



ARB Probe Service



L. Richter, R. Westram, Yadhukumar, H. Meier *, A. Bode*, W. Ludwig

Lehrstuhl für Mikrobiologie, Am Hochanger 4, 85350 Freising

* Lehrstuhl für Rechnertechnik und Rechnerorganisation, Boltzmannstr. 3, 85747 Garching
Technische Universität München.

Abstract

ARB Probe Service is a part of the ARB project and provides interactive internet access to a database comprising all potential phylogenetic gene probes generated from the current versions of the ARB rRNA databases. ARB probe service is still under development and provides an interface to the ARB probe functions for users which do not need full functionality of the comprehensive ARB software package. No local installation of the ARB package is needed. The ARB (arbor, latin: tree) project (<http://www.arb-home.de>) was started in 1993 at the Technical University of Munich as an interdisciplinary initiative of the Lehrst. f. Mikrobiologie and the Lehrst. f. Rechnertechnik und Rechnerorganisation.

ARB is based on the concept of integrated databases of raw and/or processed sequence and any other type of additional data assigned to the individual sequence entries. The work bench facilitates curation, alignment, visualisation, and analysis of databases of any genes or genomes. This is provided by a selection of directly cooperating software tools which comprise primary and secondary structure editors, automated or user guided aligners, programs for establishing profiles and filters, different treeing approaches for homogenous or concatenated heterogenous gene sets, taxon or gene specific probe/primer design and evaluation as well as other facilities. Integrated ARB databases are continuously maintained for full genomes as well as for evolutionary conserved genes (e.g. rRNAs, elongation-initiation factors, RNA polymerases, recA, heat shock proteins, DNA gyrases/topoisomerases, aminoacyl tRNA synthetases, ATP synthetases). The ARB software as well as the ARB databases are available to the public.

Among the most popular applications of ARB are design and evaluation of taxon-specific gene probes. So far, local installation of the ARB software and the ARB databases was needed to get interactive access to the potential of the data set as a source for taxon specific probe targets. There are other facilities to find published and evaluated specific probe sequences or evaluate user provided probe targets. However, databases and services providing a complete set of any potential probe targets against the background of all currently available sequences of the respective target gene(s) are still missing. The new ARB Probe Service now provides a software client to interactively retrieve the full spectrum of potential probes and additional information for any phylogenetic/taxonomic level. The ARB Probe Service is realized as Client/Server architecture where the client constitutes a platform independent component which can be used without the need for local installation of ARB or other software and a local database management. The client uses a phylogenetic tree display for database interaction and navigation as well as for the visualization of the results. Selecting single nodes or groups of nodes results in a listing of the specific probes available for the selected nodes. Detailed information for the selected probes or for tree nodes is displayed in additional text boxes.

The server itself is maintained by the ARB group. It provides the evaluated results from a complete analysis of the latest version of the respective target gene ARB database comprising all potential gene probes for any phylogenetic level. Software and databases are available from the ARB project homepage (www.arb-home.de). The ARB_genome project was supported by the bmb+f as part of the regx program for functional genomics (<http://www.regx.de>)

How Does It Work ?

Technical Realization

Existing curated rRNA databases are analyzed for potential phylogenetic gene probes by special parts of the ARB program package using expert-derived parameters. The results of this analysis are stored in separate database files. The platform independent client uses http-protocol and a common webserver as communication infrastructure. The tree description is stored locally until automatic update is triggered through the server. The client issues requests about probe information affiliated with selected tree nodes which are then answered on the base of the prefurnished analysis files using cgi scripts. Graphical rendering and representation of any kind of information is done by the client.

User customization as well as presentation of more additional probe information will be included in the future versions of the ARB Probe Service.

If subtree-specific probes exist the number is written in blue color. These probes detect 100% of the subtree and no species outside of the subtree.

If there is no 100% specific probe candidates for the subtree, then two alternatives are given:

A probe candidate which detects the complete subtree with the minimal number of hits outside of the subtree.
A probe candidate which detects the maximum number of species within the subtree but none outside of the subtree.

For these nodes the best values for both the types of candidate probes are given which states the number of non-group hits for the respective candidate and the percentage of subtree coverage for the other candidate, respectively.

Subtrees outside of the displayed section are compacted into boxes labelled with the number of represented species in the subtree.

Detailed hit list for probe: CCGGCUUAACCCUC
Melting temperature (24 rule): 46°C
%GC-Content: 53%

Mycoacterium lepre	X53999
Mycoacterium tuberculosis	S45301 X58090
Mycoacterium chirothermophilus	X79094
Actinoplanes utahensis	X60823
Frankia species	M58466
Bifidobacterium bifidum	S83624
Clavibacter xyli	M60935

This is a screen shot for a selected branch without 100% specific probes.
The probe with the maximum subtree coverage detects 53% of the species of the subtree and the matching species are scattered over the whole tree.

Tree Window

Phylogenetic tree display for graphical visualization and navigation:

A section of the complete tree is displayed with the information about the occurrence of potential probes for every subtree. The displayed section is limited by the number of branch levels to be displayed and dynamically changed upon navigation within the tree.

Probe Window

All the possible candidate probes are listed in the Probe Window

The user selects a candidate sequence (Probe) and the respective matching species are highlighted automatically in the Tree Window.

Detail Window

Detail Window allows to view additional information for the each selected candidate probe listed in the Probe Window.

Melting temperature, %GC content, list of matching species with the respective accession numbers as reference for further investigation are displayed as the additional information for the selected probe.

This is a screen shot for a selected branch with 100% specific probes.

This prototype is available upon request. Please contact Dr. Lothar Richter (richter@arb-home.de).

References: ARB - A software environment for sequence data (submitted to Nucleic Acids Research).
Wolfgang Ludwig, Oliver Strunk, Ralf Westram, Lothar Richter, Harald Meier, Yadhukumar, Arno Buchner, Tina Lai, Susanne Steppi, Gangolf Jobb, Wolfram Förster, Igor Brettske, Stefan Gerber, Anton W. Ginhart, Oliver Gross, Silke Grumann, Stefan Herrmann, Ralf Jost, Andreas König, Thomas Liss, Ralph Lüssmann, Michael May, Björn Nonhoff, Boris Reichel, Norbert Ströhlein, Alexander P. Stamatakis, Norbert Stückmann, Alexander Vilgaj, Michael Lenke, Thomas Ludwig, Arndt Bode and Karl-Heinz Schliefer.