OligoHeatMap (OHM): an online tool to estimate and display hybridizations of oligonucleotides onto DNA sequences

Olivier Croce[1,*], François Chevenet[2] and Richard Christen[1]

[1]University of Nice Sophia-Antipolis and CNRS. UMR 6543, Virtual Biology Lab. Institute of Developmental Biology and Cancer. Parc Valrose. 06108 Nice and [2]Genetics and Evolution of Infectious Diseases, UMR 2724 CNRS/IRD-IRD, 911 avenue Agropolis, B.P. 64501, 34394 Montpellier Cedex 5, France

ABSTRACT

The efficiency of molecular methods involving DNA/DNA hybridizations depends on the accurate prediction of the melting temperature (T_m) of the duplex. Many softwares are available for T_m calculations, but difficulties arise when one wishes to check if a given oligomer (PCR primer or probe) hybridizes well or not on more than a single sequence. Moreover, the presence of mismatches within the duplex is not sufficient to estimate specificity as it does not always significantly decrease the T_m. OHM (OligoHeatMap) is an online tool able to provide estimates of T_m for a set of oligomers and a set of aligned sequences, not only as text files of complete results but also in a graphical way: T_m values are translated into colors and displayed as a heat map image, either stand alone or to be used by softwares such as TreeDyn to be included in a phylogenetic tree. OHM is freely available at http://bioinfo.unice.fr/ohm/, with links to the full source code and online help.

INTRODUCTION

Design and use of oligomers for PCR and hybridization is a common practice in molecular biology. There are two basic approaches when searching for specific oligomers. The most common practice is to look for their description in scientific papers. The other possibility is a *de novo* design, using one of the numerous stand-alone softwares or web tools, such as Primers 3 (1), Primique (2) and Oligo 7 (3). In many cases, primers specific for group of sequences are looked for, while related (similar) non-target sequences should not be recognized. Some design softwares consider such nontarget sequences, but only check for exact matches, or simply use blast to estimate specificity. The problem is that some mismatches do not

significantly decrease the melting temperature (T_m) resulting in specificity problems. Also, the goal is to calculate and compare the theoretical T_m for every sequence including nonperfect matches. Finally, in the presence of many primers and sequences, the end result presented as a text file can be very difficult to analyze. OligoHeatMap (OHM) not only provides such complete text files of T_m values and additional information, but also graphical representations that are easier to analyze. OHM is a web-application (Javascript, PHP). The main window contains four frames: one to load oligomers, one to load aligned sequences, one for providing chemical options and one for outputs. The T_m is calculated using the nearest-neighbor method and the thermodynamic data published by SantaLucia *et al.* (4–6). The salt correction method used is based on Wetmur (7). Because the goal of OHM is to compare oligomers and not to precisely calculate T_m, users are encouraged to have other estimation of T_m by using tools like Melting (8), OligoCalc (9) or DNAmate (10), since different methods or tools can result in significant differences in the estimates of the T_m of the duplex (11,12).

INPUT DATA

Sequence(s) of oligonucleotide(s) is the only required input, either through manual input or through upload of a file (clustal, fasta, mase or tabulated format). A set of aligned sequences can also be provided (manual entry or file upload), in this case at least one sequence should contain the primer(s). Sequences have to be aligned because OHM extracts and compares the areas of hybridization (see the online help for more information on how to align sequences). If no sequence is provided, calculated T_m will be only for self-hybridization (dimers of primers) and invert-complement (primer on target). Finally, options like DNA and salt concentrations can be adjusted.

---

*To whom correspondence should be addressed. Tel: +33 4 92 07 69 47; Fax: +33 4 92 07 69 47; Email: croce@unice.fr
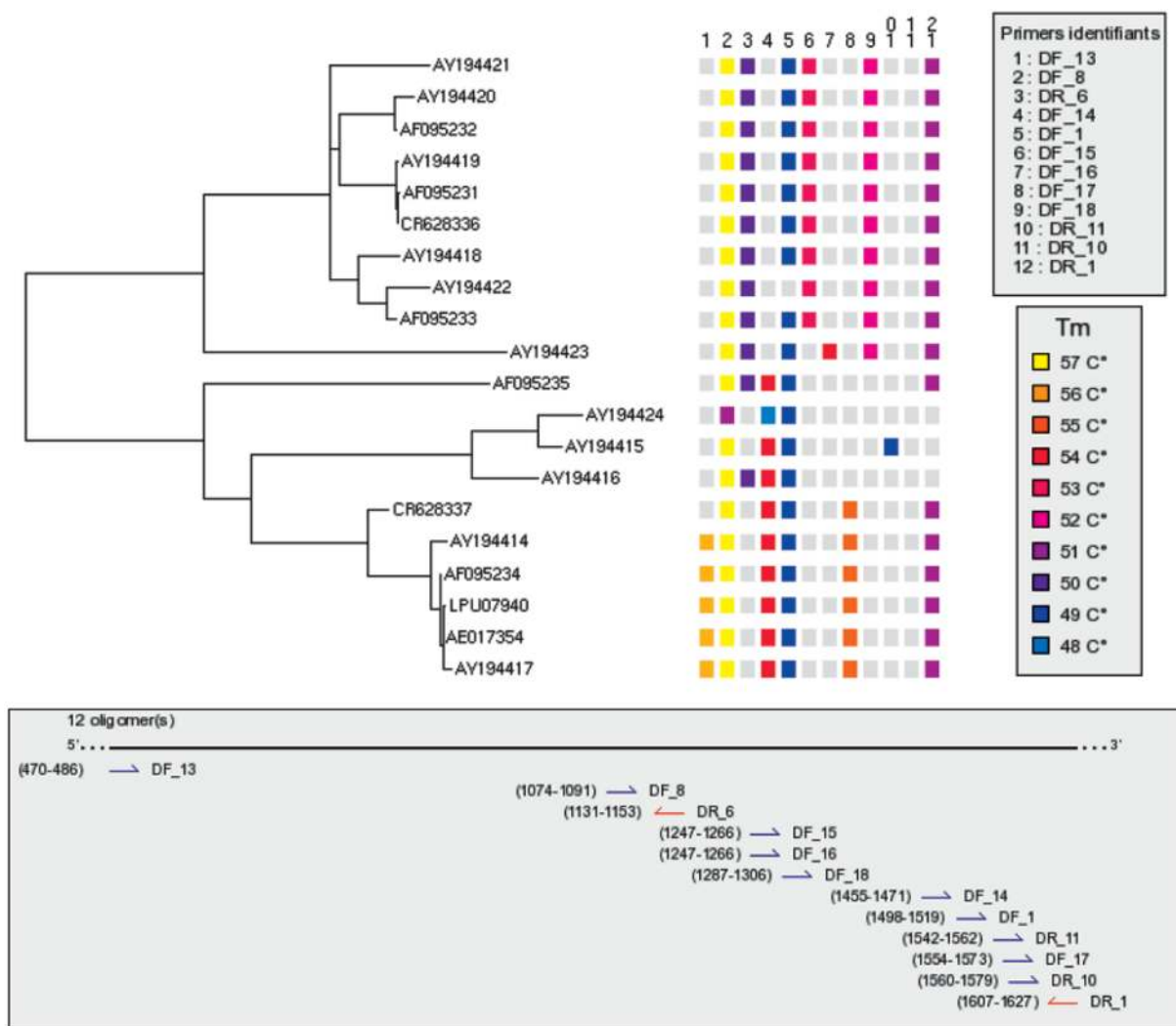
Figure 1. Analysis of primers found in papers for the amplification of the dotA gene in bacteria *Legionella pneumophila* (PubMed PMIDs: 11914343, 11927981, 12962307, 14563861, 7891566, 16000746, 16495539). Top, from left to right, the phylogenetic tree obtained using aligned dotA sequences (identified by accession numbers), the Heat Map (primers in columns, identified by numbers), list of primers, correspondence between colours and predicted T_m. Bottom, position and strand of each primer. Composite figure constructed using TreeDyn and files produced by OHM. Grey squares when the predicted T_m is below 48°C. Note that none of these published primers will well amplify every sequence. For example, we can clearly see that Primer 9 (DF_18) is specific of upper clade. The bottom of this figure shows an image generated with OHM and imported in TreeDyn (primers location with their strands and orientations along a consensus of sequences used in this example).

RESULTS

Results are available under two forms:

- 'View alignments': a display of information related to the position of each oligomer within the data set of aligned sequences (if provided) as well as information for any possible amplicon.
- 'View T_m': a display of information concerning the T_m of each oligomer on each sequence, as well as images showing how each oligomer will hybridize on each sequence of the alignement.

View alignments

This output displays oligomer's locations and strand orientations according to the alignment provided. Primers are shown in the order they were loaded or can be sorted by positions (Figure 1, bottom). Other data are the date of analysis, exact locations of primers along sequences, primers not found in aligned sequences (if any). A series of files can also be downloaded:

(1) 'Download html page' and 'download gif file' are the two files that save results described earlier. Alternatively, the entire web page can be saved.
(2) 'Extract hybridization areas (fasta)' and 'Extract hybridization areas (gif)' will work for a primer selected from a DropDown List ('Selected primer 1'). This will produce a fasta file or an image with the primer sequence, the corresponding domain in each sequence and a number (adjustable) of additional nucleotides to display at both ends. This is very useful to rapidly check that sequences are well aligned at each primer location.

(3) 'Extract PCR product' provides a fasta file corresponding to the PCR product generated by a couple of primers selected using the drop down lists; it contains products for every sequence contained in the alignment. The selection of a couple of primers is facilitated as their positions in sequences are visible on screen. Useful for a rapid analysis of PCR product size and possible differences in length and composition.

View T_m

Results provided are thermodynamic values (enthalpy, entropy and melting temperature) for each oligomer on each sequence (including mismatches). A graphical representation of hybridization appears on screen when the cursor is moved over a sequence. A 'R', if present, means that the oligomer was found as invert-complement. Positions of primers in sequences are also shown. The following files can be downloaded:

(1) Excel or text file: these files provide the predicted T_m of each oligomer on each sequence.
(2) Map of temperatures: an image file that displays, for each oligomer and each sequence, the predicted T_m as color-coded squares (i.e. a Heat Map).
(3) Matches: a file that provides an overview of results, easy to use for designing degenerated primers when necessary. For any given primer, it shows how many degenerated positions are necessary to match every sequence, as well as how many sequences match each of these modified oligomers.
(4) TreeDyn files: one of the goal of OHM is to provide an easy overview of how each primer hybridizes to each sequence of a data set. Representation of these sequences as a phylogenetic tree is useful, since hybridization is often required for a clade but not for related clades (gene families, pathogenic bacteria, etc.). TreeDyn, a powerful tree editor (13) has been modified to take annotation files produced by OHM. These annotations include the computed T_m as color squares (from yellow for high T_m to blue for low T_m) as shown in Figure 1.

CONCLUSIONS

OHM is running since September 2007 and used within the framework of a European project. Most of computing is done by the computer of the client (via Javascript), this server can support many users at the same time. Local install is possible since sources are available for download (see link in help). An online help is provided (http://bioinfo.unice.fr/softwares/ohm/help/help.html), it contains tutorials and examples of use. A demonstration data set is also available from the main OHM page (http://bioinfo.unice.fr/ohm, click on the 'Data test' link to directly import sequences in OHM). This software has already been used and results have been experimentally checked (primers for specific amplification of 16S rRNA gene sequences of some bacterial species). We encourage users to propose new features.

REFERENCES

1. Rozen,S. and Skaletsky,H. (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol. Biol.*, **132**, 365–386.
2. Fredslund,J. and Lange,M. (2007) Primique: automatic design of specific PCR primers for each sequence in a family. *BMC Bioinform.*, **8**, 369.
3. Rychlik,W. (2007) OLIGO 7 Primer analysis software. *Methods Mol. Biol.*, **402**, 35–60.
4. Allawi,H.T. and SantaLucia,J.J. (1998) Nearest-neighbor thermodynamics of internal A.C mismatches in DNA: sequence dependence and pH effects. *Biochemistry*, **37**, 9435–9444.
5. Allawi,H.T. and SantaLucia,J.J. (1998) Thermodynamics of internal C.T mismatches in DNA. *Nucleic Acids Res.*, **26**, 2694–2701.
6. Allawi,H.T. and SantaLucia,J.J. (1998) Nearest neighbor thermodynamic parameters for internal G.A mismatches in DNA. *Biochemistry*, **37**, 2170–2179.
7. Wetmur,J.G. (1991) DNA probes: applications of the principles of nucleic acid hybridization. *Crit. Rev. Biochem. Mol. Biol.*, **26**, 227–259.
8. Le Novère,N. (2001) MELTING, computing the melting temperature of nucleic acid duplex. *Bioinformatics*, **17**, 1226–1227.
9. Kibbe,W.A. (2007) OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res.*, **35**, W43–W46.
10. Panjkovich,A., Norambuena,T. and Melo,F. (2005) dnaMATE: a consensus melting temperature prediction server for short DNA sequences. *Nucleic Acids Res.*, **33**, W570–W572.
11. Panjkovich,A. and Melo,F. (2005) Comparison of different melting temperature calculation methods for short DNA sequences. *Bioinformatics*, **21**, 711–722.
12. Chavali,S., Mahajan,A., Tabassum,R., Maiti,S. and Bharadwaj,D. (2005) Oligonucleotide properties determination and primer designing: a critical examination of predictions. *Bioinformatics*, **21**, 3918–3925.
13. Chevenet,F., Brun,C., Bañuls,A., Jacq,B. and Christen,R. (2006) TreeDyn: towards dynamic graphics and annotations for analyses of trees. *BMC Bioinform.*, **7**, 439.