Designing Primer Pairs and Oligos with OligoFaktorySE


Laurent Gatto[1] and Colas Schretter[2]

[1] DNAVision, Microarray Unit, Gosselies, Belgium
[2] Université Libre de Bruxelles, Belgium

Abstract

OligoFaktorySE (Standalone Edition) is a free software for Apple Mac OS X 10.4 or more recent versions. It designs long oligonucleotides for DNA microarrays, primer pairs for PCR amplifications, siRNAs, and more… The innovative user interface emphasizes usability and is aimed at assisting researchers for a painless, rapid, automated, and reliable design. OligoFaktorySE is currently distributed as freeware on the Math & Science section of Apple Downloads (http://www.apple.com/downloads/macosx/math_science/oligofaktorystandaloneedition.html) and on its dedicated website (http://homepages.ulb.ac.be/~cschrett/oligofaktory).

Introduction

Primer and oligonucleotide design are important applications of bioinformatics in molecular biology. This is reflected by the abundance of oligonucleotide design softwares that are available. Regular primer design tools (Primer3) are targeted towards computer-savvy users and require scripting skills to automate the design of thousands of primer pairs or oligonucleotides. Alternative tools are often controlled by a graphical user interface (GUI) but only allow the design of a small number of oligonucleotides.

OligoFaktorySE is the software presented in this paper. Its main particularity is that the interface runs exclusively on Mac OS X 1.4 or newest versions of the operating system from Apple. The design of the interactive interface follows the latest's Apple Human Interface Guidelines (HIG) to easily design large batches of oligonucleotides while tuning finely the design constraints. Furthermore, the interface features advanced visualization of results.

The OligoFaktory has been first introduced to the EMBnet community in 2005 with a short article describing its main features [1]. An equivalent web service was also announced in a Bioinformatics application note [2]. Today, only the standalone version (SE) for Mac OS X is maintained.

Main Features of OligoFaktorySE

OligoFaktorySE allows the researcher to import DNA sequences in FASTA format. It also uses its own XML format to store the sequences, the oligonucleotides, and the design parameters. Several complementary actions are at hand:

- the **Design Oligos** action allows for the automated design of specific long oligonucleotides for the development of microarrays;
- the **Design Primers** action allows for the automated design of specific PCR primer pairs on an arbitrary number of regions;
- the **Design siRNAs** action allows designing optimal 19bp siRNA with the modern method described in [3];
- the **Detect Repetitions** action identifies all repetitions and microsatellites under user-specified constraints;

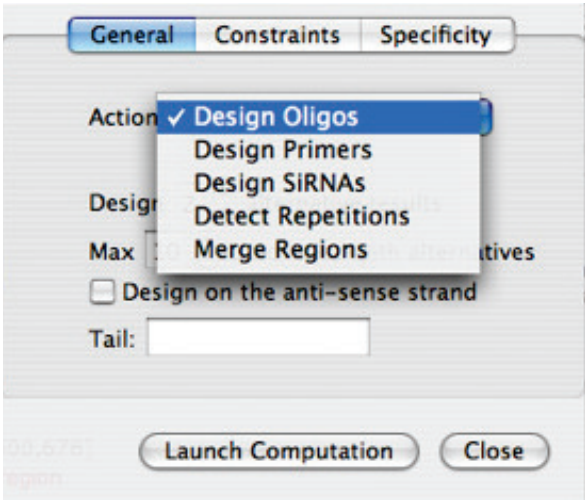

Figure 1. **Actions**. One of the five complementary design actions can be selected at each iteration of the interactive design session.

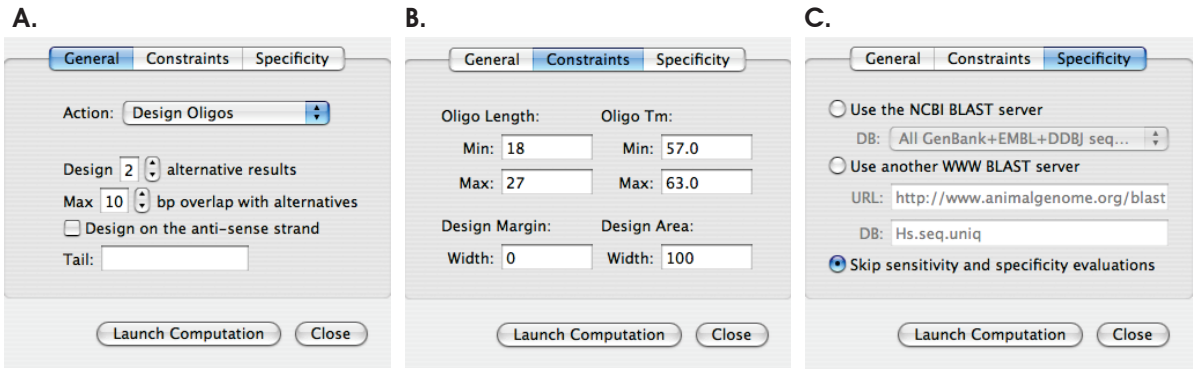

Figure 2. **Parameters.**
A. For each design action, a dedicated panel showing general design constraints appears to the user.
B. Hard constraints such as the allowed ranges for the oligonucleotide length and melting temperatures can also be set.
C. Sensitivity and specificity of oligonucleotides can be optionally checked against a BLAST database.

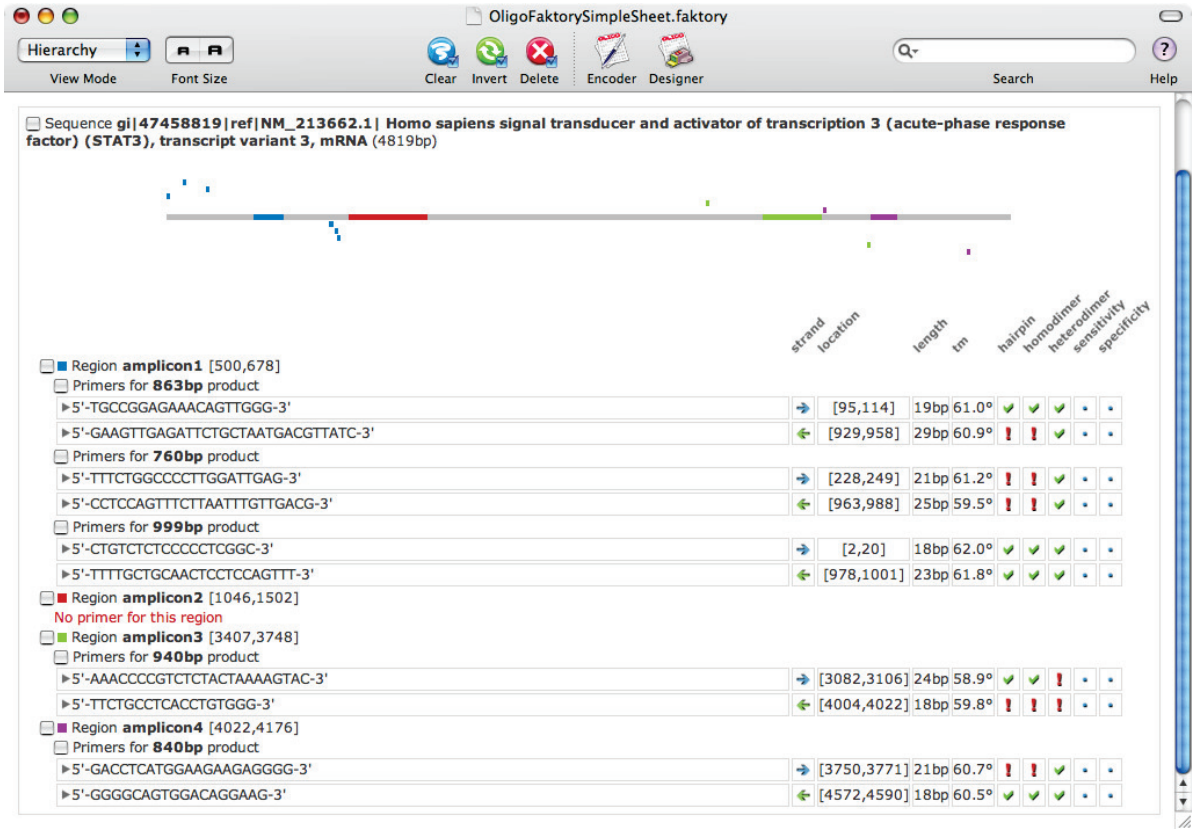

Figure 3. **Visualisation.** A colourful bar graph representation illustrates the relative position and length of primers around amplified regions.

- the **Merge Regions** action merges regions which are too close to allow designing specific primer pairs.

Each of these actions has a set of parameters that can be set by the user: design constrains like oligonucleotide length, preferred product length for PCR, melting temperatures, location of the oligonucleotides or 5'-end tail.

The underlying design algorithm is based on approaches from statistical mechanics [4] and implements successive selection of best candidates on various criteria like minimization of secondary structures (homo- and hetero-dimers).

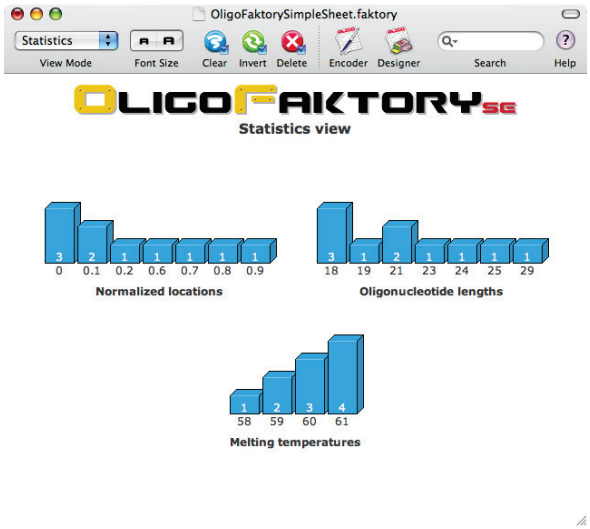

Figure 4. **Statistics.** Statistics are computed to visualize the distributions of relevant features for large batches of results.

Furthermore an optional evaluation of specificity and sensitivity can be performed using online or custom offline BLAST databases.

At the end of the design, the results can be directly visualized. The output includes the list of oligo sequences together with their corresponding locations on the query sequences, their lengths, and their melting temperatures. Easy-to-spot warning flags are shown in case of problems with secondary structures and/or with specificity evaluation. These results can be presented in full details using a hierarchical view, a short listing that emphasizes leaf results, while hiding locations information or a statistics view summarizes the results with charts showing the distributions of main features.

Finally, the results can be exported in several formats including OligoFactorySE's own XML based format (including the oligonucleotides, the design warnings and initial input data), FASTA format (oligonucleotides only) and CSV format for easy importation in spreadsheet softwares such as Excel.

Interactive User Interface

The interface presents consistent application windows, from the data import and design parameter specifications to the visualization of results. However, the novelty of OligoFaktorySE goes beyond the multiple actions presented above and the graphical interface.

The main originality is certainly the interactive user interface. Initial design constrains that are

Figure 5. **Warnings.** Easy to spot flags indicates worst-case thermodynamic properties of oligos, suggesting further inspections.

set by the user are often stringent, at least in a first design round, to maximizes the chances to yield functional results. However, there are always a set of design that do not succeed and that need to be repeated with reduced parameter stringency. The interactive design paradigm allows extracting several design subsets that can then be manipulated independently: the successful oligonucleotides can be stored and the other ones can be recycled for additional designs. At the end, the several design results can be merged together again.

As OligoFaktorySE native format is XML-based, these can be generated and parsed by third party tools. For instance, if hundreds of exons need to be amplified independently, one can easily generate a PCR primer input file based on the gene sequences and exon locations. This input file then just needs to be imported by OligoFaktorySE, the design parameters set and the design can start. Similarly, final results files can be parsed to be incorporated in oligonucleotide databases.

References

[1] Schretter, C. (2005) Discover OligoFaktory Standalone Edition. EMBnet.news, 11(4):24-25.

[2] Schretter, C. and Milinkovitch, M.C. (2006) OligoFaktory: a visual tool for interactive oligonucleotide design. Bioinformatics, 22(1):115-116.

[3] Reynolds, A. et al. (2004) Rational siRNA design for RNA interference. Nature Biotechnology, 22(3):326-30.

[4] Schretter, C. and Milinkovitch, M.C. (2005) Oligonucleotide design by multilevel optimization. Technical Report, ULB.