Interactive primer- and probe-design with CaSSiS

Kai Christian Bader* and Harald Meier
Chair of Computer Architecture, Department of Informatics,
Technische Universität München,
Boltzmannstraße 3, 85748 Garching

Abstract

Designing primers and probes for molecular diagnostic methods depends on the identification of oligonucleotide signatures, short binding sites on genome or marker gene sequences. CaSSiS [Bader et al., 2011] is able to determine such signatures even under relaxed search conditions. It uses the *Bipartite Graph Representation Tree* (BGRT), a newly developed data structure for fast sequence-to-signature mapping.

CaSSiS' command line version was designed for the computation of comprehensive signature collections from large hierarchically clustered sequence datasets, i.e. signature candidates for every defined group. It aims at maintainers of sequence databases who want to provide signature collections along with their public datasets.

For end users who are primarily interested in a few defined target groups (but still want large background data), single queries instead of a comprehensive signature set computation might be sufficient. CaSSiS was therefore extended: Computationally intensive mappings on high-performance systems can now be stored. End users can then use these mappings for querying user-defined groups.

The users are supported by a new graphical user interface. Groups can either be selected within a loaded phylogenetic tree (Figure 1) or be defined as a list of identifiers. By allowing a range of non-target hits, users may influence the specificity. Single requests are usually processed within a second. The result is a list of signatures with maximum coverage (sensitivity) for each entry in the range of allowed non-target matches, and their thermodynamical characteristics (Figure 1).

CaSSiS allows weighted mismatches [Yilmaz et al., 2008] as a measurement for the hamming distance during search. It can also be used to enforce a minimum distance to non-target matches. The number of mismatches is set before creating the mapping.

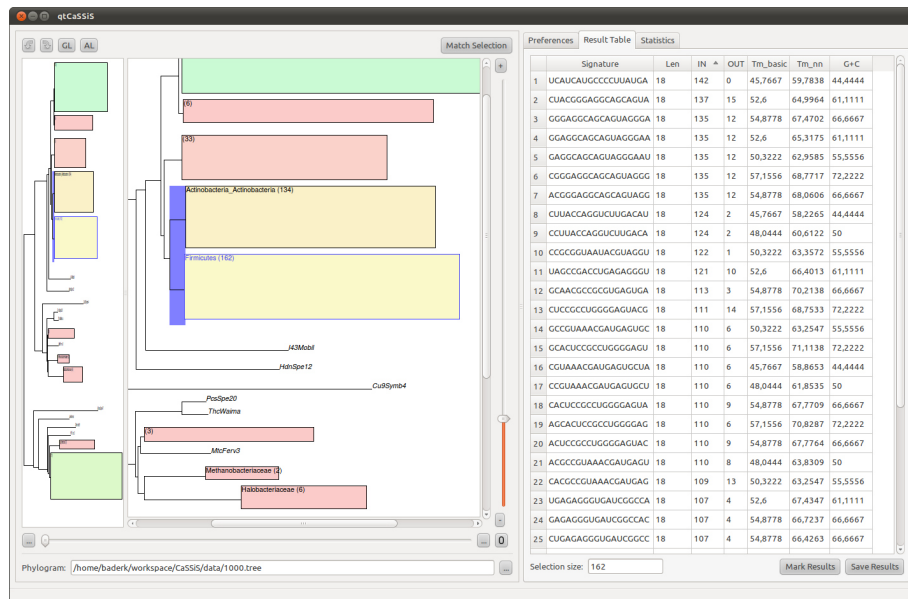*to whom correspondence should be addressed: baderk@in.tum.de

Figure 1: Screenshot: The CaSSiS user interface. Left area: group selection within a phylogenetic tree. Right area: tabs with signature candidates for a selected phylogenetic group. Not visible: Preferences and statistical data.

Availability

The command line and graphical user interface versions of CaSSiS are available for download at: `http://cassis.in.tum.de/`

References

Kai Christian Bader, Christian Grothoff, and Harald Meier. Comprehensive and relaxed search for oligonucleotide signatures in hierarchically clustered sequence datasets. *Bioinformatics*, 27: 1546–1554, Jun 2011. doi: 10.1093/bioinformatics/btr161. URL `http://dx.doi.org/10.1093/bioinformatics/btr161`.

L. Safak Yilmaz, Lindsey I. Bergsven, and Daniel R. Noguera. Systematic evaluation of single mismatch stability predictors for fluorescence in situ hybridization. *Environmental Microbiology*, 10(10):2872–2885, 2008. ISSN 1462-2920. doi: 10.1111/j.1462-2920.2008.01719.x. URL `http://dx.doi.org/10.1111/j.1462-2920.2008.01719.x`.