

如何在 NCBI 中查找基因的启动子序列

以下以人 ROCK1 基因为例，如何在 NCBI 中查找该基因的启动子序列：

(1) 在 NCBI 主页输入以下内容后确认：

Search results for ROCK1:

Name/Gene ID	Description	Location	Aliases	MM
ROCK1 ID: 5093	Rho associated coiled-coil containing protein kinase 1 [Homo sapiens (human)]	Chromosome 18, NC_000018.10 (20949740..21111851, complement)	P160ROCK, ROCK-1	601702
Rock1 ID: 19977	Rho-associated coiled-coil containing protein kinase 1 [Mus musculus (house mouse)]	Chromosome 18, NC_000084.6 (10064401..10161792, complement)	1110055K06Rik, Rock-1	
Rock1	Rho-associated coiled-coil containing protein	Chromosome 18, NC_005117.4	P160Rock.	

(2) 点击 ROCK1，找到以下信息。需要观察这个基因的方向（注意：ROCK1 基因的方向

与染色体 DNA 的方向是反的），点击：[genbank](#)

Genomic context for ROCK1:

- Location: 18q11.1
- Exon count: 34
- Chromosome: 18
- Location: NC_000018.10 (20949740..21111851, complement)

Genomic map showing ROCK1 gene structure and orientation.

(3) 看到如下信息，在右边的 from...to...表示该基因在染色体上碱基位置。

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide Search Help

Advanced

NCBI is phasing out sequence GI numbers in September 2016. Please use accession version! [Read more...](#)

GenBank

Homo sapiens chromosome 18, GRCh38.p7 Primary Assembly

NCBI Reference Sequence: NC_000018.10

[FASTA](#) [Graphics](#)

LOCUS NC_000018 162112 bp DNA linear CON 06-JUN-2016

DEFINITION Homo sapiens chromosome 18, GRCh38.p7 Primary Assembly.

ACCESSION NC_000018 REGION: complement (20949740..21113851) GPC_000001310

VERSION NC_000018.10 GI:568815580

DBLINK BioProject: [EUKM168](#)

DBLINK Assembly: [GCF_000001405.33](#)

KEYWORDS RefSeq.

SOURCE Homo sapiens (human)

ORGANISM [Homo sapiens](#)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 162112)

AUTHORS Nusbaum, C., Zody, M. C., Borowsky, M. L., Kanal, M., Kodira, C. D., Taylor, T. D., Whittaker, C. A., Chang, J. L., Cuomo, C. A., Dewar, K., Fitzgerald, M. G., Yang, X., Abouelleil, A., Allen, N. R., Anderson, S., Bloom, T., Bugalter, B., Butler, J., Cook, A., DeCaprio, D., Engels, R., Garber, M., Grirke, A., Hafez, N., Hall, J. L., Norman, C. H., Itoh, T.,

Change region shown

Whole sequence (abbreviated view)

Selected region

from: 20949740 to: 21113851

Update View

Customize view

Abbreviated view

Customize

Basic Features

Default features

Gene, RNA, and CDS features only

Display options

Show sequence

Show reverse complement

Update View

(4) 刚才我们看到, 该基因的方向与染色体 DNA 的方向是反的, 那我们从后面加上 2KB,

具体如下:

Whole sequence (abbreviated view)

Selected region

from: 20949740 to: 21113851

Update View

点击 Update View

(5) 往下看, 可以看到, 这个基因从 2001bp 开始了, 也就说明该基因的第一外显子上游 2kb 序列我们已经拿到, 而这 2000bp 加上第一外显子通常被认为是一个基因的启动子区。

```

features Annotated      :: Gene; mRNA; CDS; ncRNA
##Genome-Annotation-Data-END##
FEATURES               Location/Qualifiers
  source                1..164112
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="18"
  gene                  2001..164112
                        /gene="ROCK1"
                        /gene_synonym="P160ROCK; ROCK-I"
                        /note="Rho associated coiled-coil containing protein
                        kinase 1; Derived by automated computational analysis
                        using gene prediction method: BestRefSeq, Gnomon."
                        /db_xref="GeneID:6093"
                        /db_xref="HGNC:HGNC:10251"
                        /db_xref="HPRD:03414"
                        /db_xref="MIM:601702"
  mRNA                  join(2001..3034,43239..43320,63972..64072,64623..64760,
                        68385..68560,69666..69750,71143..71287,71617..71755,
                        74289..74380,84917..85076,90172..90232,93613..93701,
                        98373..98421,105658..105793,107062..107153,107255..107501,
                        120915..121021,122526..122676,126742..126902,
                        129317..129501,131020..131089,133848..133942,
                        143339..143504,144644..144737,144992..145080,
                        145912..146100,146776..146935,153646..153716,
                        153924..154012,158607..158685,158808..159069,
                        160067 160274 162465 164112)

```