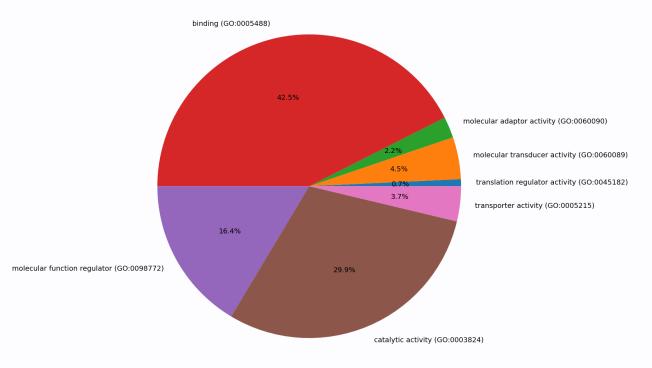
GO-анализ генов, полученных из пересечений окрестностей SNPs, связанных с редактированием РНК, с участками Z-ДНК (DeepZ / zhunt)

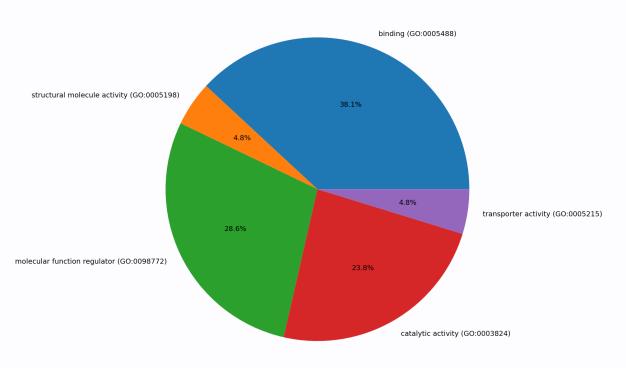
Орлов Александр
ФКН ВШЭ

Для начала воспользуемся PANTHER. Сравним классификацию генов по биологическим функциям (GO Biological Process):





Panter GO-Slim Molecular Function for DeepZ Genes



Видим, что распределения похожи. Теперь воспользуемся PANTHER Overrepresentation Test с точным тестом Фишера с подсчетом FDR на уровне значимости 0,05. Будем искать в базе GO Biological Process. Для ZHunt получаем следующие статзначимые обогащенные категории:

GO biological process	over/under	raw P- value	FDR
positive regulation of cell differentiation (GO:0045597)	+	6.02E- 06	3.16E- 02
regulation of cell communication (GO:0010646)	+	3.37E- 06	5.32E- 02
regulation of signaling (GO:0023051)	+	4.06E- 06	3.20E- 02

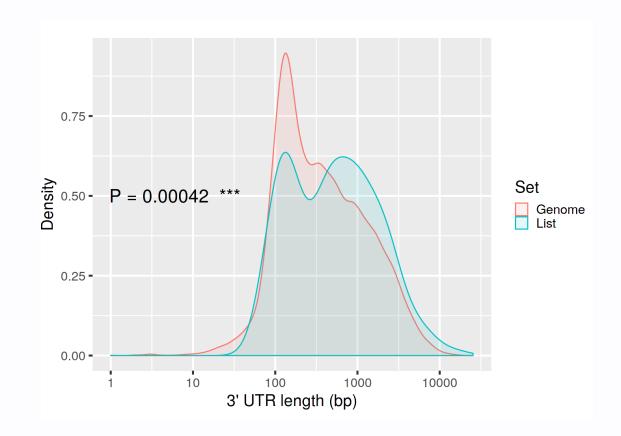
Проводя аналогичный тест для DeepZ не получаем никаких статзначимых результатов. На всякий случай я попробовал взять окна для снипов побольше (500 / 1000) и повторил тест с новым набором генов, но результат остался таким же. Также никаких статзначимых результатов не удалось найти, используя биномиальный тест с коррекцией Бонферрони.

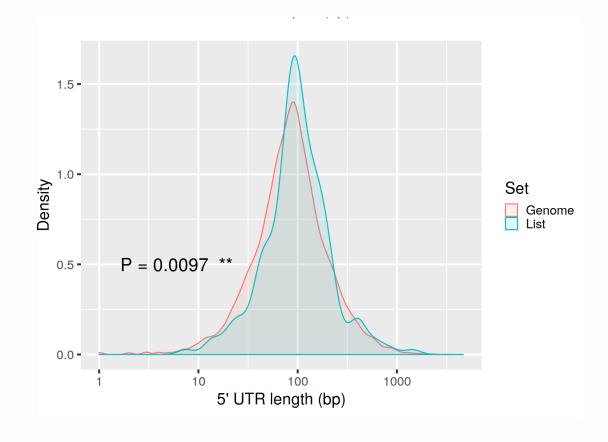
Поэтому теперь попробуем воспользоваться ShinyGO.

Уровень значимости оставляем равным 0,05. Для ZHunt получаем 106 категорий! Здесь привожу некоторые категории с самым маленьким p-value:

- Regulation of cell motility
- Morphogenesis of a branching structure
- Regulation of cell communication
- Positive regulation of cell differentiation
- Branching morphogenesis of an epithelial tube
- Cell motility
- Localization of cell
- Morphogenesis of a branching epithelium

Из интересного также можно посмотреть на длины 3'UTR и 5'UTR для наших генов:





Для DeepZ, к сожалению, получаем такой же результат, как и в PANTHER - никаких статзначимых обогащений, даже если брать снипы с большими окрестностями.