AutoPrimer3

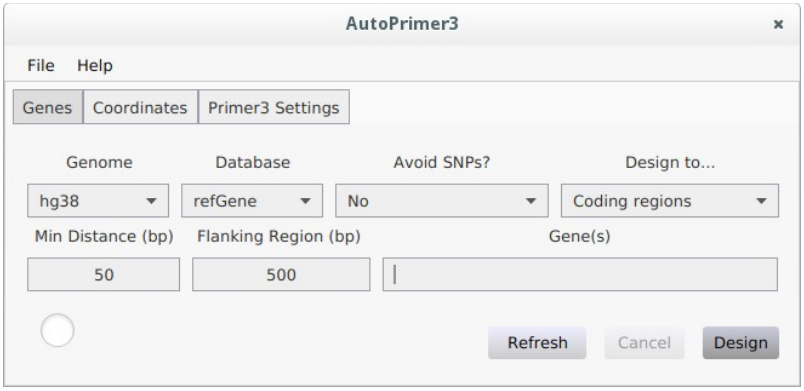
AutoPrimer3 retrieves gene information, DNA sequences and SNP information from the UCSC genome browser and uses Primer3 (http://primer3.sourceforge.net/) to automatically design primers to genes or genomic coordinate targets.

The default settings should be suitable for the design of primers for use in the amplification and sequencing of small PCR products (< 500 base pairs in length).

Note: An internet connection is required by the program to retrieve database information from the UCSC genome browser. AutoPrimer3 will not work offline.

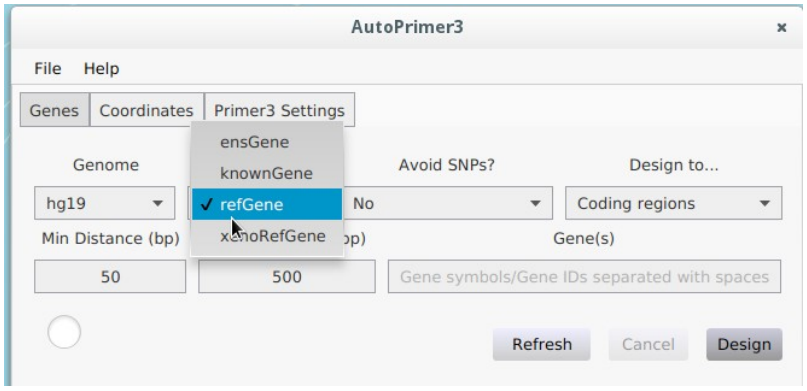
Designing to Genes or Coordinates

At startup the 'Genes' tab is shown.



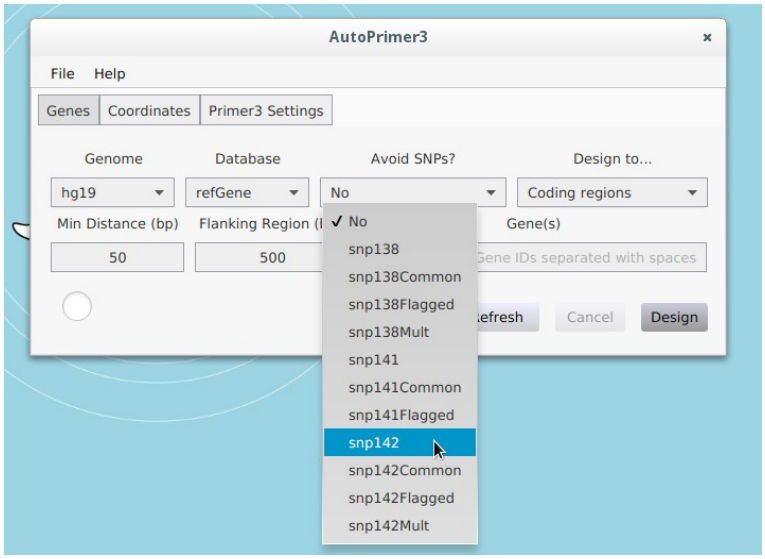
Select the desired genome for the species you want to design primers for using the **'Genome'** menu.

Hover the mouse cursor over the loaded genome for its full description



Once the genome information is loaded you may select the desired **gene database**. If there is an error loading the genome information you may try using the 'Refresh' button to attempt reloading from the UCSC database.

And, if available, you may select a **SNP database** in order to prevent primers from being designed on top of any coordinates that match SNPs in the given database.

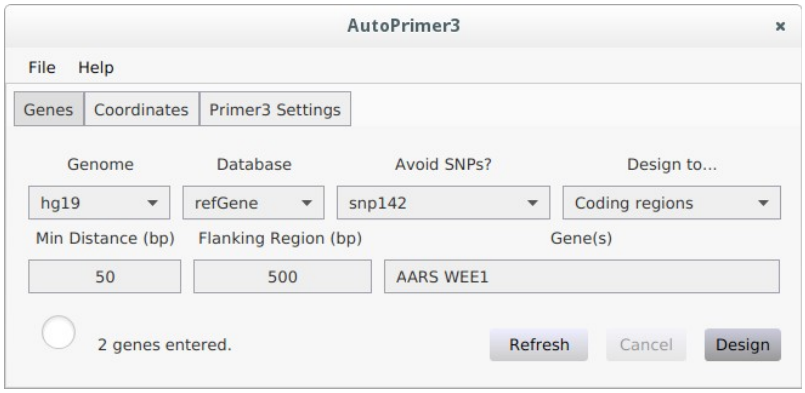


The **'Design to...'** menu allows you to choose between designing only to coding regions or to all exons of gene targets.

The **'Min Distance'** field lets you choose the minimum distance in base pairs (bp) between targets and primers.

The **'Flanking Region'** field tells AutoPrimer3 how many bp of flanking DNA to retrieve either side of target regions to use for its primer designs.

Enter one or more gene symbols or transcript Ids in the **'Gene(s)'** input box and either press Enter or click the **'Design'** button to start designing primers. If a gene symbol is used primers will be designed to all transcripts found for the given gene in the relevant database.



Alternatively, you may enter coordinates to design primers to as shown below using the input box on the **'Coordinates'** tab. The same options are available for adjusting genome, SNP databases, minimum distance from primer to target and amount of flanking sequence to use as for the 'Genes' tab. Coordinates can be typed into the **'Regions/Coordinates'** input box like below. Alternatively, you may use the **'Load File'** button to load up to 100 regions from a BED or VCF file or a text file with regions specified as intervals (i.e. in the format chr1:1000-2000).

This example also contains optional names for each region, specified as an extra field separated from the region by either space or tab characters. This also correlates to the ID field of VCF format or the Name field for BED format coordinates. This behaviour can be disabled by unchecking the 'Use region names' checkbox. If this feature is disabled or names are not provided for regions the primers will be named 'Region_n' where 'n' is 1 for the first region entered, 2 for the second etc..

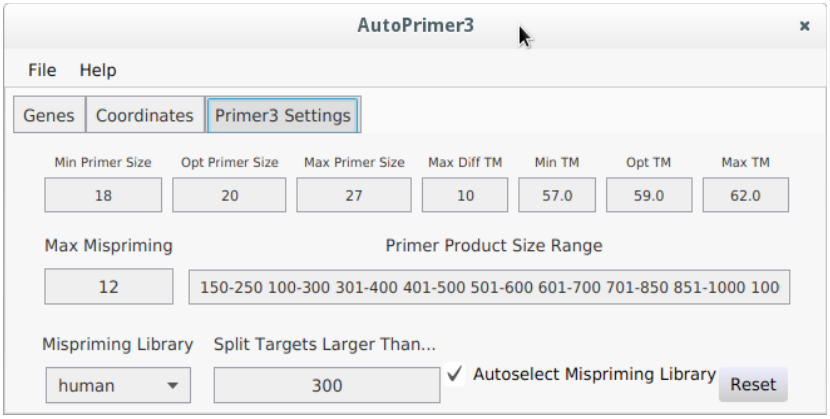
Region output will be in coordinate order and nearby regions will be merged as per the settings in the 'Primer3 Settings' tab.

The 'Regions/Coordinates' input can be cleared using the **'Clear'** button.

Adjusting Primer Design Settings

The **'Primer3 Settings'** tab provides various options to be passed to the Primer3 (http://primer3.sourceforge.net/) program when designing primers. The first two rows of options are options specific to the Primer3 tool, and more information can be found in Primer3's own documentation.

AutoPrimer3 makes three mispriming libraries available for use with Primer3: human, rodent and drosophila. By default, when you select a genome if an appropriate mispriming library is available it will be automatically selected. This behaviour can be turned off by unchecking the **'Autoselect Mispriming Library'** checkbox and you may override the programs choice for a given genome by manually selecting a different mispriming library.



Also of note, is the **'Split Targets Larger Than...'** input field. The default value of 300 is chosen such that with the default **'Min Distance'** value of 50 PCR products should generally be smaller than 500 bp. This is chosen as a sensible default in the expectation that most users will want PCR products of this size due to their relative ease of amplification and suitability for Sanger sequencing of the full length of the target sequence. If the user changes this value they may also want to change the values used in the **'Primer Product Size Range'** field, which has a default maximum of 2000 bp.

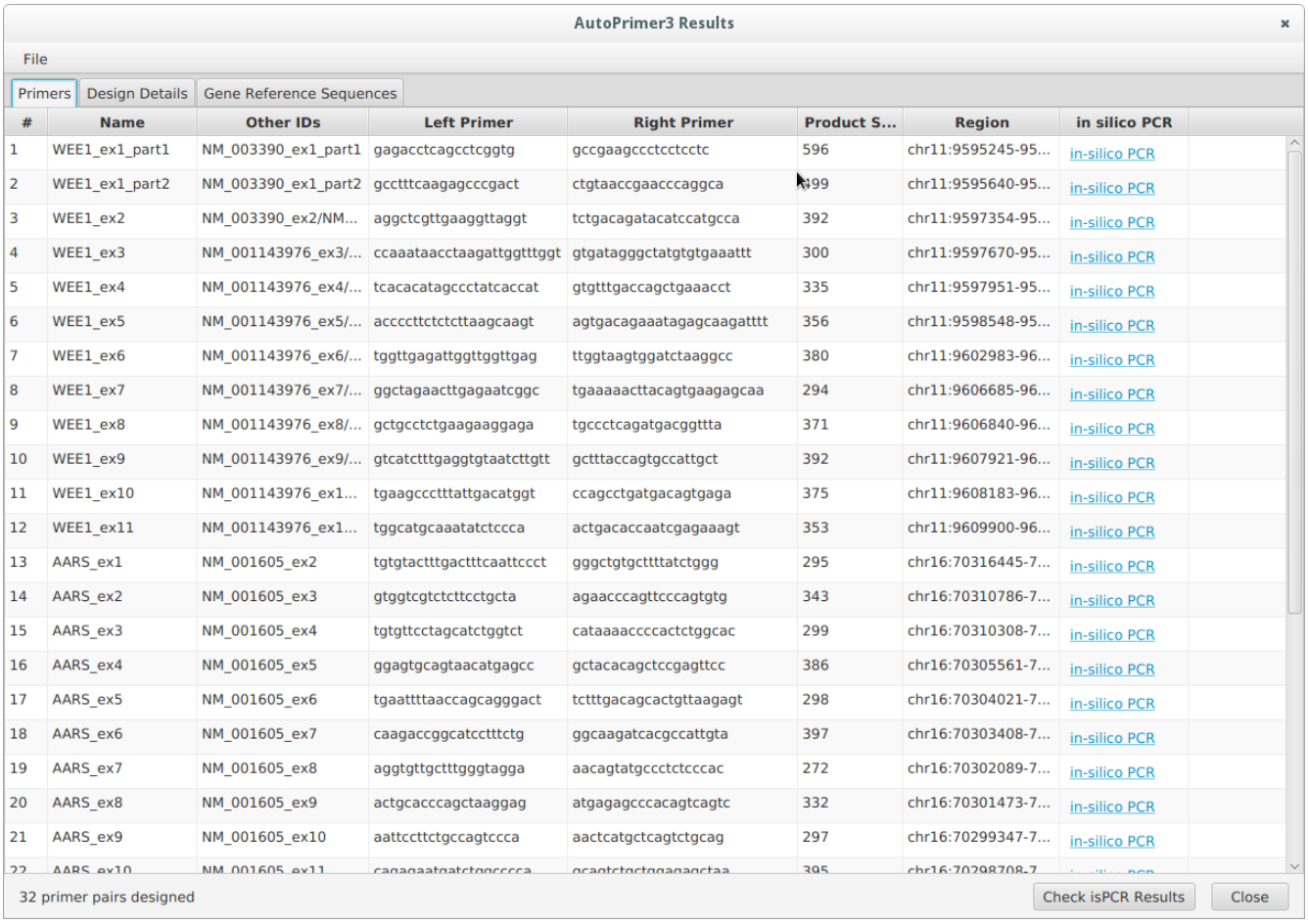
Results

Once designs are completed a new window will appear with the results. You may use the associated **'File'** menu to **save** the output of each tab. Note that if choosing .xlsx output for saving primers the output consists of two worksheets, one giving the primers in a list (one row per primer, which may be useful for copying and pasting into certain primer purchasing websites) and one giving more detailed output like that shown by the AutoPrimer3 Results window.

The **'Primers'** tab of the results window contains a table of the primer targets and designs. Most of the fields are self-explanatory but the final column provides a hyperlink to perform in-silico PCR at the UCSC genome browser in order to check for the specificity of the primer

pairs.

You may also choose to use the **'Check isPCR Results'** button to perform an automatic check of the number of predicted PCR products in your target genome from UCSC's in-silico PCR program. Once complete, a new column will be added with the number of predicted PCR products for each primer pair and a warning will be displayed if any primer pair appears to amplify more than one genomic region. Please note that sometimes the in-silico PCR tool appears to fail to find any matches for a pair when one of the primers is short, but this does not necessarily mean that the primers are of no use.

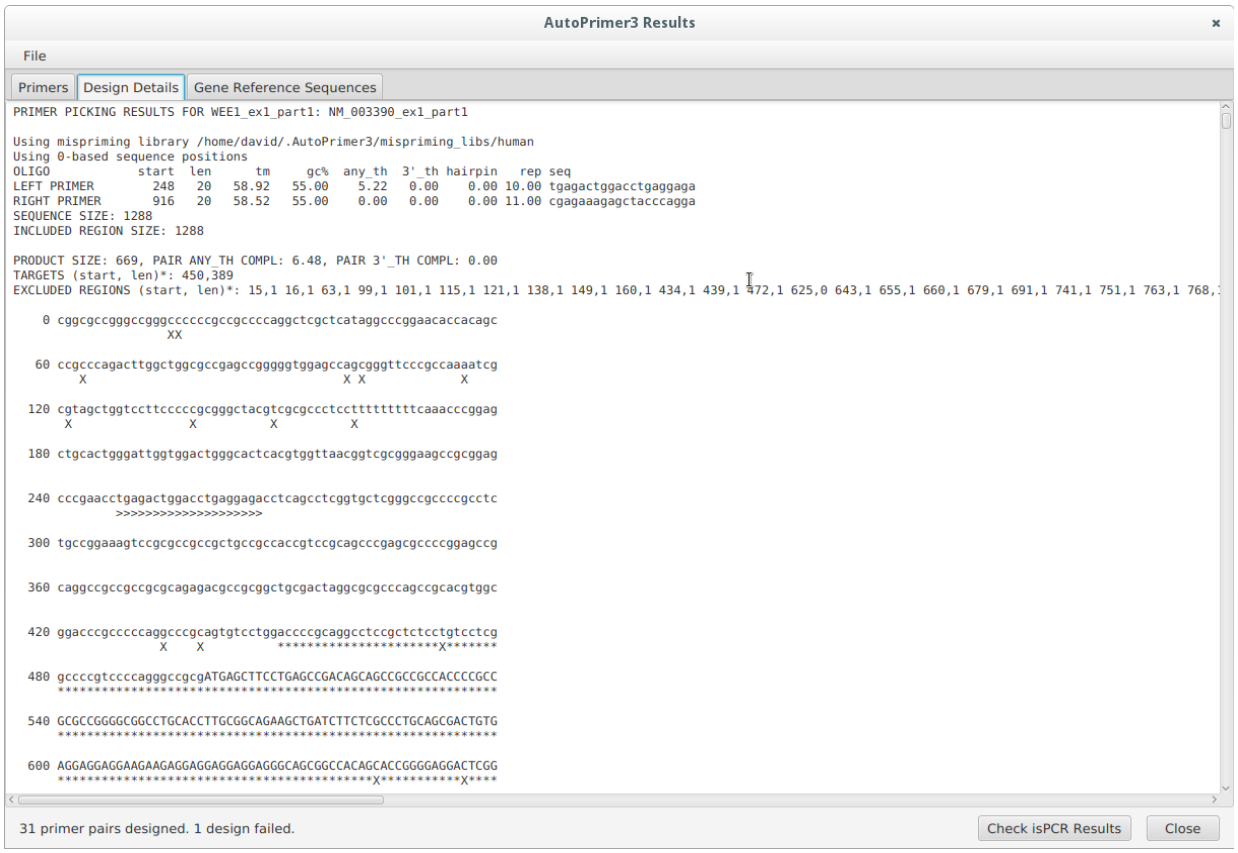


#	Name	Other IDs	Left Primer	Right Primer	Product S...	Region	in silico PCR
1	WEE1_ex1_part1	NM_003390_ex1_part1	gagacctcagcctcggtg	gccgaagccctcctcctc	596	chr11:9595245-95...	in-silico PCR
2	WEE1_ex1_part2	NM_003390_ex1_part2	gcctttcaagagcccgact	ctgtaaccgaacccaggca	399	chr11:9595640-95...	in-silico PCR
3	WEE1_ex2	NM_003390_ex2/NM...	aggctcgttgaaggttaggt	tctgacagatacatccatgcca	392	chr11:9597354-95...	in-silico PCR
4	WEE1_ex3	NM_001143976_ex3/...	ccaaataacctaagattggtttggt	gtgatagggctatgtgtgaaattt	300	chr11:9597670-95...	in-silico PCR
5	WEE1_ex4	NM_001143976_ex4/...	tcacacatagccctatcaccat	gtgtttgaccagctgaaacct	335	chr11:9597951-95...	in-silico PCR
6	WEE1_ex5	NM_001143976_ex5/...	acccctttctctcttaagcaagt	agtgacagaaatagagcaagatttt	356	chr11:9598548-95...	in-silico PCR
7	WEE1_ex6	NM_001143976_ex6/...	tggttgagattggttggttgag	ttggtaagtggatctaaggcc	380	chr11:9602983-96...	in-silico PCR
8	WEE1_ex7	NM_001143976_ex7/...	ggctagaacttgagaatcggc	tgaaaaacttacagtgaagagcaa	294	chr11:9606685-96...	in-silico PCR
9	WEE1_ex8	NM_001143976_ex8/...	gctgcctctgaagaaggaga	tgccctcagatgacggttta	371	chr11:9606840-96...	in-silico PCR
10	WEE1_ex9	NM_001143976_ex9/...	gtcatctttgaggtgtaatcttgtt	gctttaccagtgccattgct	392	chr11:9607921-96...	in-silico PCR
11	WEE1_ex10	NM_001143976_ex1...	tgaagccctttattgacatggt	ccagcctgatgacagtgaga	375	chr11:9608183-96...	in-silico PCR
12	WEE1_ex11	NM_001143976_ex1...	tggcatgcaaatatctccca	actgacaccaatcgagaaagt	353	chr11:9609900-96...	in-silico PCR
13	AARS_ex1	NM_001605_ex2	tgtgtactttgactttcaattccct	gggctgtgcttttatctggg	295	chr16:70316445-7...	in-silico PCR
14	AARS_ex2	NM_001605_ex3	gtggtcgtctcttcctgcta	agaacccagttcccagtgtg	343	chr16:70310786-7...	in-silico PCR
15	AARS_ex3	NM_001605_ex4	tgtgttcctagcatctggtct	cataaaaccccactctggcac	299	chr16:70310308-7...	in-silico PCR
16	AARS_ex4	NM_001605_ex5	ggagtgcagtaacatgagcc	gctacacagctccgagttcc	386	chr16:70305561-7...	in-silico PCR
17	AARS_ex5	NM_001605_ex6	tgaattttaaccagcagggact	tctttgacagcactgttaagagt	298	chr16:70304021-7...	in-silico PCR
18	AARS_ex6	NM_001605_ex7	caagaccggcatcctttctg	ggcaagatcacgccattgta	397	chr16:70303408-7...	in-silico PCR
19	AARS_ex7	NM_001605_ex8	aggtgttgctttgggtagga	aacagtatgccctctcccac	272	chr16:70302089-7...	in-silico PCR
20	AARS_ex8	NM_001605_ex9	actgcacccagctaaggag	atgagagcccacagtcagtc	332	chr16:70301473-7...	in-silico PCR
21	AARS_ex9	NM_001605_ex10	aattccttctgccagtccca	aactcatgctcagtctgcag	297	chr16:70299347-7...	in-silico PCR
22	AARS_ex10	NM_001605_ex11	cagagaatgatctggcccca	gcagtctgctggagagctaa	395	chr16:70298708-7...	

Note that for simplicity, exons for each gene (i.e. those in the 'Name' column) are counted sequentially from the first relevant exon (e.g. first coding exon if designing to coding regions only). Exon numbering for transcripts counts from the first exon of that transcript regardless of whether it is coding or non-coding. As such, numbering of exons in transcripts should be accurate while exon numbering at the whole gene level is only an approximation due to the various complexities that can arise from multiple splice isoforms and merely serves as a useful shorthand.

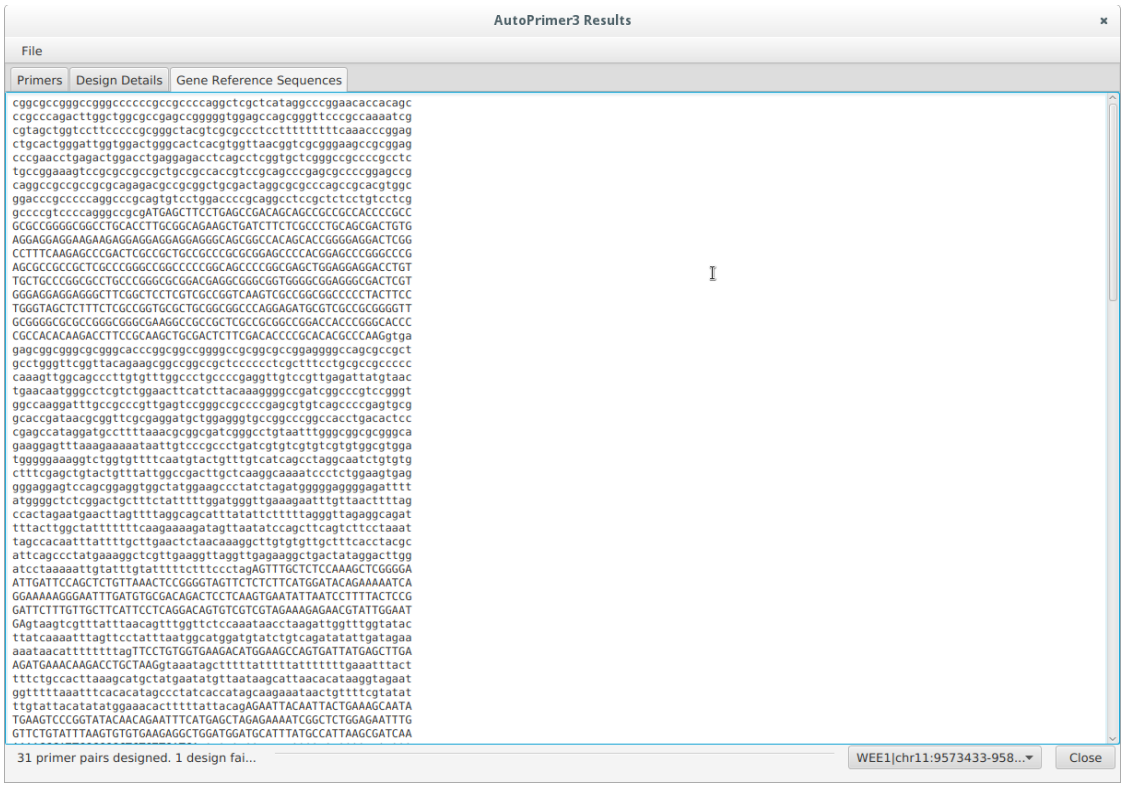
The **'Design Details'** tab (see below) provides output from primer3 for each design attempted. This output should be familiar to users of Primer3 on the web. AutoPrimer3 will

put the user's target sequence (e.g. an exon) in uppercase and the rest of the sequence in lowercase. Sequence with asterisks underneath consists of the target sequence plus the minimum amount of sequence between the target and primer as specified by the **'Min Distance'** input box. Any known SNPs that are to be avoided are marked underneath with an 'X'. Note that each design also supplies alternative primer pairs at the bottom.



When designing primers to genes, the **'Gene Reference Sequences'** tab (see below) provides the DNA sequence for your target exons plus flanking sequence either side (the length of which is determined by the value used in the **'Flanking region'** field when the primers were designed). If there were multiple target regions (either multiple genes/transcripts specified or genes that map to multiple genomic regions) the choice box towards the bottom right of the window can be used to switch between the relevant reference sequences.

Target sequences (e.g. exons or coding regions) are shown in uppercase while non-target sequence (e.g. introns) are shown in lowercase.



Credit

AutoPrimer3 uses primer3 (http://primer3.sourceforge.net/). AutoPrimer3 was written by David A. Parry and is available from:

https://github.com/gantzgraf/autoprimer3

or alternatively:

https://sourceforge.net/projects/autoprimer3/

It was originally available as a perl script and a perl/perl + objective C based GUI application for Windows and Mac OS X. This version is a complete rewrite using java and is available for Windows, Mac OS X and linux.

If you use AutoPrimer3 for primer designs that are used in published work, please cite the URL 'https://github.com/gantzgraf/autoprimer3'.

License