

PublisherInfo		
PublisherName	:	BioMed Central
PublisherLocation	:	London
PublisherImprintName	:	BioMed Central

## Prediction of intron splice sites

ArticleInfo		
ArticleID	:	3597
ArticleDOI	:	10.1186/gb-2000-1-1-reports223
ArticleCitationID	:	reports223
ArticleSequenceNumber	:	88
ArticleCategory	:	Web report
ArticleFirstPage	:	1
ArticleLastPage	:	4
ArticleHistory	:	RegistrationDate : 1999-11-18 Received : 1999-11-18 OnlineDate : 2000-3-17
ArticleCopyright	:	BioMed Central Ltd2000
ArticleGrants	:	
ArticleContext	:	130591111

Todd Richmond

## Abstract

The NetGene2 server uses a neural network combined with a rule-based system to predict intron splice sites in *Arabidopsis thaliana*, *Caenorhabditis elegans* and humans.

## Content

The NetGene2 server uses a neural network combined with a rule-based system to predict intron splice sites in *Arabidopsis thaliana*, *Caenorhabditis elegans* and human genes. The server returns a variety of information. First are tables of the predicted donor and acceptor sites, as well as branch points (for *A. thaliana* only), in both the (+) strand and the (–) strand. Included in these tables are the position, the phase, the predicted confidence level, and the 20 base pairs around the predicted splice site. Following these tables are a series of graphics showing the predicted coding regions, with the donor and acceptor sites drawn below. The graphics can be downloaded in GIF format (manually from the original output page) or in Postscript format (as a separate file). A complete scoring table for either strand can be downloaded as well.

## Navigation

The submission page is very simple and easy to use. There are two options for inputting sequence. You can either read a sequence in FASTA format from a local drive or paste your sequence into a dialog box. Then choose your organism by clicking a button and pressing submit. There are links at the bottom of the submission page that point to abstracts of papers, instructions, output format, mail server and performance report.

## Reporter's comments

## Timeliness

There is no indication of when the site was last updated.

## Best feature

The table includes the 10 base pairs preceding a predicted site, and the 10 base pairs following the site. This makes it very easy to locate the exact position of the splice site in your sequence.

## Worst feature

The graphical representations of the predicted coding sequences are much too small to be of practical use. Although you can download the Postscript files (which theoretically can be scaled to any size you want), most Macintosh and Windows users do not have the ability to view or manipulate Postscript files easily. In addition, these files are compressed with GZIP (a Unix compression format), which may pose a problem for some users.

## Wish list

It would be nice to be able to select whether or not the images were generated, or include an option to download a single archive file with all of the images (in a more accessible format with more detail). There are some potential problems, for example large sequences generate a large number of image files, which may cause memory problems for some browsers. I have run into problems printing when the data expires from the cache, and I had to resubmit the sequence in order to get a printout. For very large sequences, it is probably best to use the e-mail server instead of the web server.

## Related websites

There are a number of sites that offer splice site and gene prediction, including [GENIO/splice](#) (human only) and [BCM GeneFinder](#).

## Table of links

[NetGene2 world wide web prediction server](#)

GENIO/splice

BCM GeneFinder

## References

1. NetGene2 world wide web prediction server.