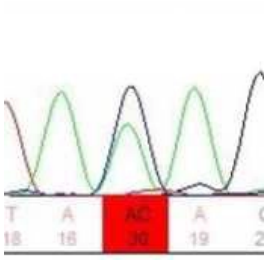
May 06, 2024

Genotyping of a candidate variant by PCR and Sanger sequencing



DOI
dx.doi.org/10.17504/protocols.io.rm7vzxp32gx1/v1

Cecile Grohs[1]

[1]Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350, Jouy-en-Josas, France

Cecile Grohs
Université Paris-Saclay, INRAE, AgroParisTech, GABI, 78350, ...

OPEN ACCESS

DOI: dx.doi.org/10.17504/protocols.io.rm7vzxp32gx1/v1

Protocol Citation: Cecile Grohs 2024. Genotyping of a candidate variant by PCR and Sanger sequencing. **protocols.io** https://dx.doi.org/10.17504/protocols.io.rm7vzxp32gx1/v1

Protocol status: Working
We use this protocol and it's working

Created: February 01, 2024

Last Modified: May 06, 2024

Protocol Integer ID: 94550

Keywords: PCR, genotyping, Sanger sequencing, sequences, SNP, INDEL, genetic mosaicism

Abstract

This procedure describes a classic process for genotyping a simple SNP or small INDEL by PCR followed by Sanger sequencing, which can be easily analysed. It is used for the validation or invalidation of a candidate variant that has been identified from a whole genome sequence. It combines several steps using online tools and a free downloadable software. It is easy to use, inexpensive and has the advantage of displaying a large number of samples in the same window, aligned simultaneously to a reference sequence. Alleles are detected automatically.
It is also possible to inspect bases that have not been identified as polymorphic by the software, particularly in the case of somatic mosaicism. If so, it is possible to distinguish the two alleles and obtain a rough estimate of the relative allelic proportions.

Guidelines

This classic process combines several steps using online tools and a free downloadable software.

Before start

To set up a genetic test, it is recomended to have acces to an available reference genome.

As this procedure is designed to describe the validation of a candidate variant, we assume that a candidate mutation has already been identified (i.e. its position in the reference genome and the gene to which it relates).

Download NovoSNP software.

Recover complete sequence of the gene of interest in the species of interest

10m

1 First select the genome assembly of the species you wish to genotype in a genome browser (e.g. https://genome-euro.ucsc.edu/cgi-bin/hgGateway). Then, select the gene whose DNA you want to retrieve. 5m

2 Use the appropriate tool to obtain the DNA sequence in a fasta format (here Menu: **View**, Submenu: **DNA**) and copy-paste the sequence in a text file, with no formatting such as line breaks. This file will be needed for the following steps. It will be your reference sequence. Make sure you name the file correctly so that you can recognise it later, especially if you are working on several genes at once. 5m

Recover the surrounding sequence of the variant of interest for PCR purposes

10m

3 Select the exact position on the chromosome you would like to genotype in the genome browser, and zoom out to obtain a sequence of about 500-600 base pairs around the mutation of interest that is suitable for PCR followed by Sanger sequencing. 5m

4 Use the tool mentioned above to obtain the DNA sequence. Then save the selected sequence to a new file and copy and paste the sequence into the online Primer3 software (see references). 5m

Select primers for PCR purposes

3d

5 In **Primer3**, it is recommended to mark the variant to check that the primers chosen are correctly positioned (at least 70 base pairs on either side of the variant to be genotyped) so that the sequences can be unambiguously read. 1m

6 Choose a product size range of 400-600bp. Standard options are sufficient in most cases. 1m

Expected result

- A best choice primer pair with specifications as in the exemple below

OLIGO ***start*** ***len*** ***tm*** ***gc%*** ***any*** ***3'*** ***seq***
LEFT PRIMER *55 21 58.94 47.62 8.00 1.00 TCCGCTAATTAGGTCCTGAGA*
RIGHT PRIMER *555 20 60.26 60.00 3.00 0.00 GCTGGTAGGGACACAGAGGA*
SEQUENCE SIZE: 601
INCLUDED REGION SIZE: 601

PRODUCT SIZE: 501, PAIR ANY COMPL: 4.00, PAIR 3' COMPL: 2.00
TARGETS (start, len): 331,18*

- The sequence of interest, the position of each primer, around the selected variant

```
  1 GAGGTTATCGCCTGAAACTAAAGGTTTATTCACTAACTAATCAGACTTTAAGTGTCCGCT
                                                          >>>>>>

 61 AATTAGGTCCTGAGACATTCAATAAAGAAATTCAACAGGAAATTACCTATCGGTGCATTC
    >>>>>>>>>>>>>>>

121 AGCCTGAATTTTTATTCTACTTCAGAAAATGTCAAGGTCTGTTATAAGACTTGAGAATAA

181 CAAATGTGGCTTTGACTTTTTCAAAATCCAAAAGGGAAAAATGAACTATAAAGGGGGGAA

241 AAAGCCTTCTTGACTCTTCATAGGCAAATGGAAAGAGGGAAAGAAGAAAAAGTCAGACCC

301 AAGAGGATTATGTGATATGTTTATGAGATTAGACTTGTCTTATAGAAGATTTCATGAAAA
                                  ******************

361 CCATTTAAGTTTTGCTTCAAAATGGACTCGAGGCCTTTAAAGATTACAGCTGGTTAGTAT

421 GTCTTCTGACCCTCAGACAATGTCGGGGTGGTGTGGTTAAATCACAGGCATTATTTATGC

481 AGTGAAAAGATGTAGTCGTGTAAGAAGGTTGTATTTTCAGGTTAGGTTAACAACTTCCTC
                                                          <<<<<

541 TGTGTCCCTACCAGCTCAGCAGTTTTTCCACCTTACTCACTTCCCATTCATCCAGTTCCC
    <<<<<<<<<<<<<<<

601 T
```

>>>> indicates the position of primers, ***** indicates the position of variants

- Additional oligos

7 Make sure you save your selections to the new file you created. 1m

8 Order one or more primer pairs from your usual supplier. 3d

Setting up PCR conditions 3h

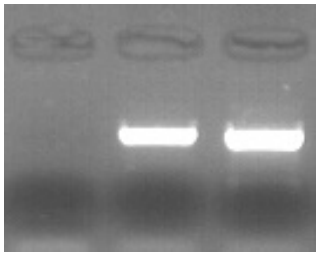
9 Use a standard PCR kit and follow the manufacturer's recommendations. Select two DNA samples from the species of interest as positive controls and add a negative control tube (containing water instead of DNA). 15m

10 Perform the PCR using the conditions recommended by PRIMER3. Alternatively, you can use conditions that are accepted as the most common in your laboratory (i.e. considered robust). If these conditions work, you can perform several tests in the same experiment. 1h 30m

11 Load PCR products onto a 2% agarose gel with BET in the presence of a size marker. Migrate at 110 Volts for approximately 00:25:00 . 25m

Expected result

The PCR product obtained must be unique and its size must match that predicted by primer3. The expected band must be clearly visible.



Example of PCR test

Prepare samples for Sanger sequencing 3d

12 Select the samples to be genotyped: ideally, take 1 to 3 DNA samples corresponding to the different genotypes you expect. For a recessive disease, you will expect

- homozygote alternative, which should be affected individuals, 30m

- heterozygous witch should be carriers (e. g. parents of cases)
- homozygous wild type.

13 Perform the PCR using the conditions developed in the previous step in a final volume of at least 35µL to obtain sufficient PCR product for Sanger sequencing in either direction. 3h

14 The amplified and gel-checked PCR fragments and primers are then sent to any supplier for Sanger sequencing. 2d

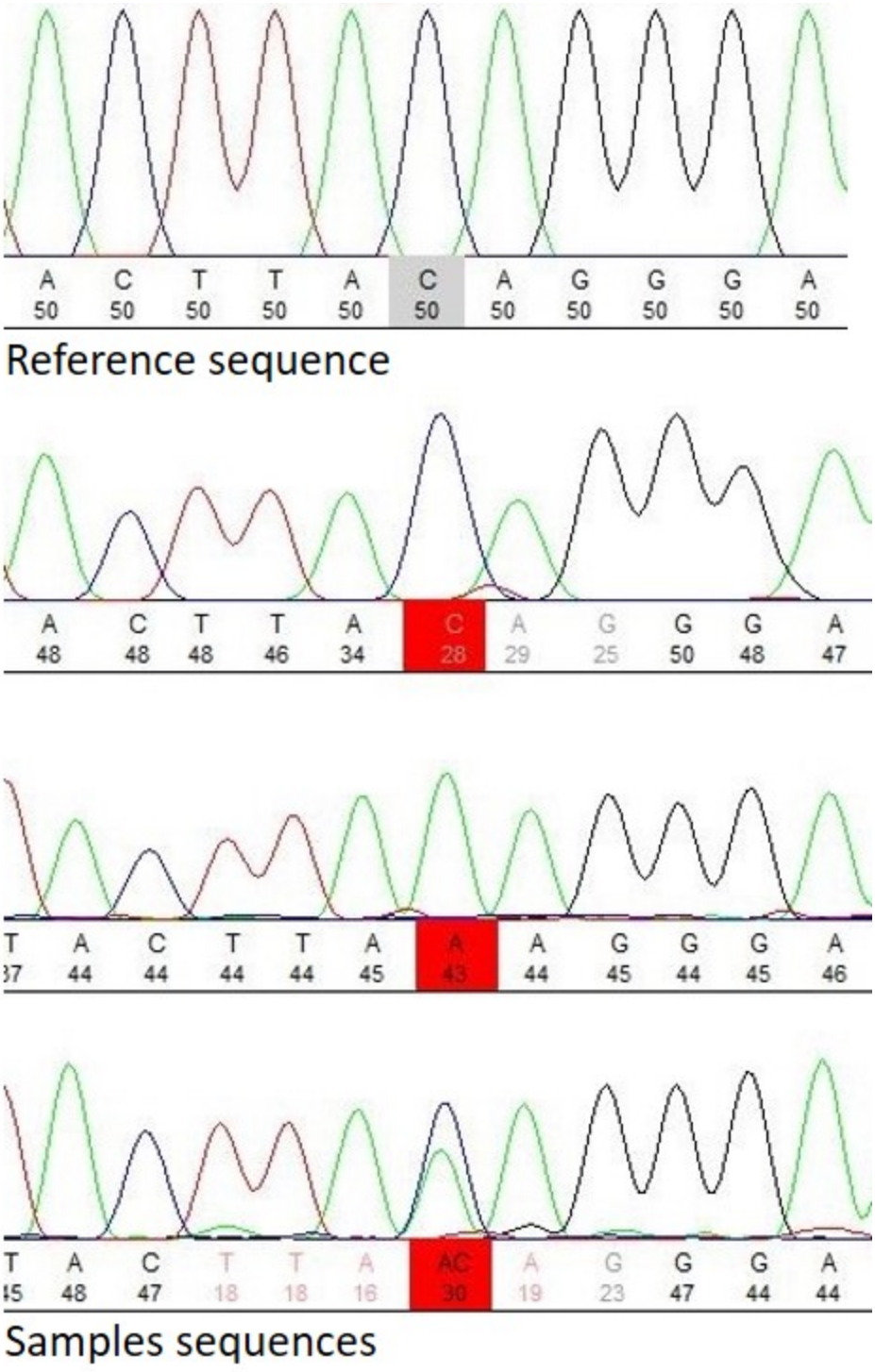
Genotyping with NovoSNP 12m

15 In NovoSNP (see references), 10m
First load the gene reference sequence you prepared in step 2
Then load the files from the Sanger sequencing supplier into NovoSNP and let the software process the files.

16 The PCR fragments are aligned to the reference sequence. 2m
Novo SNP detects putative SNPs and small INDELs, and provides a quality score for each detected position.

17 Find the position to be studied and check that the sequences match the phenotype and the expected genotype perfectly.

Expected result



Three different genotypes are displayed here: homozygous reference, homozygous alternative, and heterozygous.

18 If the sequences match the phenotype and the expected genotype perfectly, the mutation and the test are validated.
Any new unknown sample can be genotyped using this method.

Note

If you are analysing a mosaic animal (i.e. one that carries the mutation in a small proportion of its cells), the software may not annotate the two alleles, but it is possible to distinguish them visually and estimate their relative proportions.

Protocol references

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M and Rozen SG. **Primer3--new capabilities and Interfaces**. Nucleic Acids Res. 2012 Aug 1;40(15):e115.

Weckx S, Del-Favero J, Rademakers R, Claes L, Cruts M, De Jonghe P, et al. **novoSNP, a novel computational tool for sequence variation discovery.** Genome Res. 2005;15:436–42.