

TP M2 bioinfo

Thomas Karaouzene

07 novembre 2017

Informations :

L'objectif de ce TP d'appréhender l'analyse bioinformatique des données issues de séquençage exomique. Les comptes rendus sont à rendre par **binomes**. Sauf exceptions les monomes et trinomes ne seront pas acceptés !!!

Les comptes rendus seront à envoyer à l'adresse tkaraouzene@gmail.com avant le **mardi 14 novembre minuit**. L'horloge de la messagerie faisant foi, les CR reçus après minuit se verront retirer 5 points.

Ce TP est très rapide si vous maîtrisez les outils présentés ($\simeq 20$ minutes), l'objectif ici est donc de vous laisser travailler en **AUTONOMIE**.

Prenez le réflexe d'explorer par vous-même les différents outils.

Analyse de deux frères azoospermes

Afin d'identifier la cause génétique entraînant un phénotype d'azoospermie chez deux frères, vous décidez d'effectuer un séquençage exomique de ceux-ci. En plus de ces deux frères, vous séquencez un troisième individu sain.

En utilisant le logiciel *Variant Effect Predictor* pour annoter les exomes de ces trois individus et en appliquant les filtres classiques, identifiez le variant ayant le plus de risque d'être le responsable du phénotype des deux frères.

Vous détaillerez dans votre rapport l'ensemble des critères de filtre utilisés en justifiant leur utilisation (une simple phrase suffit).

Vous pouvez appliquer les filtres dans l'ordre qui vous convient, néanmoins, certaines étapes seront plus rapide si vous avez appliqué vos filtres dans un ordre logique.

NB: Dans les paramètres de VEP :

. **Identifiers and frequency data** : cochez gnomAD (exomes) allele frequencies

Données

Frère 1 :

CHROM	POS	ID	REF	ALT
1	92647250	.	C	T
2	97818236	.	A	C
3	239675	.	C	T
3	239798	.	G	C
4	57676326	.	G	GA
5	112824068	.	C	CCGC
5	156479452	.	GTTG	G
8	41369893	.	G	A
8	41387610	.	C	T
8	41387642	.	T	G
8	41387809	.	T	C
11	196070	.	G	C

11	199813	.	A	G
15	50540273	.	C	A
15	50546971	.	C	G
16	4924426	.	C	T
16	29912807	.	G	GTGG
16	33629700	.	G	A
20	61919110	.	C	T
22	25334178	.	G	GG
22	29921848	.	A	G
22	46656674	.	TTC	T

Frère 2 :

CHROM	POS	ID	REF	ALT
1	92647250	.	C	T
3	239675	.	C	T
4	57676326	.	G	GA
5	112824068	.	C	CCGC
5	156479452	.	GTTG	G
8	41369893	.	G	A
8	41387610	.	C	T
8	41387809	.	T	C
9	79118329	.	G	A
9	79118400	.	C	T
9	79118475	.	G	A
9	79118628	.	C	T
9	79118633	.	CC	C
11	196070	.	G	C
11	199813	.	A	G
15	50540273	.	C	A
15	50546971	.	C	G
16	4924426	.	C	T
16	29912807	.	G	GTGG
16	33629700	.	G	A
20	61919110	.	C	T
22	25334178	.	G	GG
22	29921848	.	A	G
22	46656674	.	TTC	T

Individu sain :

CHROM	POS	ID	REF	ALT
1	92647250	.	C	T
3	69113153	.	T	C
3	69113155	.	G	A
3	69113185	.	T	A
8	41369893	.	G	A
9	120176925	.	C	CGGC
9	120176929	.	G	A
9	120176967	.	T	C
15	50546971	.	C	G
20	45355480	.	G	A
20	45357878	.	A	G

20	45358005	.	G	T	
22	41743863	.	T	G	
22	41743947	.	CCTC		C
22	41744022	.	C	G	
22	41744334	.	A	G	
22	41745025	.	C	T	

Liens utiles

Variant Effect Predictor : annotation des variants.

Pubmed : bibliographie