highlight reste visible à tout moment.
- **Lane** : piste d'une flowcell.
- **NGL_BI** = NGL_Bio-Informatique : gestion des analyses primaires de Bio-Informatique. Interfaces de suivi et d'évaluation des runs, readsets et analyses ; visualisation des statistiques issues des pipelines ngs-rg, ngs-qc et ngs-ba.
- **Readset** = lot de séquence ; correspond à un sample ayant subi une prep. de banque, déposé sur une piste avec un index donné (ou non indexé).
- **Résolution** : correspond à un compte-rendu (d'abandon ou non ; peut être informatif) à notifier au niveau du run, de la lane, du readset ou de l'analyse.
- **Run :** expérience de séquençage sur N cycles *(ex : de 50 cycles à 2x300 cycles)*, qui correspond à 1 flowcell, SR *(1 read de n cycles)* ou PE *(2 reads de n cycles),* déposée sur un séquenceur et générant des readsets.
- **Traitement** (NGL) : un traitement bio-informatique génère des statistiques qui sont importées dans NGL_BI au niveau du run, des lanes, des readsets ou analyses ; il existe plusieurs catégories de traitements : séquençage / transfert / qualité / analyse biologique ; exemples :
  - **SAV** *(cat. séquençage)* : ce traitement permet d'importer les statistiques en cours de séquençage (ou à la fin) à partir du logiciel SAV (Sequencing Analysis Viewer) présent sur les séquenceurs ;
  - **NGS-RG** *(cat. transfert)* : correspond au transfert et démultiplexage ;
  - **Read Quality** *(cat. qualité)* : analyse qualité des readsets en terme de distribution nucléotides, taille des lectures, contamination adaptateurs…
  - **Assembly** *(cat. ba)* : statistiques d'assemblage effectué sur une analyse.
- **Workflow** : enchaînement des états pris par un run, readset ou analyse *(workflows différents qui peuvent interagir entre eux)* ;
  les états s'enchaînent automatiquement selon des règles métier définies ; certains changements manuels sont toutefois possibles.

CRITÈRES D'ÉVALUATION DÉFINISSANT LES HIGHLIGHTS

Lien html (= nom du document) actif en mode diaporama : cliquer sur le lien => ouverture d'une page Alfresco => visualiser dans le navigateur ou télécharger le document.

- **RUN :**
  - CNS_Highlight_RUN.xlsx (NGS-RG et SAV) : regroupe les **ValuationCriteria** suivants :
    - RHS2000 2x101
    - RHS2500 2x101
    - RHS2500R 2x101
    - RHS2500R 2x151
    - MISEQ 2x101
    - MISEQ 2x151
    - MISEQ 2x251
    - MISEQ 2x301
  - CNG_critère évaluation_highlight_run.xlsx : regroupe les **ValuationCriteria** suivants :
    - RHS2000 2x101
    - RHS2500 2x101
- **READSET :**
  - CNS_Highlight_READSET masse_blé.xlsx : regroupe les **ValuationCriteria** suivants :
    - critères éval. blé MP masse
    - critères éval. blé PE masse
  - CNS_ValuationCriteria_RNAseq.xlsx
  - CNS_ValuationCriteria_PhyloAlps.xlsx
- **ANALYSE :**
  - CNS_Highlight_ANALYSES_blé.xlsx : regroupe les **ValuationCriteria** suivants :
    - critères éval. Analyse blé

REPORTING CONFIGURATION

Lien html (= nom du document) actif en mode diaporama : cliquer sur le lien => ouverture d'une page Alfresco => visualiser dans le navigateur ou télécharger le document.

- **REPORTING CONFIGURATION :**
  - CNS - Config report_readsets_tara.xlsx
  - CNS - Config report_readsets_blé.xlsx
  - Config report_readsets_RNAseq.xlsx
  - Config report_readsets_PhyloAlps.xlsx
  - CNS - Config report_analyses_blé.xlsx

COLONNES SUPPLÉMENTAIRES ET COMPORTEMENT SI LA FONCTIONNALITÉ « GROUPE » DU TABLEAU EST UTILISÉE

Lien html (= nom du document) actif en mode diaporama : cliquer sur le lien => ouverture d'une page Alfresco => visualiser dans le navigateur ou télécharger le document.

- **Colonnes supplémentaires :**
  - Liste des statistiques_mode reporting_mode graphique.xlsx : regroupe l'ensemble des propriétés disponibles dans « Colonnes supplémentaires » et le comportement (Somme, moyenne, …) à adopter lors de l'utilisation de la fonctionnalité GROUPE dans le tableau résultat.

WORKFLOW READSET – CHANGEMENTS MANUELS À FAIRE (CNG)

Lancement QC wrapper **automatique**
Lancement Pipeline **manuel**

Changement automatique

Changement manuel

Read generation en cours

Read generation terminéé

QC en attente

QC en cours

QC wrapper récupère les readsets à QC en attente et met à jour le workflow : QC en cours, QC terminé.

QC wrapper

QC terminé

Evaluation en attente

Pipeline Exome / WG / ChiPseq / …

Evaluation en cours

Donner au pipeline qualité correspondant les readsets à « évaluation en attente » et mettre à jour l'état des readsets :

Aller sur Readsets > Changement d'état (en masse) :

Utiliser les filtres de recherche à disposition :

- Sélectionner un état : « évaluation en attente » ;
- Sélectionner éventuellement un projet, un experimental type (avec le regex), un run…, etc.

Sélectionner les readsets à changer d'état et cliquer sur éditer.

Mettre à l'état « évaluation en cours » et sauvegarder.

Valide QC ? OUI ou NON en masse

Aller sur Readsets > Evaluation (en masse) :

Utiliser les filtres de recherche à disposition :

- Sélectionner un état : « évaluation en cours » ;
- Sélectionner éventuellement un projet, un experimental type (avec le regex), un run…, etc.

Sélectionner les readsets à évaluer et cliquer sur éditer.

Mettre l'évaluation QC à Oui ou Non et sauvegarder.

Lancement QC wrapper **manuel**
Lancement Pipeline **manuel**

Read generation en cours

Read generation terminéé

QC en attente

QC en cours

Donner à QC wrapper tous les readsets à QC en attente et mettre à jour manuellement le workflow : QC en cours, QC terminé.

QC wrapper

QC terminé

Evaluation en attente

Pipeline Exome / WG / ChiPseq / …

Evaluation en cours

Donner au pipeline qualité correspondant les readsets à « évaluation en attente » et mettre à jour l'état des readsets.

Valide QC ? OUI ou NON en masse

Evaluer en masse les readsets à « évaluation en cours » (sortis du pipeline qualité)