

STAGE M1 BIOINFORMATIQUE

LBBE UCBL

Juin 2024

Analyse des insertions d'éléments transposables
dans un génome viral par séquençage nanopore

Maitres de stage :

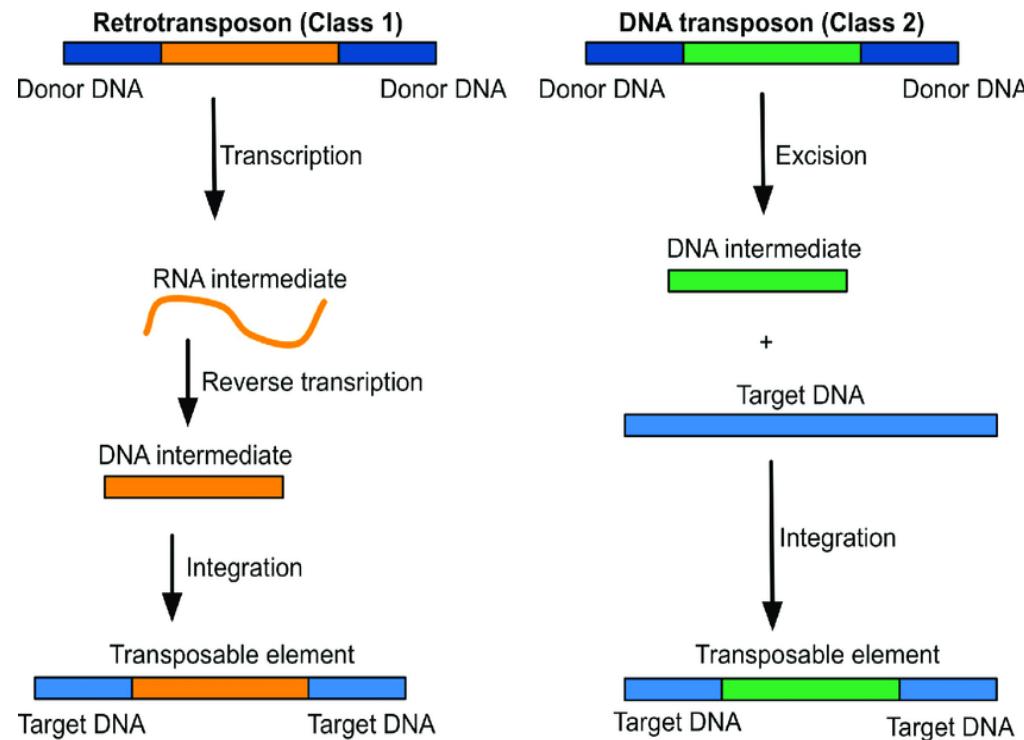
Vincent LACROIX (LBBE : équipe Baobab)

Marie FABLET(LBBE : équipe GEI)

Présenté par :

Tidjani FOUSSENI SALAMI CISSE

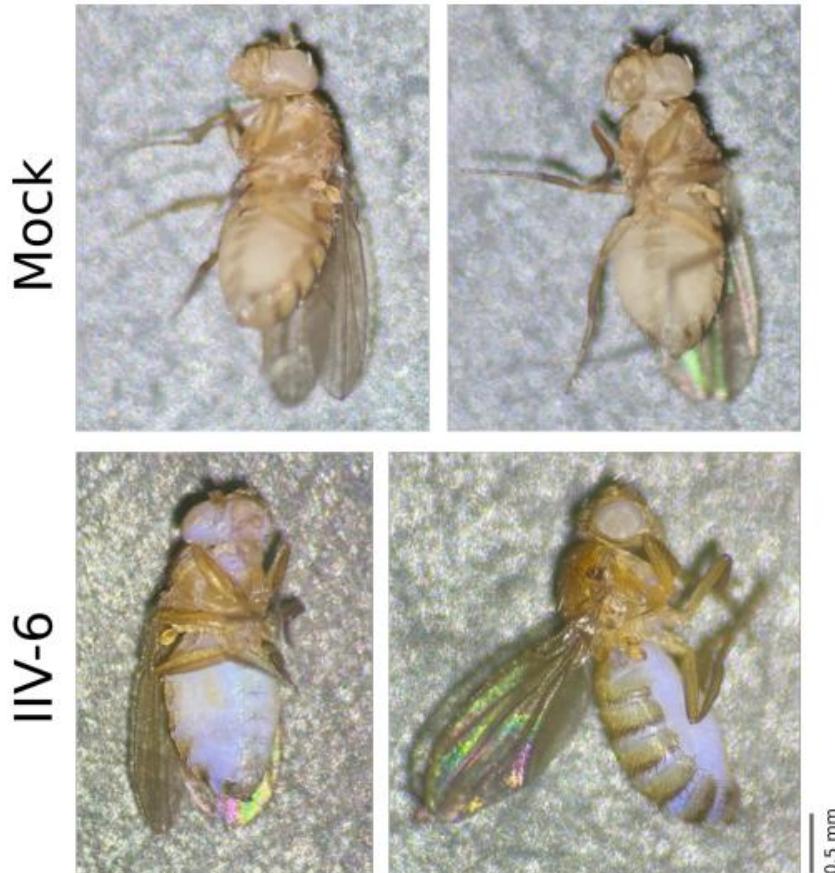
Introduction



Infection de *Drosophila* par le virus

- Augmentation transcript des ET chez Dmel.
- Virus à ADNdb.
- Présence des ET chez IIv6

Introduction



■ Deux hypothèses:

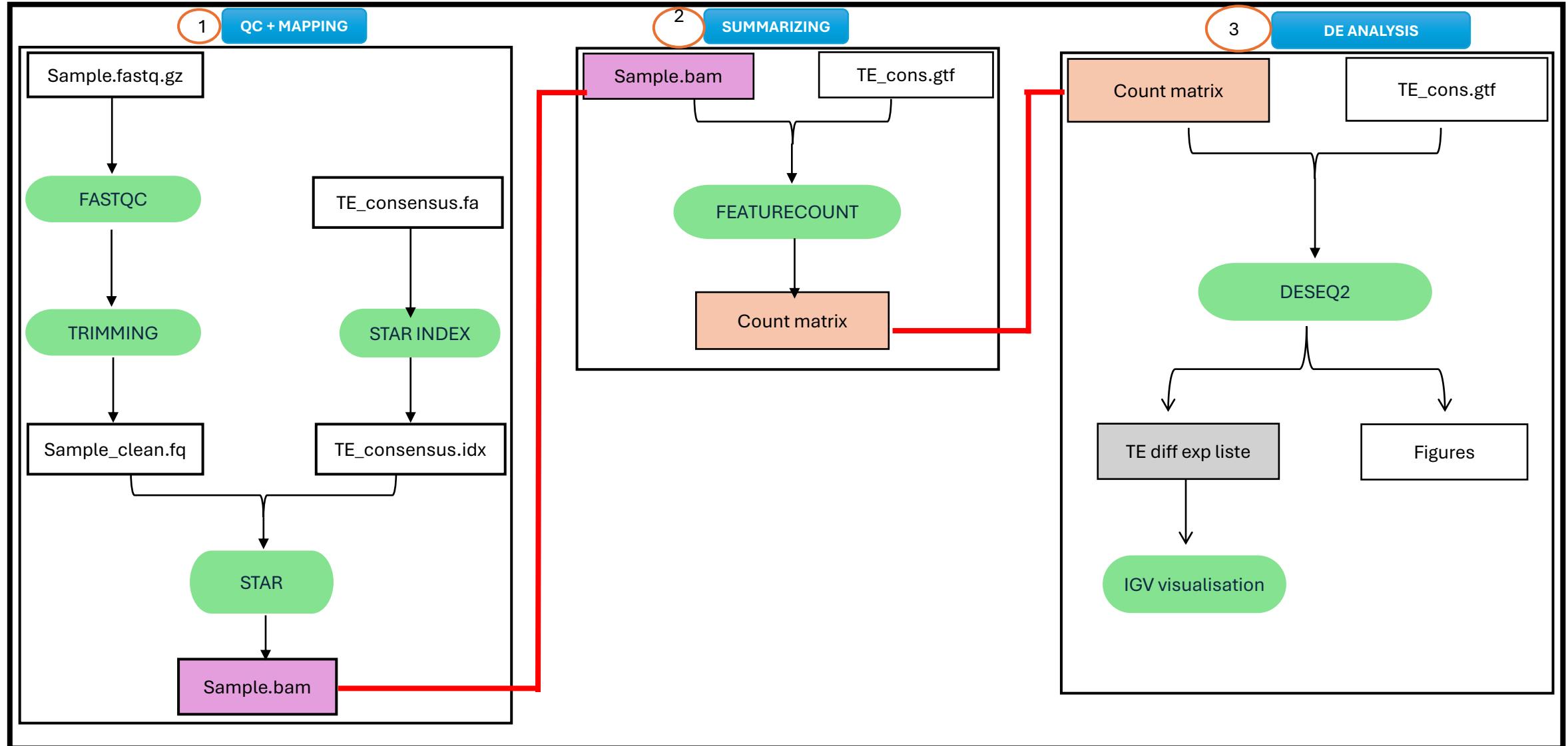
■ ET du IIV6 et de Dmel s'expriment ?

■ Surexpression des ET induite par infection virale ?

Les jeux de Données

RNaseq (short reads)				DNaseq(long reads)
Carcasses		Ovaires		Génome du virus IIv6
mock	Infectée par IIv6	mock	Infectée par IIv6	
R X 3	R X 3	R X 3	R X 3	

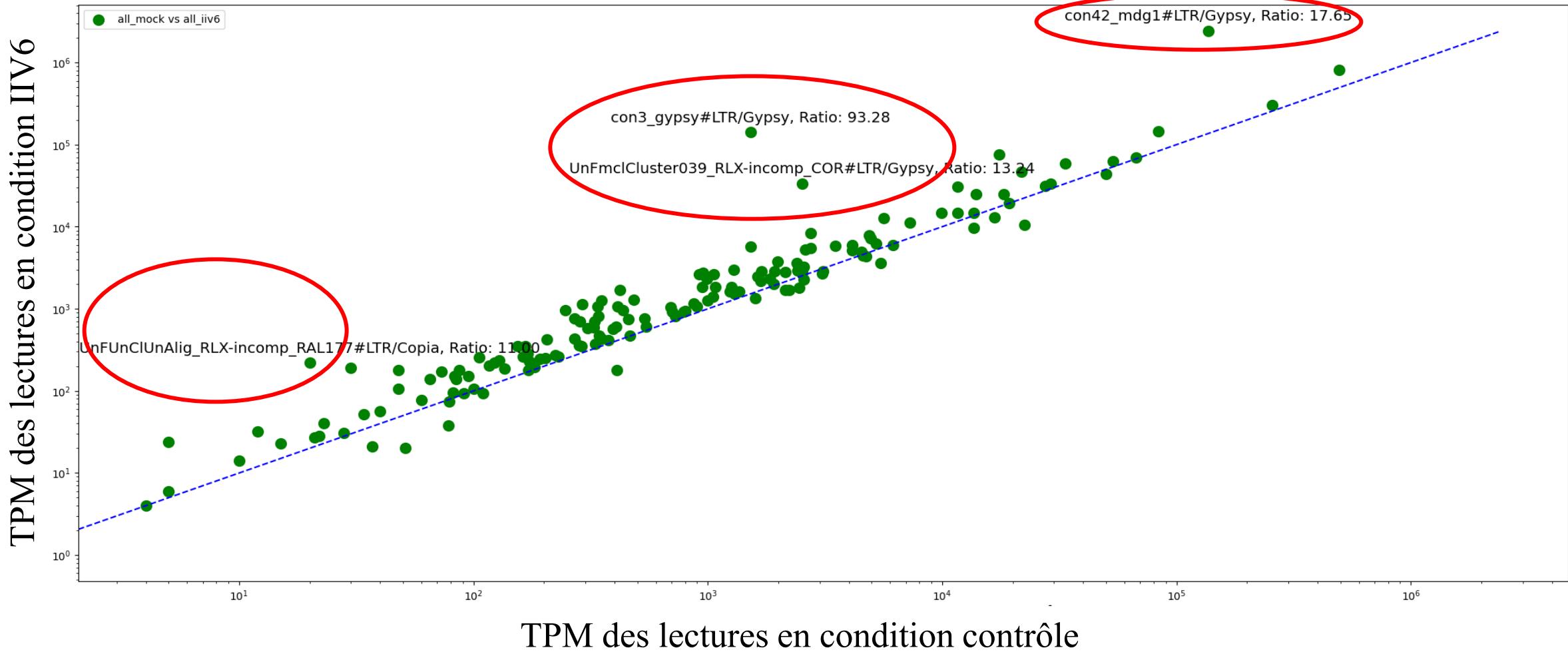
Pipeline de traitement des données utilisées



Analyse des données RNAseq (short reads)

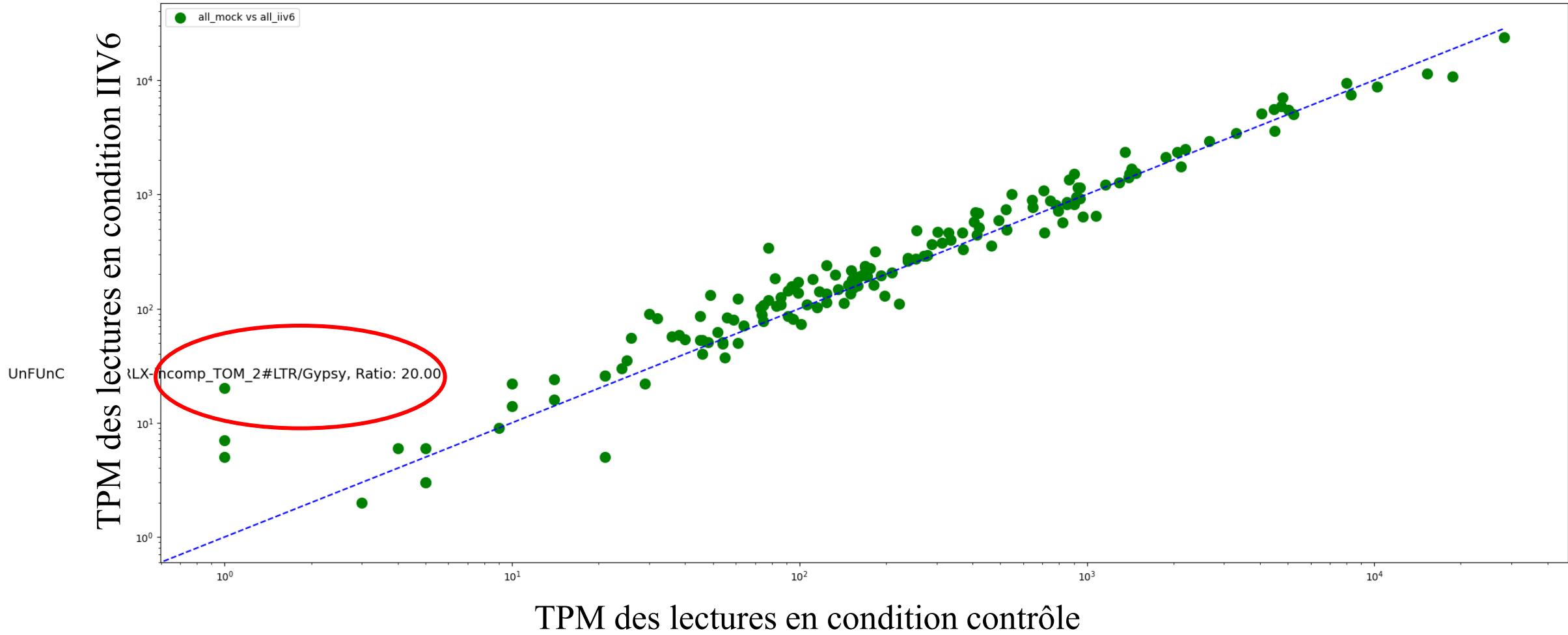
Variations des niveaux d'expression des TE en conditions IIV6 (Dmel : carcasses)

Variation du niveau d'expression des TE en condition infectée par IIV6 Vs contrôle



Variations des niveaux d'expression des TE en conditions IIV6 (Dmel : ovaires)

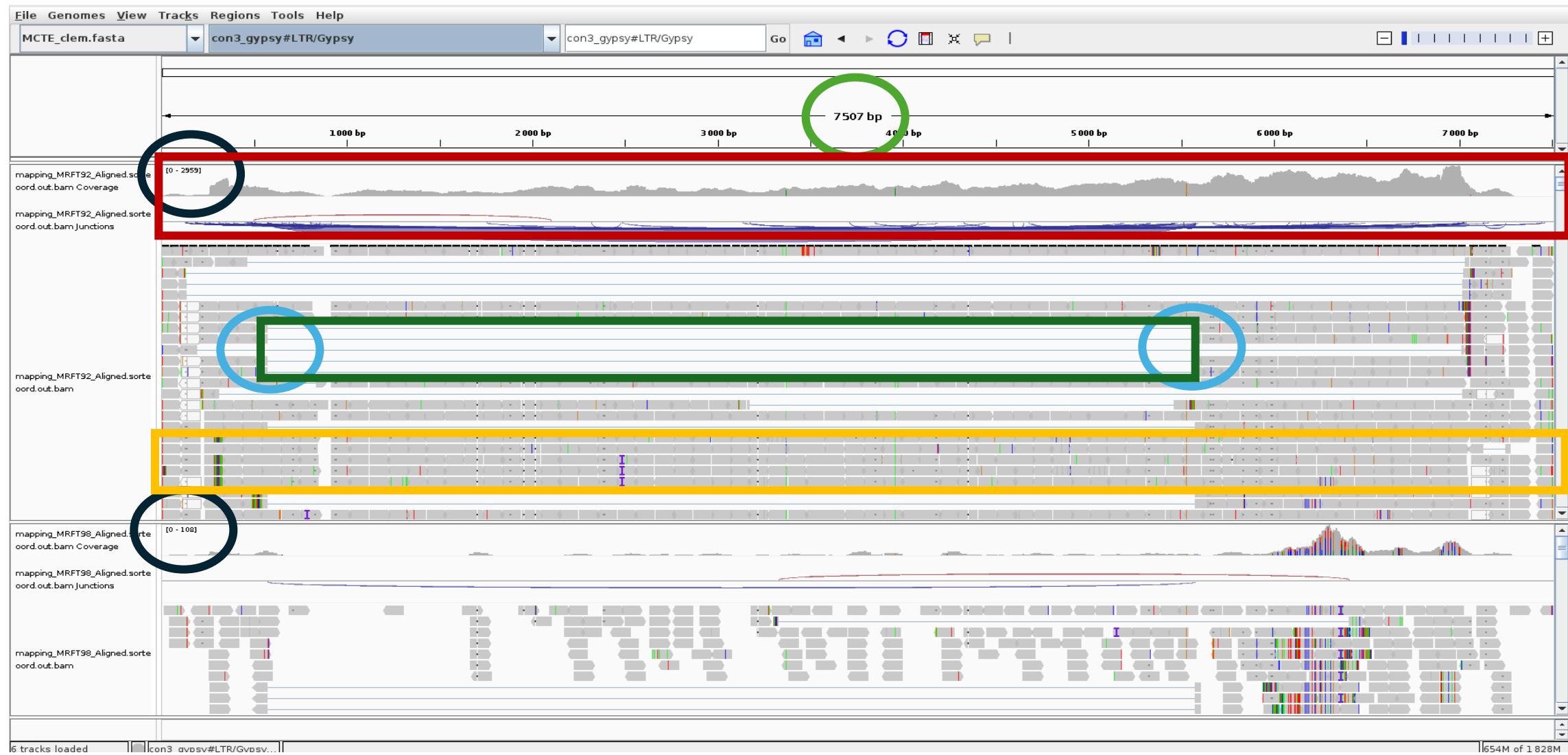
Variation du niveau d'expression des TE en condition infectée par IIV6 Vs contrôle



Résultats DESeq2 des ET différentiellement exprimés entre les conditions expérimentales dans les carcasses et dans les ovaires.

Target_id	Basemean	Log2FC	LfcSE	Stat	P_value	FDR	Target_id	Basemean	Log2FC	LfcSE	Stat	P_value	FDR
Con42_md_g1	359512,61	3,63	0,35	-10,10	5,51E-24	1,37E-22	Con42_md_g1	1647,42	-0,12	0,11	-1,10	0,26	0,68
Con3_gypsy	19471,87	5,97	0,35	-16,78	3,29E-63	2,45E-61	Con3_gypsy	65,82	-1,91	0,33	-5,68	1,34	1,06
UnFmclCluster039_RLX-incomp_COR	5125,16	3,18	0,13	-24,29	2,38E-130	3,55E-128	UnFmclCluster039_RLX-incomp_COR	38,49	-0,46	0,34	-1,33	0,18	0,51

Image IGV des reads alignés sur le consensus de con3_gypsy#LTR/Gypsy



Présence des sites d'épissages pour les reads alignés sur TE con3_gypsy

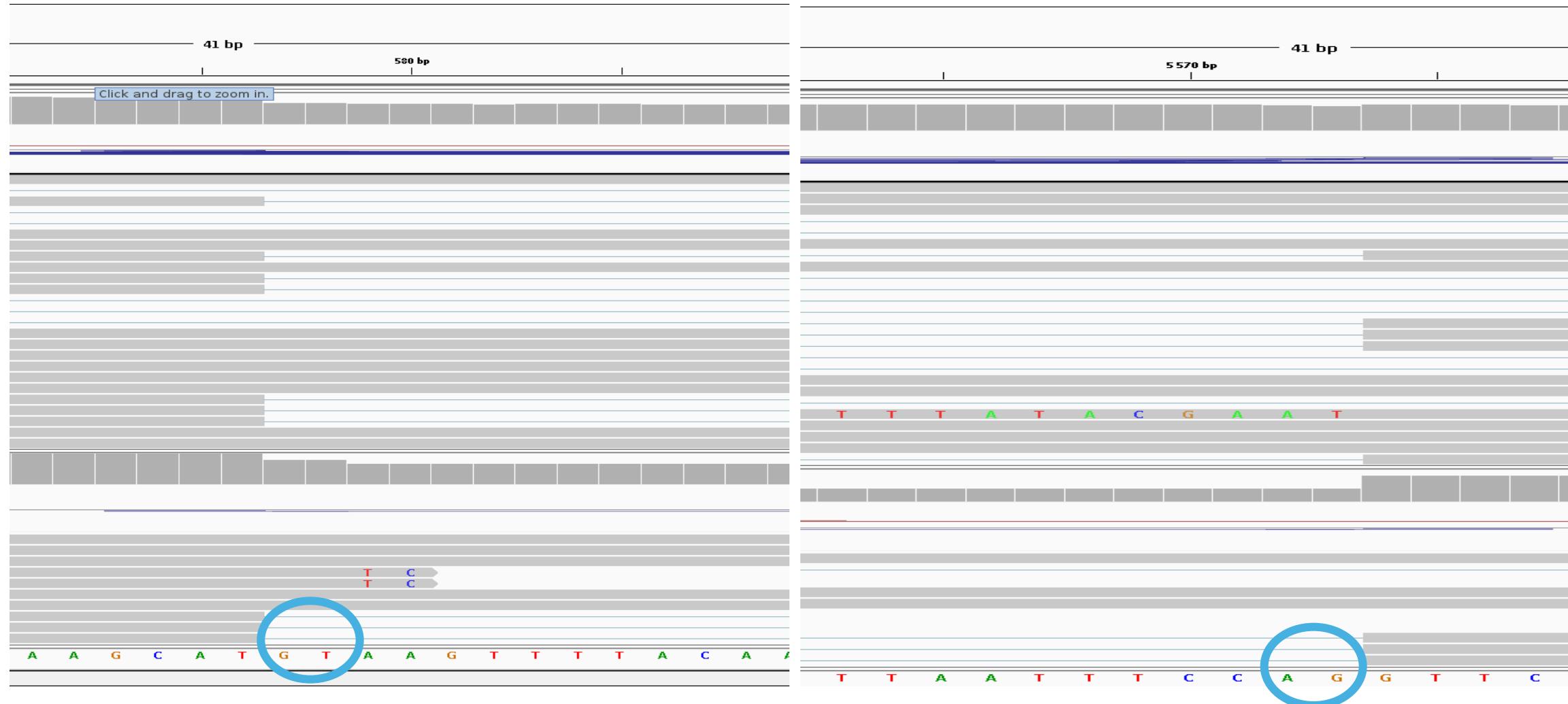
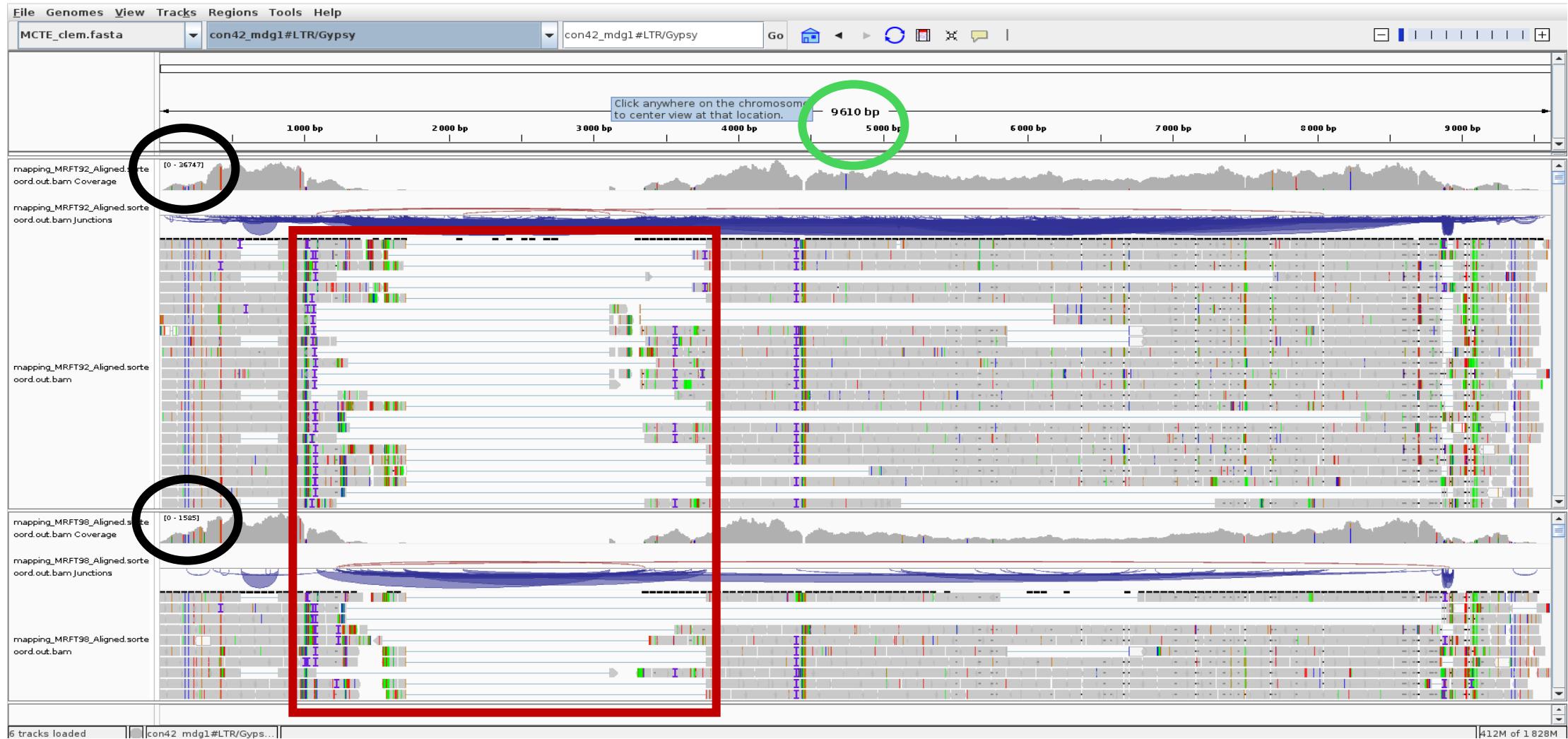


Image IGV de con42_mdg1#LTR/Gypsy



ET différentiellement exprimés chez Dsim en condition infectée

- Con2_gypsy
- Con10_unFmclCluster019_RLX-incomp_COR

ET différentiellement exprimés chez Dmel en condition infectée

- Con3_gypsy
- Con42_mdg1
- UnFmclCluster039_RLX-incomp_COR

Bonne nouvelle !!!

Piste sur la surexpression des ET chez *Drosophila*.

Analyse des données DNaseq (long reads)

Contrôle qualité et stat alignment des reads DNAseq sur les génomes

- Score Q = seuil de qualité des lectures nanopore.
 - Nombre de lectures: 62284
 - 99,8 % > Q10

	Génome
IIV6	43,26%
Dmel	57,91%
Lectures avec soft-clipping	399 soit 0,64 % et 1,48 %

IGV lectures alignées partiellement avec soft-clipping sur le génome de référence de IIv6 (ET con10_3S18#LTR/Pao)



Lectures soft-clippées correspondant aux ET

Eléments transposables	Sites d'insertion génome HIV6	Code CIGAR sur la séquence consensus	Identité (%)	Taille reads alignés
Con10_3S18	(Inséré en amont) : 5686	6S16M1D16M2I10M1I5M1D36M 1D70M3I1M1I95M	91	249 pb
Con14_F-elements	51882 : (Inséré en aval)	447S23M1D68M1D350M3D98M 1D242M14D72M1D166M3D70M	67	1089 pb
Con16_blood	(Inséré en amont) : 1806	2S135M1I191M1D373M1D220M 3D14M3D54M5D73M1D65M1D4 8M1D13M3D39M1D14M1D80M 1D51M3I1M1I233M5421S	77	1604 pb
Con3_invider2	59559 : (Inséré en aval)	54M1D11M1D116M1D33M2D18 M1D49M1D20M2D79M1D91M1 D21M1D89M2D185M7D3M8D29 7M2D40M1883S	35	1106 pb
Con25_1360	135490 : (Inséré en aval)	4S47M2D29M1D148M1D13M2D 7M1D24M3D24M8D4M3D14M1 D42M1D110M1D26M1D8M6D14 M1D3M4D21M3D10M3D7M2I7 M1927S	20	564 pb

ET détectés dans le génome de IIV6

- Con14_F-éléments
- Con10_3S18
- Con16_blood
- Con3_invider2
- Con25_1360

ET différentiellement exprimés chez Dmel en condition infectée

- Con3_gypsy
- Con42_mdg1
- UnFmclCluster039_RLX-incomp_COR

Conclusion

■ Deux hypothèses:

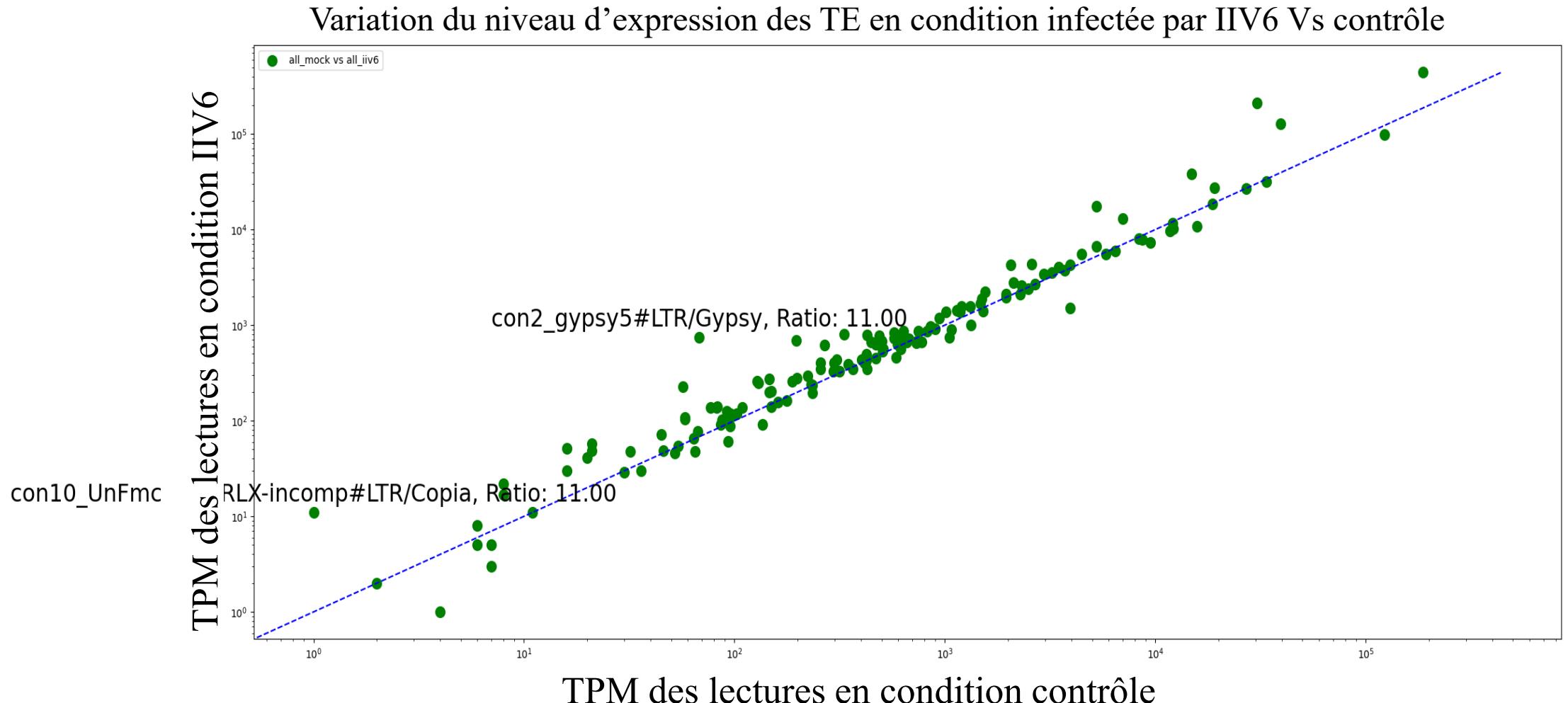
■ ET du IIV6 et de Dmel s'expriment ? 

■ Surexpression des ET induite par infection virale ? 

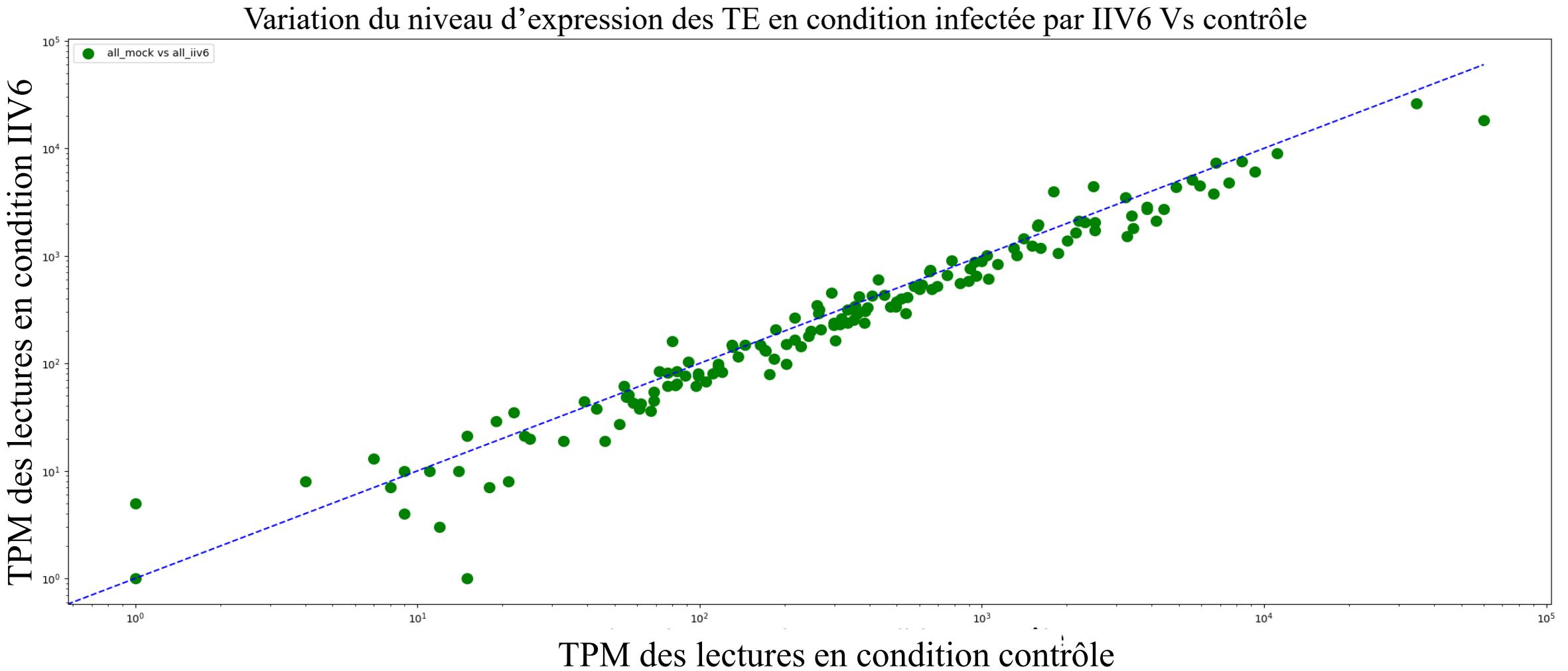
- Remerciements :
- Vincent LACROIX
- Marie FABLET
- Equipes Baobab et GEI
- Corps professoral de bioinformatique

ANEXES

Variations des niveaux de transcrits des TE en conditions IIv6 (Dsim : carcasses)



Variations des niveaux de transcrits des TE en conditions IIV6 (Dsim : ovaires)



IGV lectures alignées partiellement sans soft-clipping sur le génome de référence de IIv6

