

Transcription vs Metilation

Lucas Michel Todó

October 13, 2017

Contents

1	Generalitats	2
2	Heatmaps Percentatge de Coverage	3
2.1	TSS+Gene+TTS	3
3	Heatmaps Coverage	5
3.1	Coverage Gene Body + 5'	5
3.2	Diferència de Coverage, filtrat per diferència de transcripció . . .	6
3.3	Diferència de Coverage, filtrat per diferència de metilació	8
3.4	Diferència de Coverage: auto-clustering	10
4	Coverage a pics diferencials	12
4.1	Coverage a pics diferencials	12
4.2	Diferència de coverage a pics diferencials	13
5	Resta de gràfics	15
6	Percentatge de metilació	15
6.1	Gene-body	15
6.2	TSS	16
6.3	Gene + TSS	17
6.4	TTS	18
7	Coverage	19
7.1	Coverage Gene Body	19
7.2	Coverage 5'	20
7.3	Coverage 3'	21
8	Diferència de coverage	22
8.1	Diferència de Coverage a 5', filtrat per diferència de transcripció	22
8.2	Diferència de Coverage a genebody, filtrat per diferència de transcripció	23
8.3	Diferència de Coverage 5'genebody, filtrat per diferència de transcripció	24

8.4	Diferència de Coverage a 3', filtrat per diferència de transcripció	25
9	Multiple-regression Model	26

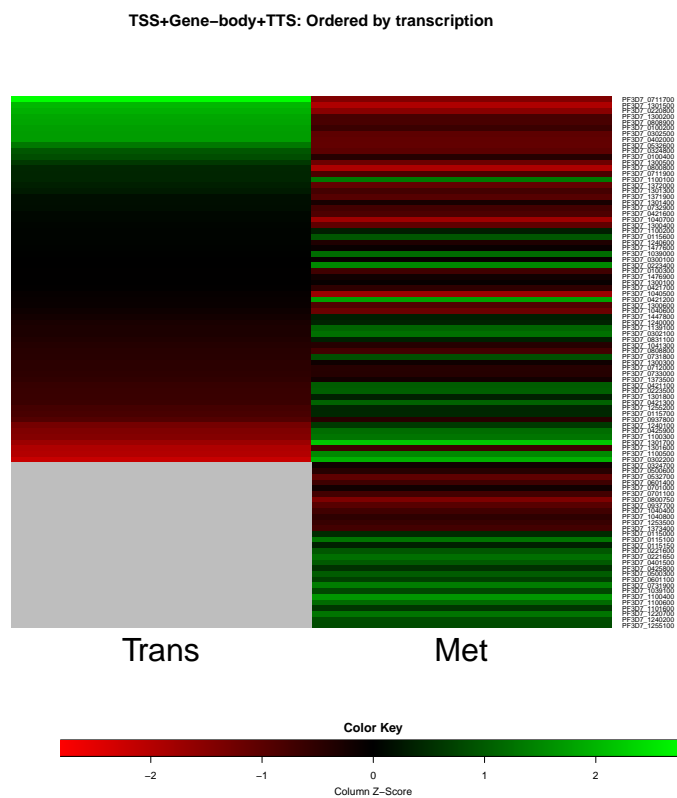
1 Generalitats

En TOTS els gràfics la columna de més a l'esquerra correspon als valors de diferència de transcripció (1.2B - 10G). Les columnes següents corresponen a algun paràmetre relacionat amb la metilació (percentatge del gen metilat, coverage mitjà...). En alguns casos hi ha valor de metilació per a 1.2B i per a 10G i en altres una sola columna que correspon a la diferència de metilació. Per a cada tipus de gràfic se n'han fet diversos tenint en compte la metilació a tres zones diferents: TSS, gene-body i TTS. Per a cada tipus de gràfic he posat el més representatiu a tall d'exemple. Al final del document hi ha tota la resta de gràfics.

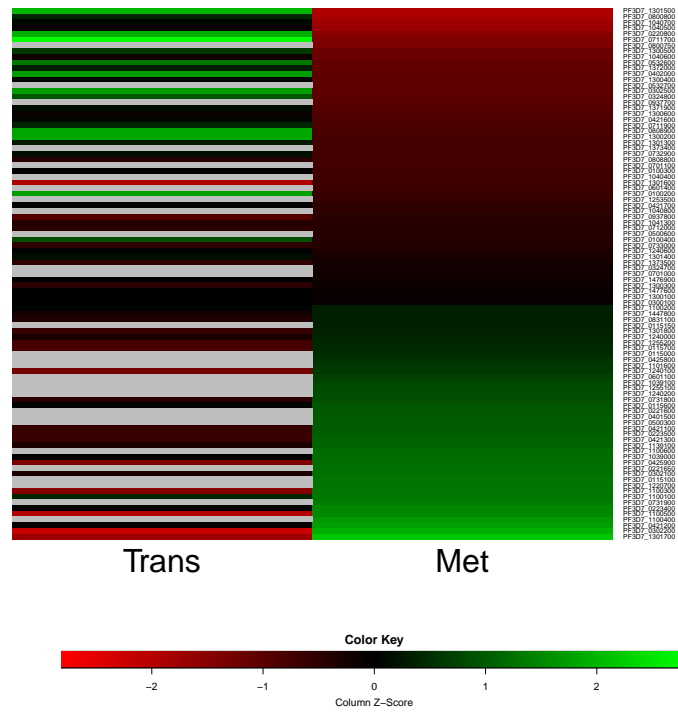
2 Heatmaps Percentatge de Coverage

En aquest gràfic els valors corresponents a les columnes 1.2B i 10G corresponen al percentatge del gen cobert en pics diferencials. Per a obtenir aquests valors s'han unificat les dues llistes de pics diferencials anotats. Tant per a 1.2B com a 10G els gens que corresponen a "pics de l'altra llista" tenen valors de 0.

2.1 TSS+Gene+TTS



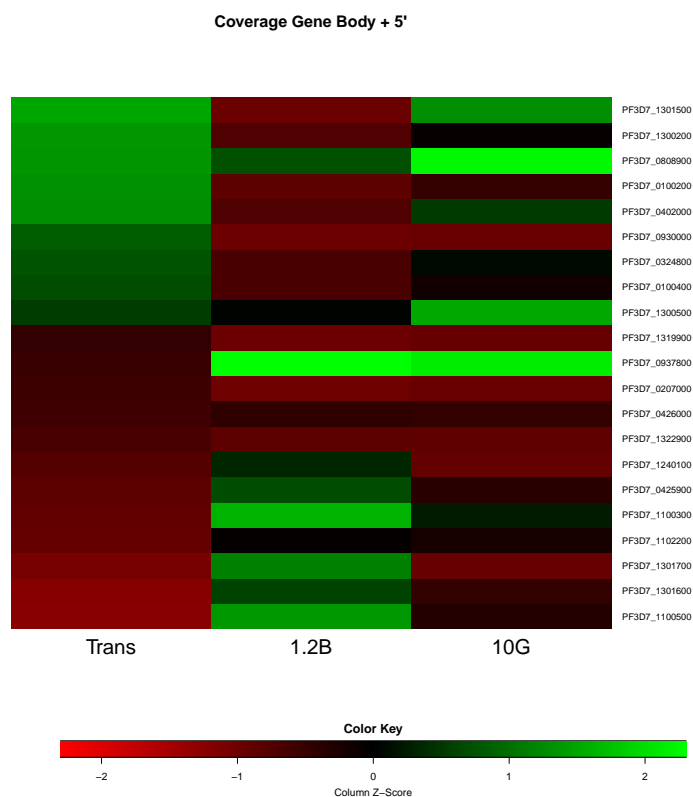
TSS+Gene-body+TTS: Ordered by methylation



3 Heatmaps Coverage

En aquests gràfics el valor de les columnes 1.2B i 10G corresponen al valor de coverage de cada gen. El coverage s'ha calculat per a tots els gens i s'ha "normalitzat" tant per nombre de reads (per a compensar el fet que els alineaments amb més reads tenen més coverage en general) com pel tamany del gen (per compensar el fet que els gens més llargs tenen més reads). Els gens que es mostren al gràfic estàn filtrats: només apareixen aquells que tenen una diferència de transcripció (primera columna) major a 0.7 (en valor absolut).

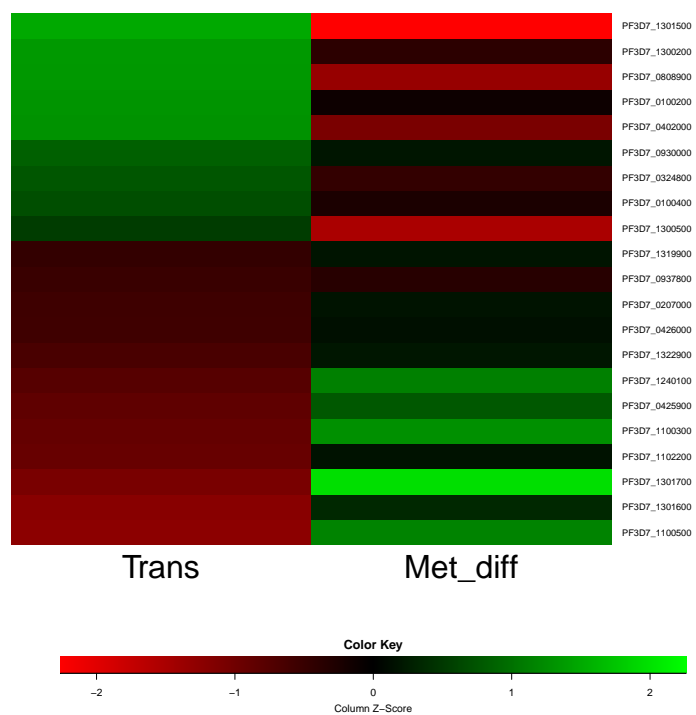
3.1 Coverage Gene Body + 5'



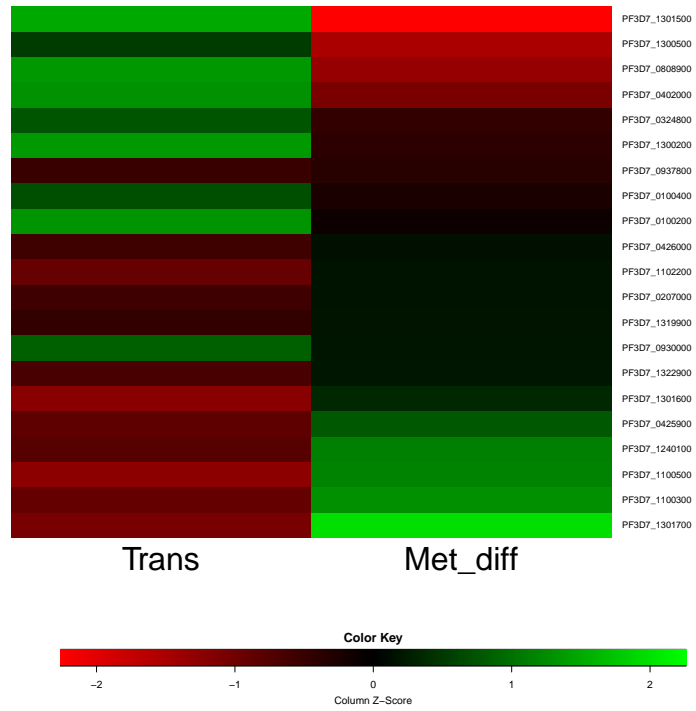
3.2 Diferència de Coverage, filtrat per diferència de transcripció

En aquests gràfics hi ha una sola columna corresponent a metilació que correspon a diferència de coverage. Aquesta diferència és simplement la resta (1.2B - 10G) dels valors obtinguts en l'apartat anterior.

Diferència de Coverage ordenat per transcripció

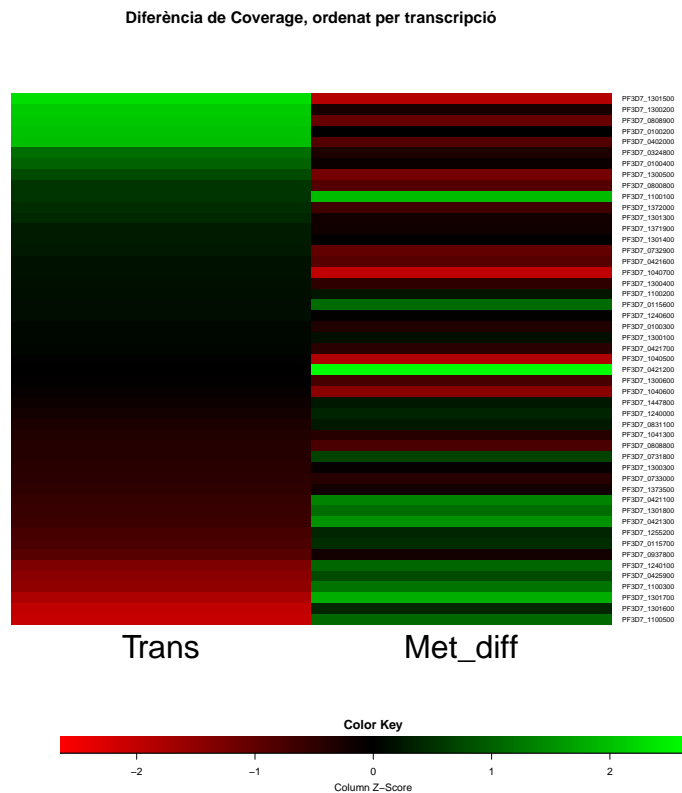


Diferència de Coverage ordenat per Metilació

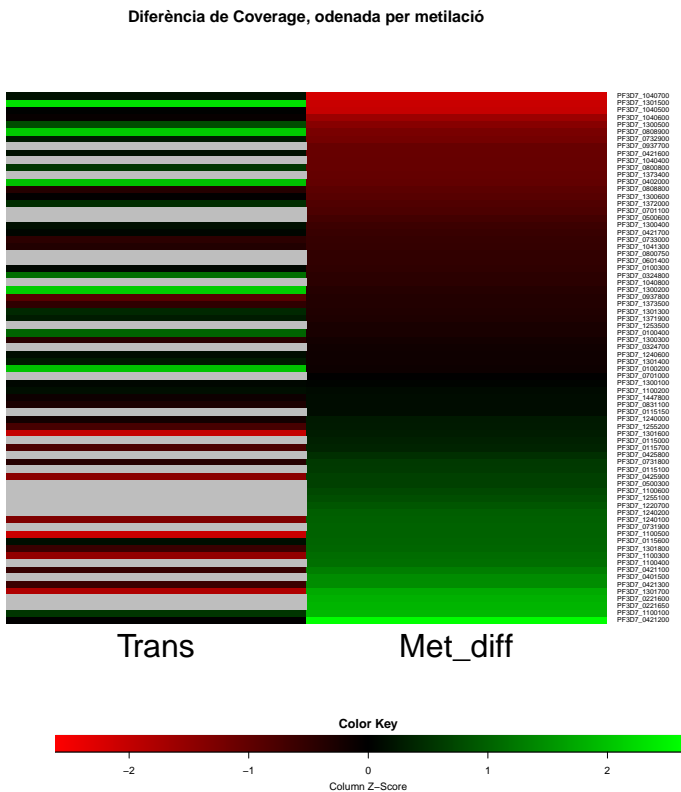


3.3 Diferència de Coverage, filtrat per diferència de metilació

Aquest gràfic és el mateix que l'anterior però en comptes de filtrar per valors de transcripció els gens que hi apareixen estan filtrats per valors de metilació. Els gens que apareixen al gràfic són aquells que apareixen a alguna de les llistes de gens pics diferencials.



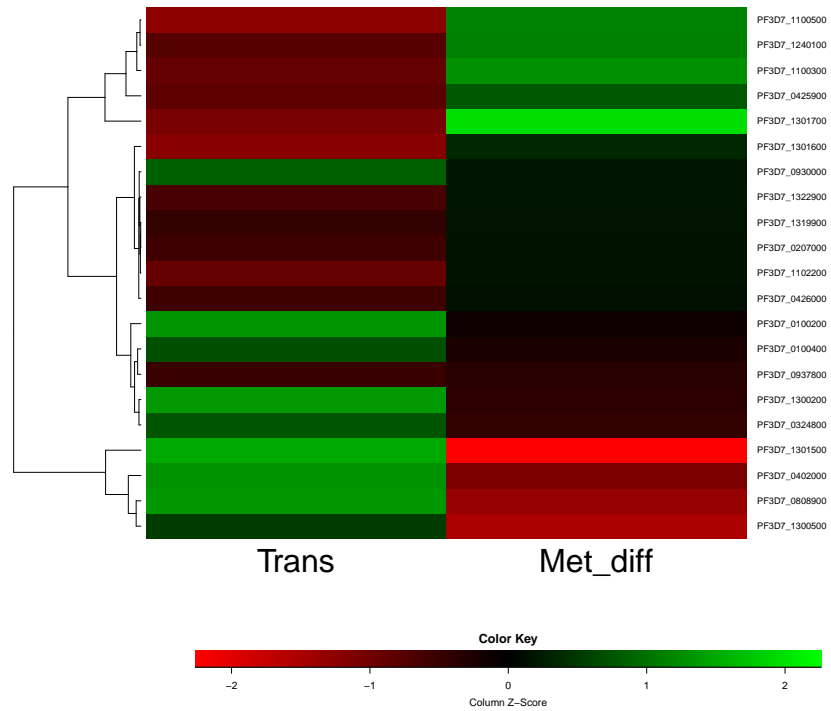
En aquest cas hem deixat els gens per als que no tenim valors de transcripció.



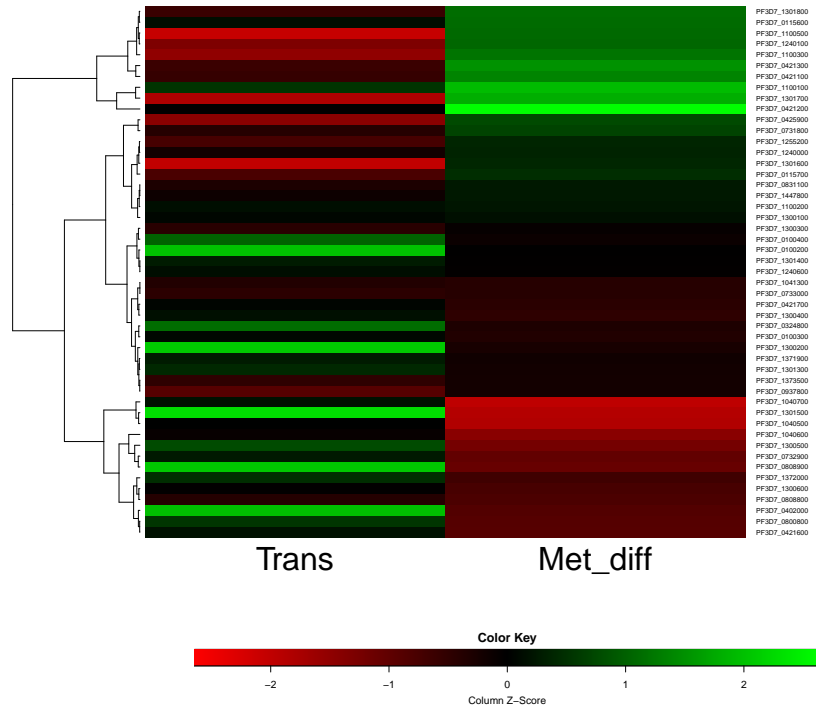
3.4 Diferència de Coverage: auto-clustering

En aquest cas hem deixat que els gens s'ordini "sol" fent servir l'algoritme de clustering de la funció "heatmap2" d'R.

Diferència de Coverage filtrat per transcripció, clustering



Diferència de Coverage filtrat per metilació, clustering

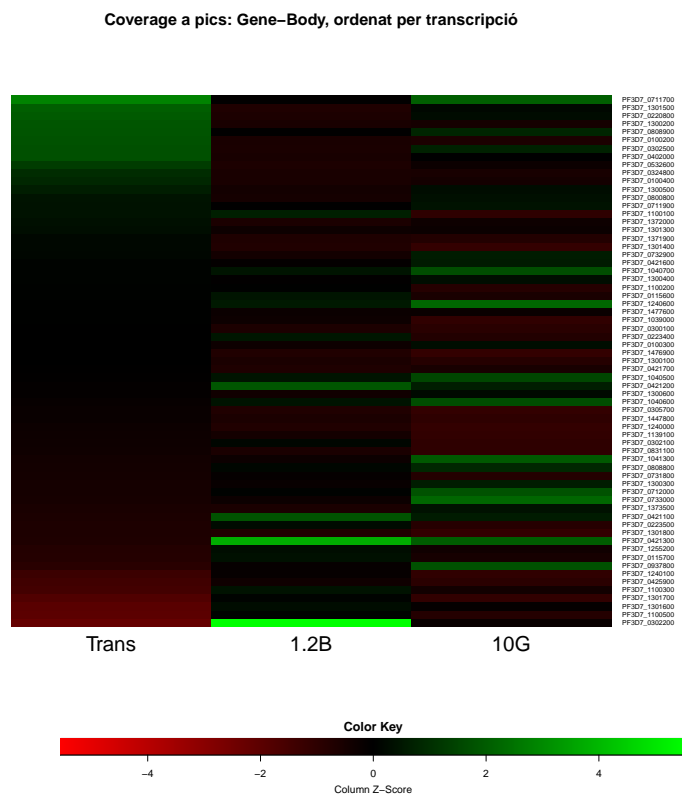


4 Coverage a pics diferencials

En aquests gràfics s'ha calculat el coverage mitjà de la mateixa manera que en els anteriors, però l'anàlisi s'ha restringit a aquells gens que apareixen a qualsevol de les dues llistes de pics diferencials (pics diferencials a 1.2B i a 10G).

4.1 Coverage a pics diferencials

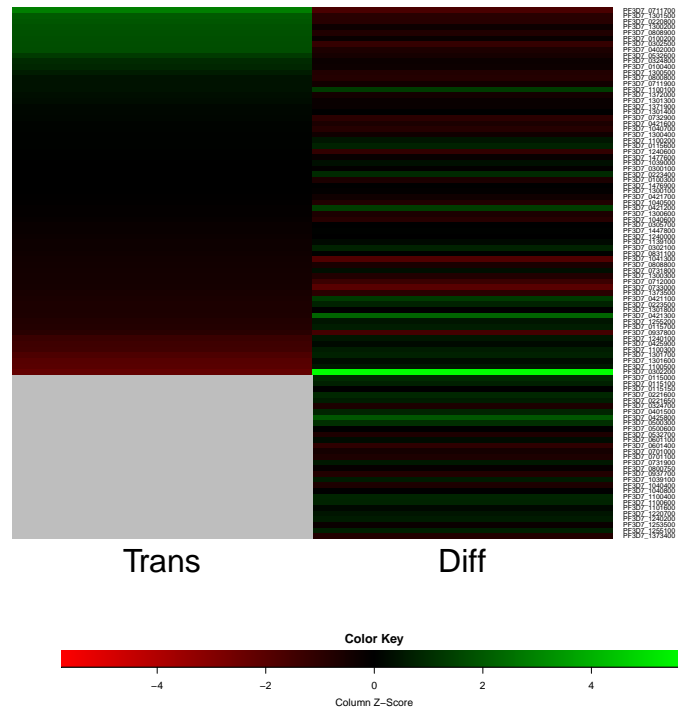
En aquests gràfics els valors de les columnes 1.2B i 10G corresponen al coverage mitjà de cada gen només a les zones on hi ha pics diferencials (de metilació).



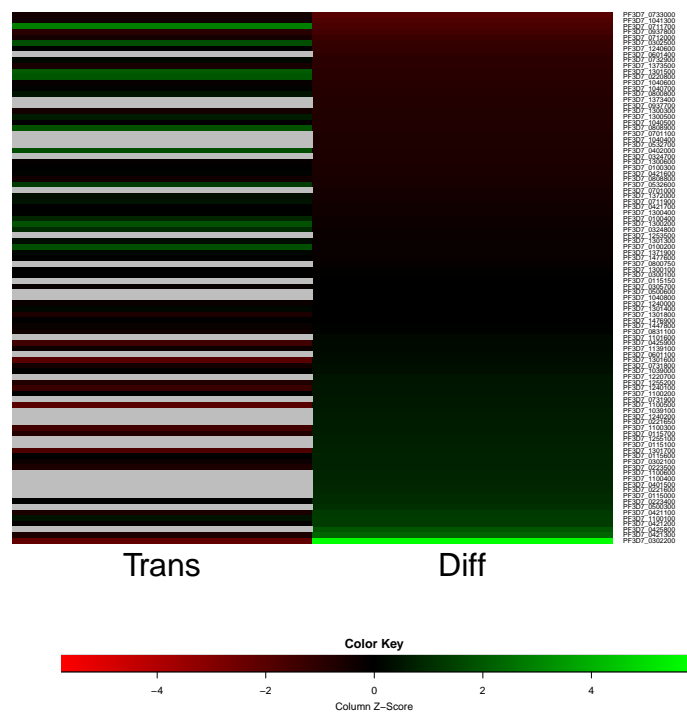
4.2 Diferència de coverage a pics diferencials

En aquests gràfics tenim una sola columna a la dreta que correspon a la diferència de coverage entre 1.2B i 10G només a les zones on hi ha pics.

Diferència de coverage a pics: Gene-Body, ordenat per transcripció



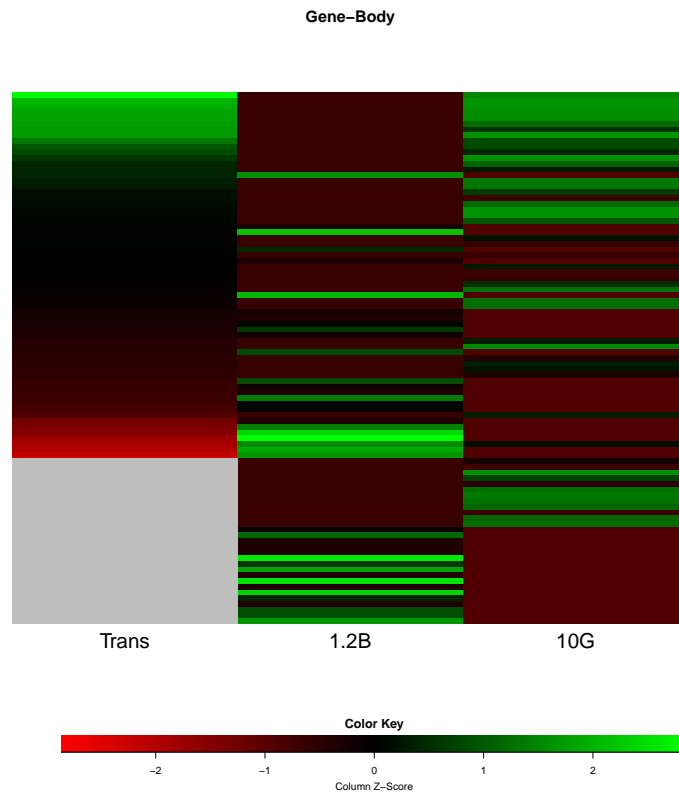
Diferència de coverage a pics: Gene-Body, ordenat per metilació



5 Resta de gràfics

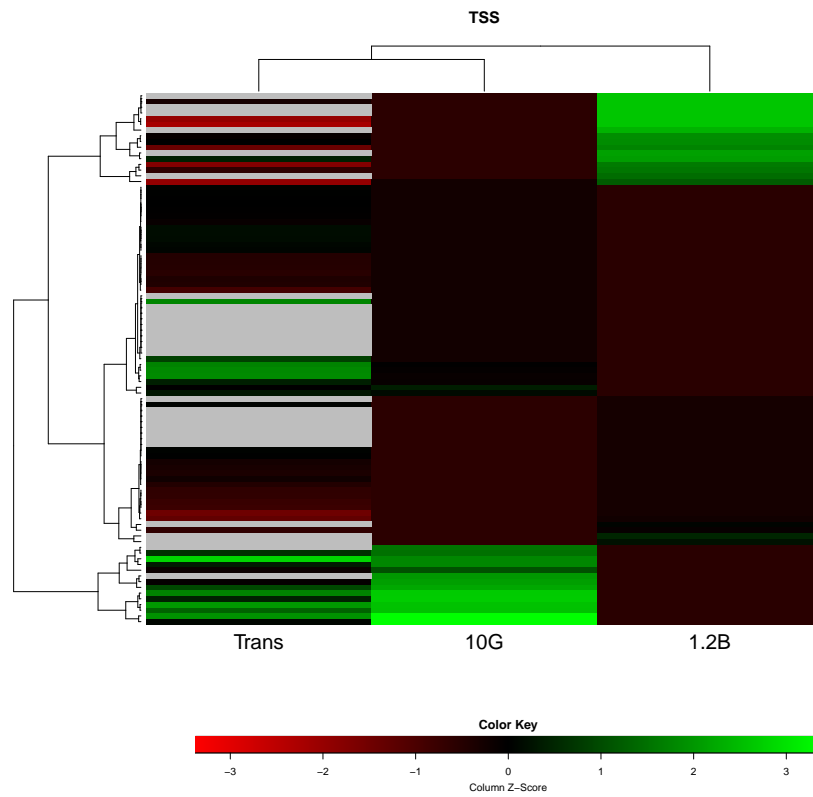
6 Percentatge de metilació

6.1 Gene-body

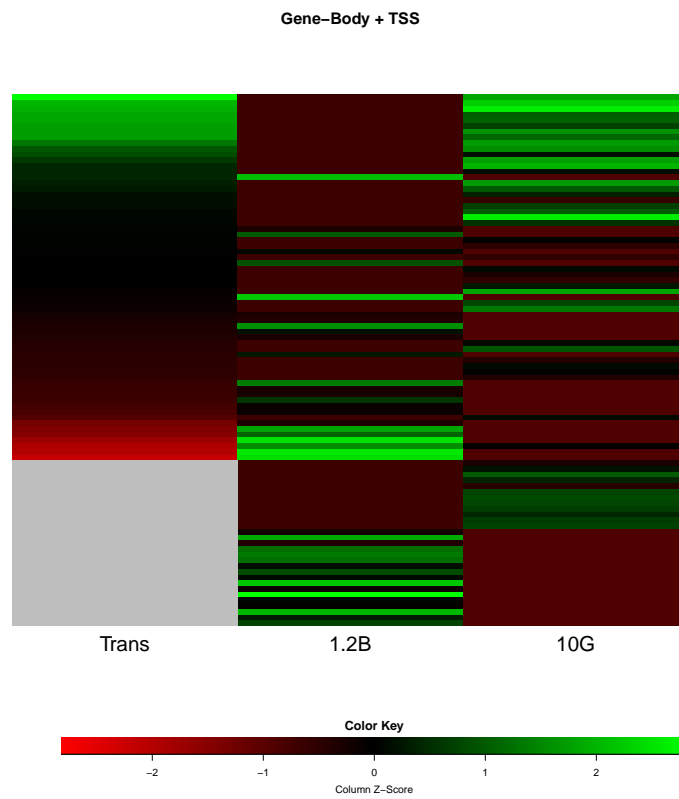


6.2 TSS

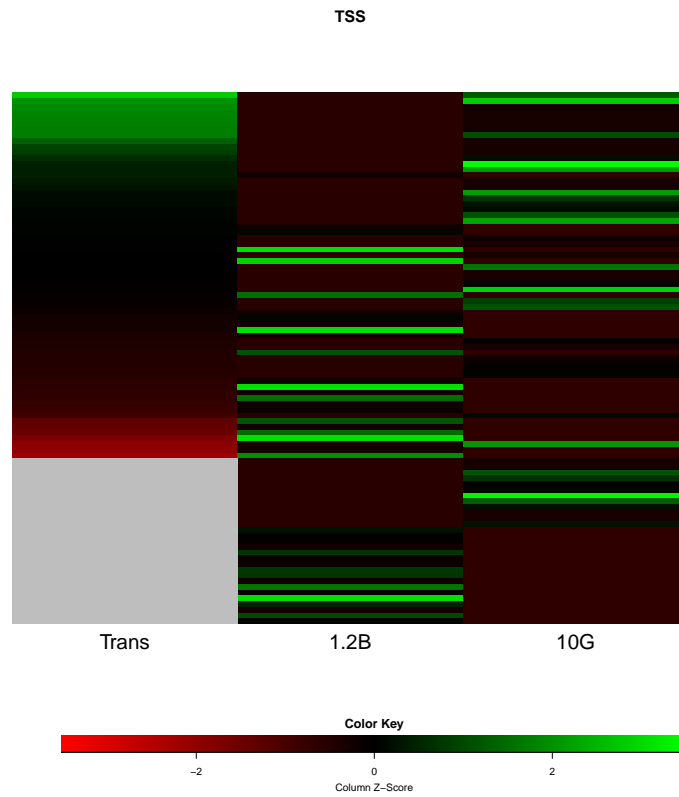
En aquest gràfic en comptes d'ordenar els gens en funció de la primera columna s'han ordenat "sols" seguint l'algoritme de clustering de la funció "heatmap" de R.



6.3 Gene + TSS

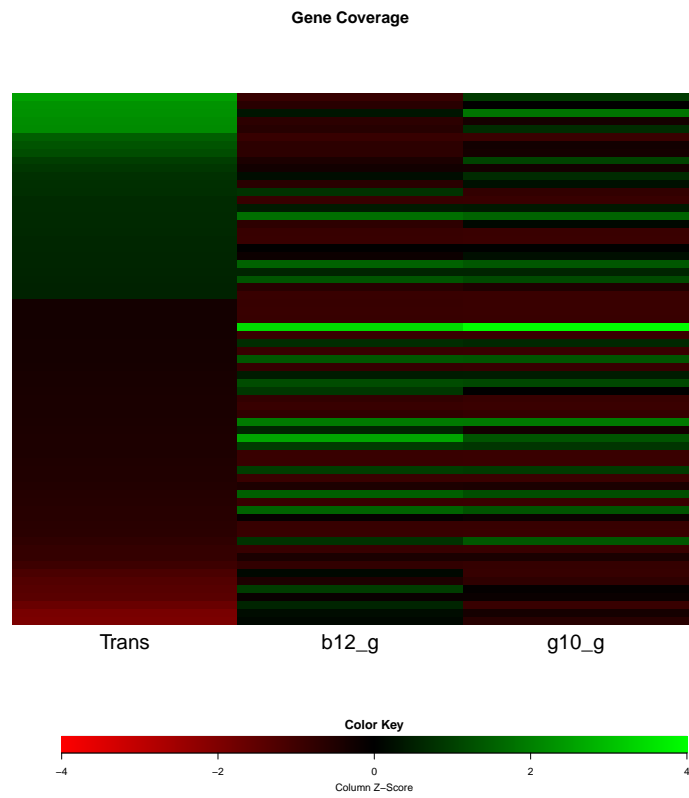


6.4 TTS

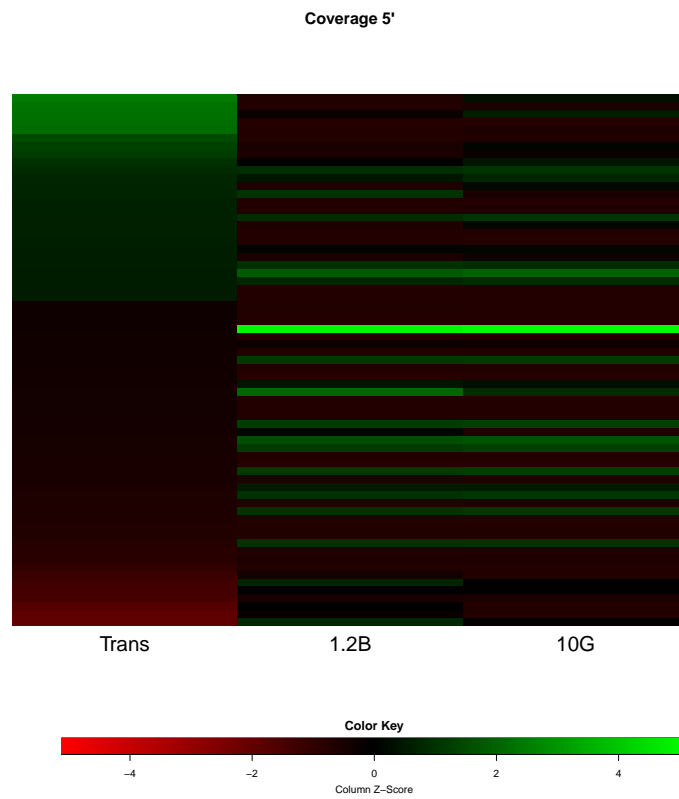


7 Coverage

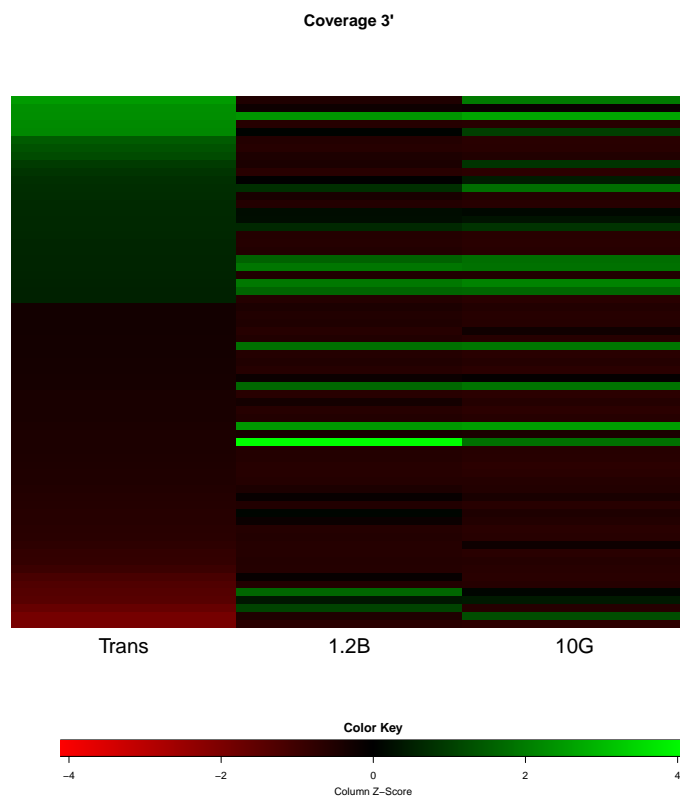
7.1 Coverage Gene Body



7.2 Coverage 5'

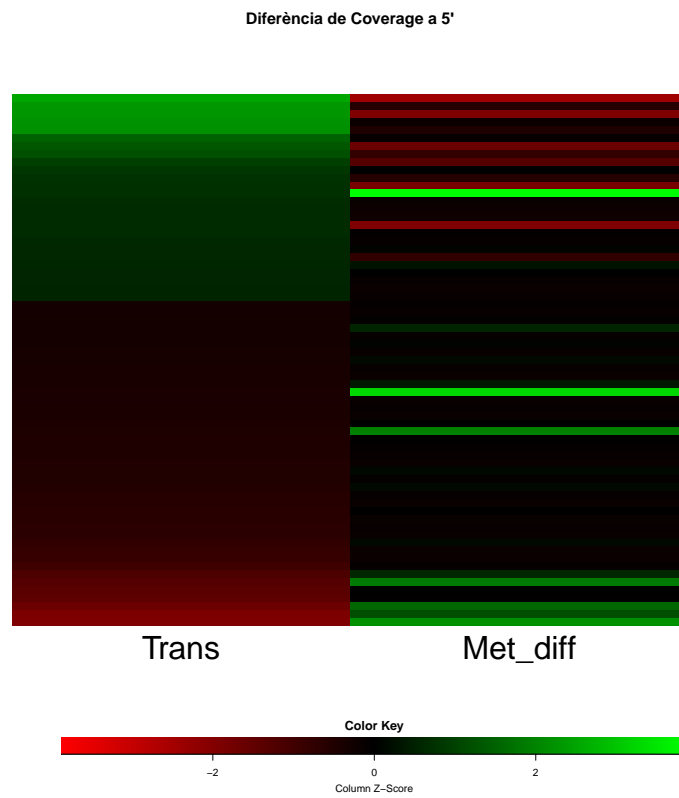


7.3 Coverage 3'

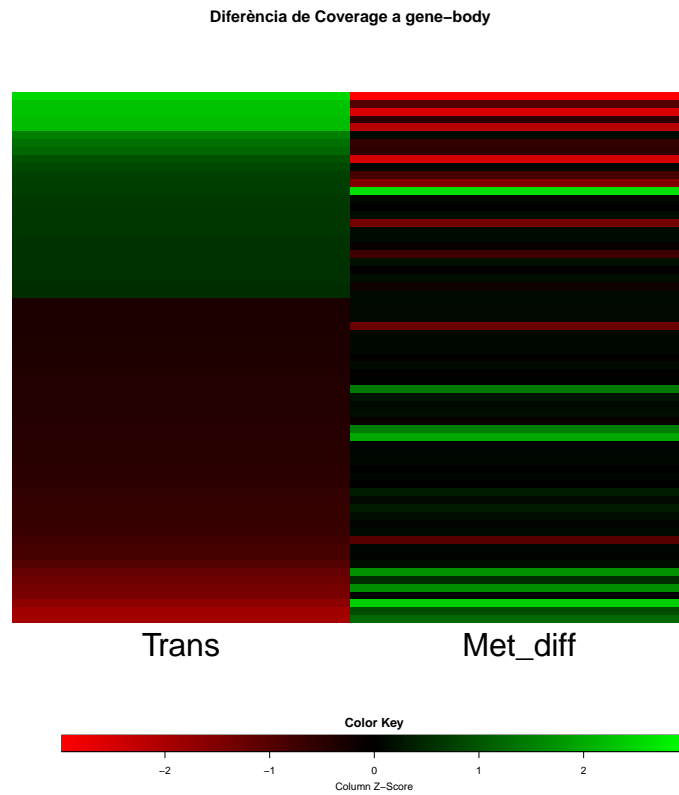


8 Diferència de coverage

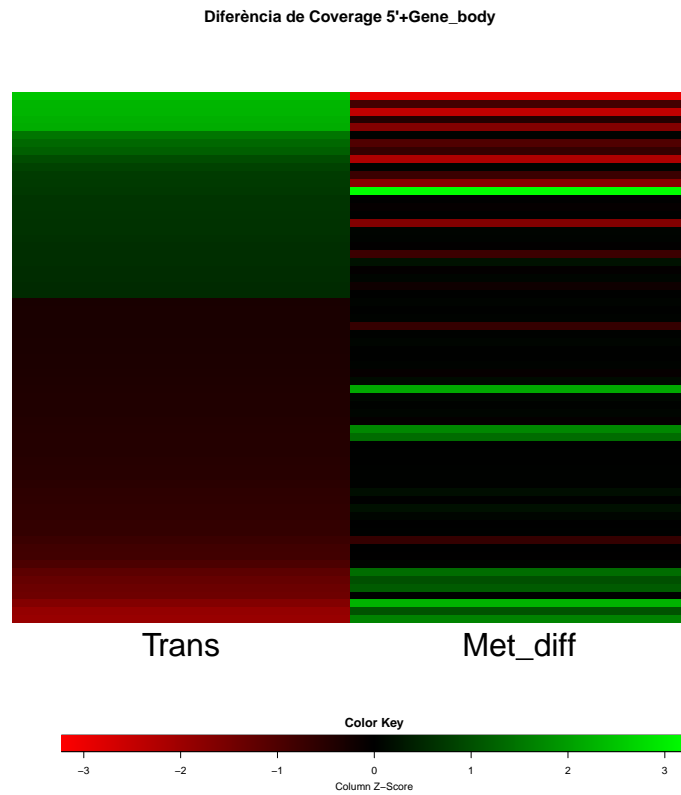
8.1 Diferència de Coverage a 5', filtrat per diferència de transcripció



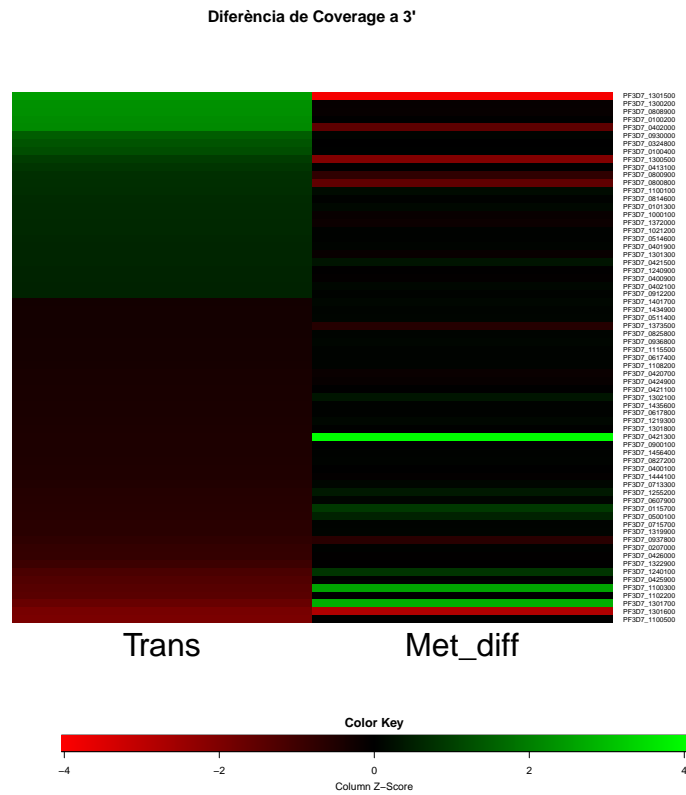
8.2 Diferència de Coverage a genebody, filtrat per diferència de transcripció



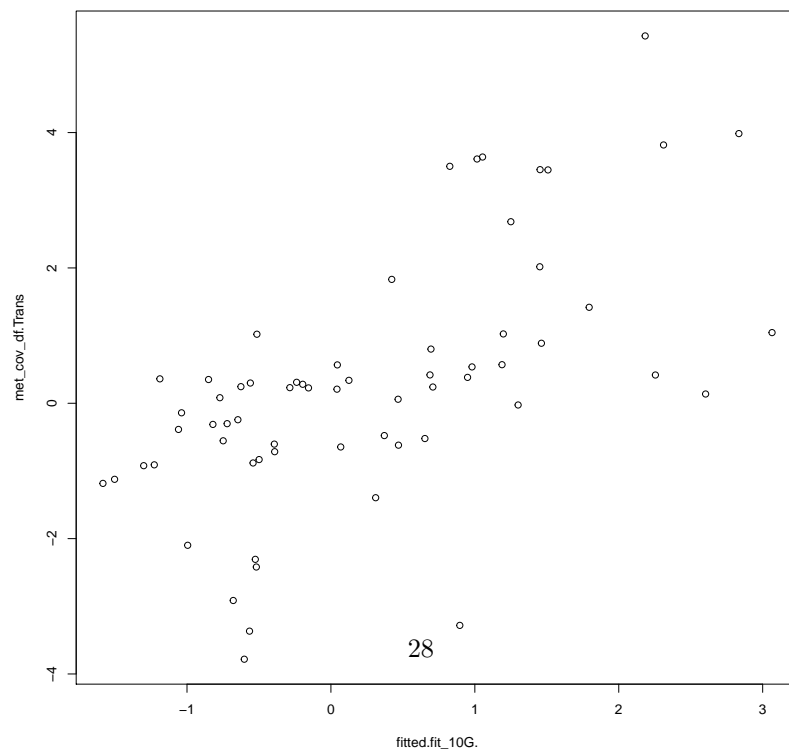
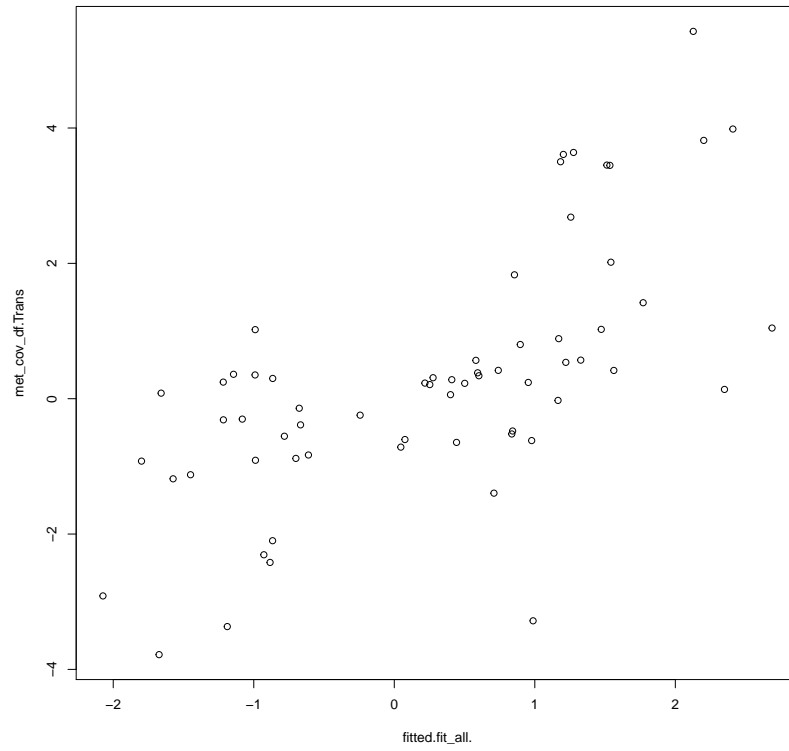
8.3 Diferència de Coverage 5'genebody, filtrat per diferència de transcripció

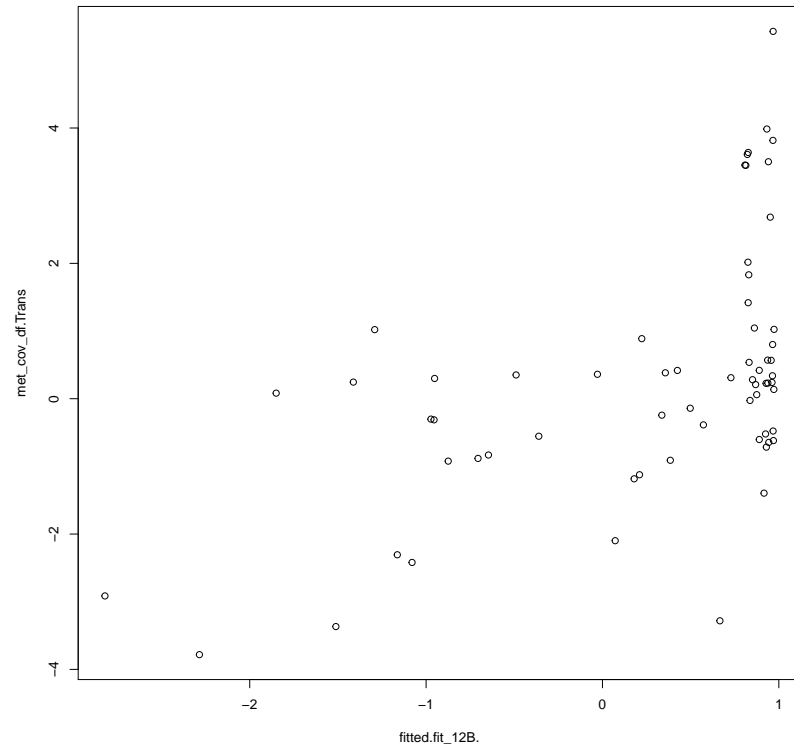


8.4 Diferència de Coverage a 3', filtrat per diferència de transcripció



9 Multiple-regression Model





Diferença de Coverage a 3'

