



# ВЫСШАЯ ШКОЛА ЭКОНОМИКИ МАЙНОР БИОИНФОРМАТИКА

Домашняя работа

## Протоонкогены

*Студент:*  
Ведерников Михаил

*Преподаватель:*  
Попцова М. С.

Москва, 2023

## Содержание

1	Поиск информации о гене . . . . .	1
2	Поиск мутаций у пациента . . . . .	4

Анализ данных производился в питоне ([файл на гите](#))

# 1 Поиск информации о гене

Мною был выбран ген KRAS ([IntOGen](#)).

1) Посмотрим с какими типами рака ассоциируются мутации в этом гене:

Show 

25

 entries

Search:

Cancer type	Methods	Samples	Samples (%)
Pancreas adenocarcinoma	CBaSE dNdScv HotMAPS MutPanning CLUSTL smRegions	728	81.25
Colorectal adenocarcinoma	CBaSE dNdScv HotMAPS MutPanning CLUSTL smRegions	515	40.2
Multiple myeloma	CBaSE dNdScv HotMAPS MutPanning CLUSTL smRegions	306	27.27
Lung adenocarcinoma	CBaSE dNdScv HotMAPS MutPanning CLUSTL smRegions	203	26.85
Endometrial cancer	CBaSE combination dNdScv HotMAPS MutPanning CLUSTL smRegions	85	14.66
Non small cell lung cancer	CBaSE dNdScv HotMAPS MutPanning CLUSTL smRegions	65	16.93
Cholangiocarcinoma	CBaSE combination dNdScv HotMAPS MutPanning CLUSTL smRegions	61	15.6
Acute lymphoblastic leukemia	CBaSE combination dNdScv HotMAPS MutPanning smRegions	41	11.29
Stomach adenocarcinoma	CBaSE combination dNdScv HotMAPS MutPanning	36	5.09
Bladder cancer	combination dNdScv HotMAPS MutPanning	26	3.16
Ovary cancer	CBaSE dNdScv HotMAPS MutPanning CLUSTL	23	3.33
Acute myeloid leukemia	CBaSE dNdScv HotMAPS	19	3.03
Cervix squamous cancer	dNdScv HotMAPS MutPanning	16	4.69
Esophageal cancer	combination dNdScv HotMAPS MutPanning	14	1.48
Breast adenocarcinoma	dNdScv HotMAPS MutPanning CLUSTL	13	0.48
Male germ cell tumor	CBaSE dNdScv HotMAPS MutPanning CLUSTL smRegions	12	8.33
Bowel cancer	CBaSE dNdScv HotMAPS MutPanning CLUSTL	11	73.33
Cutaneous melanoma of the skin	HotMAPS	9	0.86
Hepatic cancer	combination	7	0.43
Uterine carcinosarcoma	CBaSE dNdScv HotMAPS MutPanning CLUSTL	7	12.28
Chronic lymphoblastic leukemia	combination dNdScv CLUSTL	6	0.82
Diffuse large B-cell lymphoma	combination	5	1.36
Pilocytic astrocytoma	dNdScv	5	0.81
Renal papillary cell carcinoma	dNdScv MutPanning CLUSTL	5	1.81
Anus cancer	dNdScv	4	23.53

Showing 1 to 25 of 29 entries

Previous

1

2

Next

Рис. 1.1

2) Посмотрим на известные типы мутаций:

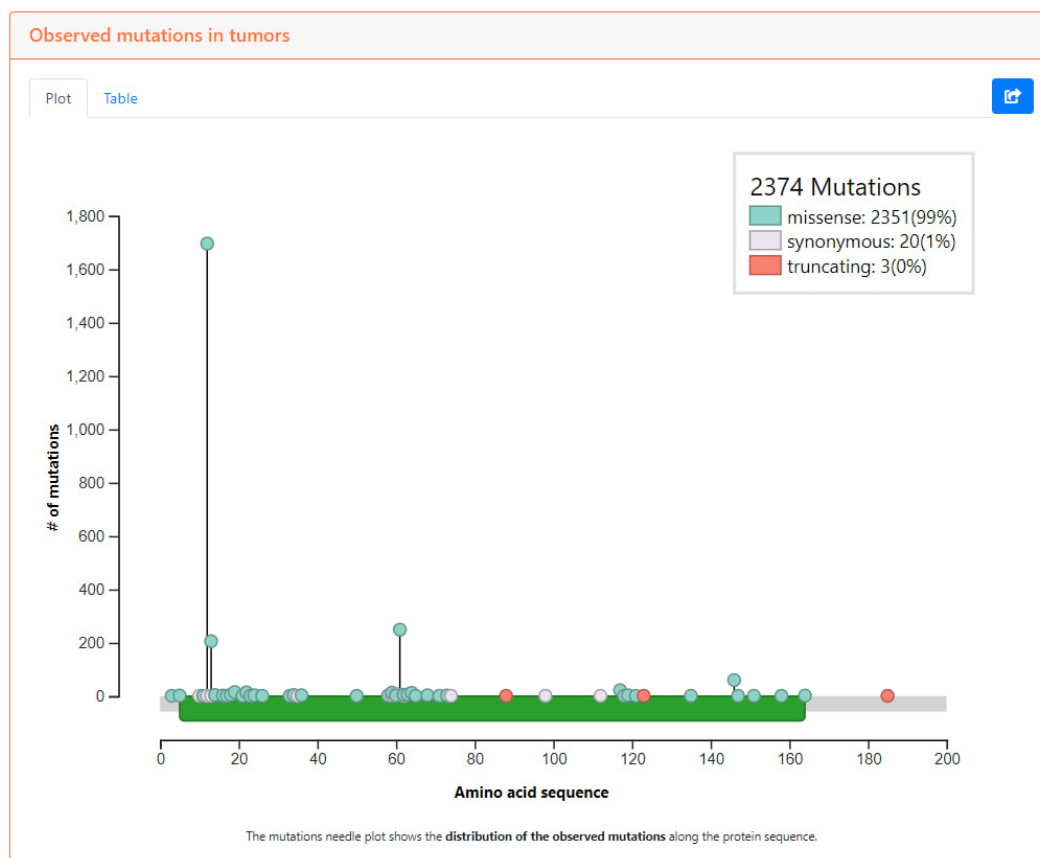


Рис. 1.2

Видим, что большинство мутаций - точечные (missense).

3) Посмотрим ассоциируемые типы рака на портале ICGC (и проекты их изучающие):

## Cancer Distribution

[Open in Advanced Search](#)

2,216 DONORS AFFECTED BY 990 MUTATIONS ACROSS 70 PROJECTS



Showing 1 - 10 of 70 projects

Filter table

Project	Site	Tumour Type	Tumour Subtype	# Donors affected ▼	# Mutations
PACA-CA	Pancreas	Pancreatic cancer	Ductal adenocarcinoma	247 / 268 (92.16%)	42
PACA-AU	Pancreas	Pancreatic cancer	Ductal adenocarcinoma	350 / 391 (89.51%)	43
PAAD-US	Pancreas	Pancreatic cancer	Adenocarcinoma	118 / 177 (66.67%)	13
LMS-FR	Mesenchymal	Soft Tissue cancer	Leiomyosarcoma	37 / 67 (55.22%)	52
COAD-US	Colorectal	Colon cancer	Adenocarcinoma	182 / 402 (45.27%)	32
BTCA-SG	Gall Bladder	Biliary Tract cancer	Gall bladder cancer / Cholangiocarcinoma	31 / 71 (43.66%)	25
READ-US	Colorectal	Rectal cancer	Adenocarcinoma	58 / 143 (40.56%)	15
MELO-AU	Skin	Skin cancer	Melanoma	68 / 183 (37.16%)	116
COCA-CN	Colorectal	Colorectal cancer	Adenocarcinoma, non-Western	118 / 321 (36.76%)	43
LUAD-US	Lung	Lung cancer	Adenocarcinoma	159 / 516 (30.81%)	14

Showing 10 rows

<<< < 1 2 3 4 5 > >>>

Рис. 1.3

Мутации гена KRAS часто обнаруживают при раке толстой кишки, раке поджелудочной железы и раке легких.

## 2 Поиск мутаций у пациента

Выберем пациента DO229010, участвовавшего в проекте COCA-CN. Цель этого проекта - исследование колоректального рака (colorectal cancer), который имеет сильную ассоциацию с мутациями в нашем гене.

Остальные гены выберем таким же способом. Например, на сайте IntOGen можно посмотреть [сводку](#) по колоректальному раку:

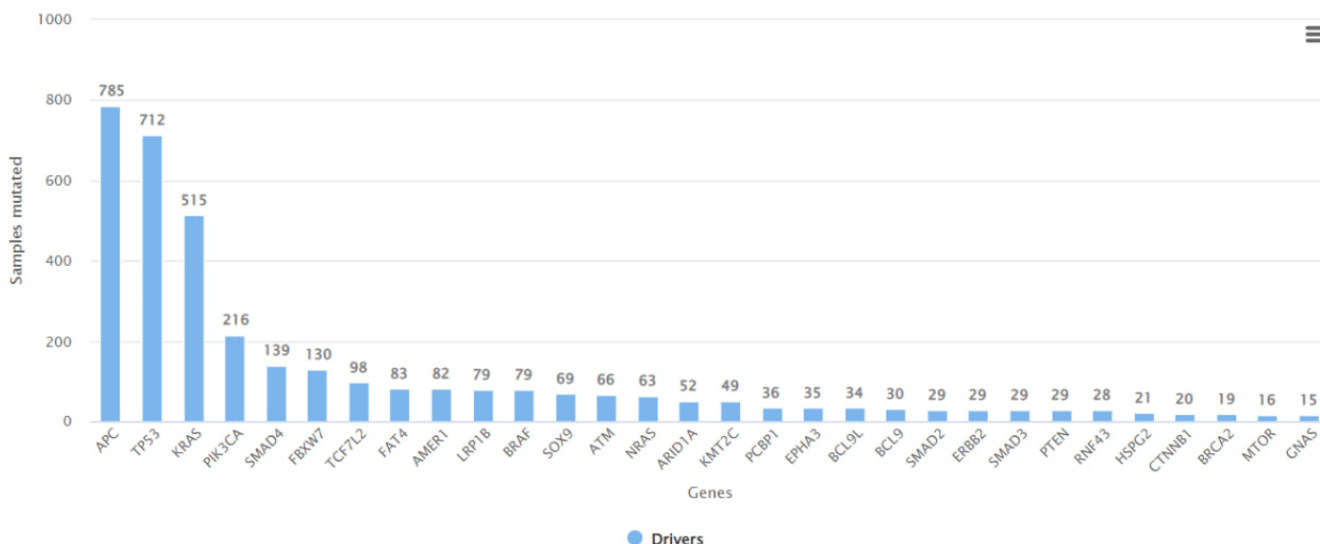


Рис. 2.1. Mutational cancer driver genes for colorectal cancer

Возьмем первые пять генов и выпишем их Ensembl ID для поиска по загруженному файлу:

Name	Ensembl ID
APC	ENSG00000134982
TP53	ENSG00000141510
KRAS	ENSG00000133703
PIK3CA	ENSG00000121879
SMAD4	ENSG00000141646

Заносим их в питон код и получаем:

```
Mutations in KRAS gene: 10
Mutations in TP53 gene: 51
Mutations in SMAD4 gene: 14
Mutations in APC gene: 65
Mutations in PIK3CA gene: 2
```

Рис. 2.2. Результаты выполнения [программы](#)

У нашего пациента оказались мутации во всех выбранных генах.