



Lettre d'information de SIGENAE



N°29

du 12 mai 2010 – Rédacteur : CK

Cette note est destinée aux membres des équipes de recherche utilisateurs des outils et services mis à disposition par la plate-forme SIGENAE. Elle a pour but de vous informer sur les évolutions de l'équipe, les nouveaux outils, services et formations mis en place.

Évolutions de l'équipe :

Gaëlle Rauffet et Jennifer Dupiot ont quitté l'équipe respectivement les 15 janvier et 28 février 2010.

Appels à projets 2010 :

L'arrivée des nouvelles technologies de séquençage à fortement fait augmenter le nombre de demandes d'accompagnement formulées par les utilisateurs auprès de la plate-forme.

Le tableau ci-dessous indique le nombre de demandes par type d'appel à projet. Dans les années passées la plate-forme était en moyenne sollicitée par 7 ou 8 équipes. Cette années 30 demandes d'accompagnement ont été enregistrées dont 26 sur les traitement de séquences issues des NGS.

Appel à projet	demandes
AIP	10
ANR Blanc	5
ANR Taïwan	1
ANR Génomique	8
AAP Génétique animale	5
ACI Phase	1

En fait, plusieurs tendances sont combinées dans cette évolution ; une augmentation du nombre d'équipes utilisant le séquençage, une augmentation des volumes de données à traiter (certains projets prévoient de re-séquencer plusieurs dizaines de génomes) et une diversification du type d'analyse (analyse du transcriptôme, assemblage de génome, recherche de SNP par RRL, par AFLP, par capture, par Long Range PCR, analyse des petits ARN, ...).

Ces projets, vu leur nombre, la diversité des approches et les volumes générés ne pourront pas être conduits de front dans le périmètre actuel de la plate-forme. Une réflexion est en cours avec les chefs de département, le directeur d'AGENAE, la présidente du comité des utilisateurs et la responsable de la cellule bioinformatique de l'INRA pour proposer une organisation capable de relever ce défi.

Formations 2010 :

Sigenae organise conjointement avec la plate-forme Bio-info Genotoul une formation à l'alignement de séquences issues des NGS et à la recherche de polymorphismes. Cette formation a pour but d'aider les biologistes à traiter les séquences issues des NGS en particulier les plates-formes Roche 454 et Illumina Solexa.

Cette formation se compose de 4 modules comportant une grande variété d'exercices.

Horaires	Contenus
09h30 - 11h	Les séquences produites par les plate-formes de NGS Roche 454 et Solexa Illumina, leur caractéristiques, leurs biais, la qualité des séquences Récupération des séquences dans SRA (Short Read Archive) du NCBI Récupération de la séquences de référence

Horaires	Contenus
11h15 - 12h30	L'évolution des algorithmes d'alignement : Smith-Waterman / Blast / MAQ / bwa Les formats d'alignement sam et bam
14h - 15h30	La visualisation des alignements avec IGV
15h45 - 17h	La recherche de polymorphismes à partir de fichiers bam Le format pileup

Pré-requis : savoir utiliser un environnement Unix.

Cette formation est pour l'instant dispensée sur le site de Toulouse. Les trois sessions prévues sont complètes mais d'autres seront organisées en septembre. Elle sera ensuite étendue aux autres sites.

Pour nous faire part de votre intérêt pour cette formation ou pour nous transmettre d'autres besoins vous disposez de la page : <http://www.sigene.org/index.php?id=213>

Contigbrowser / DDDView (Digital Differential Display) :

Cette nouvelle fonctionnalité va être disponible sur le Sigene Contig Browser au fur et à mesure de la mise à jour des assemblages des différentes espèces (browsers public et privé). L'objectif de cet outil est de comparer le nombre de représentants de différentes librairies assignés à un contig Sigene particulier. Pour compenser l'hétérogénéité du nombre de séquences par librairie, le DDD utilise le [Test Exact de Fisher](#) comme test statistique de comparaison. L'idée sous-jacente est de déterminer quels sont les transcrits différentiellement exprimés entre deux tissus et/ou conditions physiologiques distincts.

La réalisation d'un Digital Differential Display se déroule en 2 étapes :

La première consiste à sélectionner les librairies à comparer en fonction des tissus et/ou conditions physiologiques à partir desquels elles ont été créées et ce en composant deux pools de librairies au moyen de cases à cocher. On passe à la seconde étape en cliquant sur le bouton « Continuer ».

Public Sigene Contig Browser Sheep DddView

Based on Ensembl release 40 - Aug 2006

Use Ensembl to... Search Contig Browser, Data mining (BioMart), Export data, Download data, Digital Differential Display

Docs and downloads: About Contig Browser, Contig Browser data

Other links: Home, Sigene, Ageneae

Sheep Digital Differential Display

DDD is a tool for comparing EST profiles in order to identify contigs with significantly different expression levels. Define two library pools by selecting one or more libraries in each +Lib. Pool+ column and click on Continue button.

Tissue	Lib. name	Dev. stage	Nb ESTs	Lib. Pool 1	Lib. Pool 2
Abomasal mucosa	KN511 Ovis aries abomasal mucosa (gastric fold) from Teladorsagia infected sheep	Unknown	1733	<input type="checkbox"/>	<input type="checkbox"/>
	KN511 Ovis aries abomasal mucosa (gastric fold) from (Teladorsagia) uninfected sheep	Unknown	1421	<input type="checkbox"/>	<input type="checkbox"/>
Corpus luteum	OD10L	Unknown	2086	<input type="checkbox"/>	<input type="checkbox"/>
	ONDC	Unknown	22828	<input type="checkbox"/>	<input type="checkbox"/>
Dendritic cells					
Efferent gastric lymph	KN511 Ovis aries efferent gastric lymph from uninfected sheep	Unknown	3107	<input type="checkbox"/>	<input type="checkbox"/>
	KN511 Ovis aries efferent gastric lymph from Teladorsagia infected sheep	Unknown	1796	<input type="checkbox"/>	<input type="checkbox"/>
Endometrium	POE14_Day_14_pregnant_ovine_endometrium	Day 14 pregnant	3388	<input type="checkbox"/>	<input type="checkbox"/>
Fetal ovary	Ovine foetal ovary tissues (Early folliculogenesis) - F (oca)	82 days of gestation	2551	<input checked="" type="checkbox"/>	<input type="checkbox"/>
	Ovine foetal ovary tissues (Prophase I meiosis) - M (oca)	55 days of gestation	2474	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Gall bladder	OMGB2	Unknown	2595	<input type="checkbox"/>	<input type="checkbox"/>
	OMGB1	Unknown	6510	<input type="checkbox"/>	<input type="checkbox"/>
Gastric lymph node	KN511 Ovis aries gastric lymph node from Teladorsagia infected sheep	Unknown	1096	<input type="checkbox"/>	<input type="checkbox"/>
Healing fracture callus	Adult sheep fracture callus 10d	Unknown	17243	<input type="checkbox"/>	<input type="checkbox"/>
	Adult sheep fracture callus 7d	Unknown	14926	<input type="checkbox"/>	<input type="checkbox"/>
	Adult sheep fracture callus pool	Unknown	15224	<input type="checkbox"/>	<input type="checkbox"/>
Hind limb	OM135	Unknown	2754	<input type="checkbox"/>	<input type="checkbox"/>

Le DDD calcule alors le test de comparaison pour chaque contig Sigene contenant des ESTs venant des deux pools de librairies et génère un tableau de résultats

ne regroupant que les contigs pour lesquels le test de comparaison est significatif. Sont indiqués les tailles de chaque pool ainsi que les effectifs présents dans le contig. Enfin, la couleur de ligne indique le pool de librairies statistiquement sur-représenté.

Public Sigene Contig Browser Sheep DddView

Based on Ensembl release 40 - Aug 2006

Use Ensembl to... Search Contig Browser, Data mining (BioMart), Export data, Download data, Digital Differential Display

Docs and downloads: About Contig Browser, Contig Browser data

Other links: Home, Sigene, Ageneae

Sheep Digital Differential Display

First library pool (FP): Ovine foetal ovary tissues (Early folliculogenesis) - F (oca) / Ovary (fetal) / 82 days of gestation / 2551 ESTs

Second library pool (SP): Ovine foetal ovary tissues (Prophase I meiosis) - M (oca) / Ovary (fetal) / 55 days of gestation / 2474 ESTs

DDD only reports on a contig if at least one pairwise comparison (FP/SP) results in a statistically significant difference. Each row in the results table represents a Sigene contig.

Sigene contig	Description	FP size	SP size	FP in contig	SP in contig	Fisher exact test
CD2827259.p.ov.8	(nc)P004151 Cyclochrome c oxidase subunit 3 OS=Hos taurus GN=MT-CO3 PE=1 SV=2	2546	2441	5	33	1.705464146907956-06
AF125041.p.ov.8	(BBH) P052_SHEEP (sp)Q12722) Deocinin OS=Ovis aries GN=DCN PE=2 SV=1	2532	2472	19	2	0.00022289293130419
CD2823187.p.ov.8	(BBH) H590A_BOVIN (sp)Q12762) Heat shock protein HSP 90-alpha OS=Hos taurus GN=HSP90A1 PE=2 SV=3	2550	2460	1	14	0.000429754903325413
CD288357.p.ov.8	(nc)Q37430) Cyclochrome c oxidase subunit 2 OS=Capra hircus GN=MT-CO2 PE=3 SV=1	2550	2483	1	11	0.00297652247567053
CD2826005.p.ov.8	(BBH) RL27A_BOVIN (sp)Q29093) 60S ribosomal protein L27a OS=Hos taurus GN=RL27A PE=2 SV=3	2539	2473	12	1	0.00342467625511838
CD2804617.p.ov.8	(BBH) OBP_BOVIN (sp)Q12143) Quinine nucleoside-binding protein subunit beta-2-like 1 OS=Hos taurus GN=OBNB2L1 PE=2 SV=3	2535	2471	16	3	0.004437461810203073
BD974273.p.ov.8	(BBH) EF1AT1_RABIT (sp)P68105) Elongation factor 1-alpha 1 OS=Oryzotilapia auriculata GN=EEF1A1 PE=1 SV=1	2541	2468	10	26	0.00976920792305659
CD2806619.p.ov.8	(nc)Q21282) Glutathione S-transferase Mu 5 OS=Rattus norvegicus GN=Gstm5 PE=1 SV=3	2535	2470	16	4	0.0118465750627552
U08283.p.ov.8	(BBH) ACTG_XENLA (sp)A28080) Actin, cytoplasmic 2 OS=Macropus lewisii GN=actg1 PE=2 SV=1	2536	2470	15	4	0.019250058994021
CD2828987.p.ov.8	(BBH) RL2_BOVIN (sp)Q28090) 60S ribosomal protein L4 OS=Hos taurus GN=RL4 PE=2 SV=3	2550	2466	1	8	0.0196191597784135
CD2826917.p.ov.8	(BBH) HBPIC_BOVIN (sp)P19120) Heat shock cognate 71 kDa protein OS=Hos taurus GN=HSP71 PE=1 SV=2	2550	2466	1	8	0.0196191597784135
QV477339.p.ov.8	(BBH) RAB14_PONAB (sp)Q28282) Rab-related protein Rab-14 OS=Phago apheli GN=RAB14 PE=2 SV=3	2542	2473	9	1	0.0215562743140996
DV510854.p.ov.8	(BBH) STT3B_HUMAN (sp)Q12022) Deltalysin4-delta-4-epitope-associated-protein glycosyltransferase subunit STT3B OS=Homo sapiens GN=STT3B PE=1 SV=1	2540	2472	11	2	0.0024907174867946
DV514610.p.ov.8	PREDICTED: Bos taurus similar to RAB4A, member RAB oncogene family, transmembrane 2 (RAB4A), mRNA	2540	2463	3	11	0.03241730705124456
DV598734.p.ov.8	(BBH) TQPL_BOVIN (sp)Q23208) T-complex protein 1 subunit theta OS=Hos taurus GN=TCPT1 PE=2 SV=3	2550	2467	1	7	0.0362152369082299
EE772042.p.ov.8	PREDICTED: Meacoa mullata similar to zinc finger protein 395, transcript variant 5 (LOC113700), mRNA	2543	2473	8	1	0.0390926082297539
QV477335.p.ov.8	(BBH) ATX10_BOVIN (sp)Q21870) Ataxin-10 OS=Hos taurus GN=ATX10 PE=2 SV=1	2543	2473	8	1	0.0390926082297539
CD2827554.p.ov.8	(BBH) RL2_BOVIN (sp)Q28071) 60S ribosomal protein L7 OS=Hos taurus GN=RL7 PE=2 SV=1	2543	2473	8	1	0.0390926082297539
DE117513.p.ov.8	(BBH) SNU3_BOVIN (sp)Q232816) Sorting-nuclein-3 OS=Hos taurus GN=SNU3 PE=2 SV=1	2543	2473	8	1	0.0390926082297539