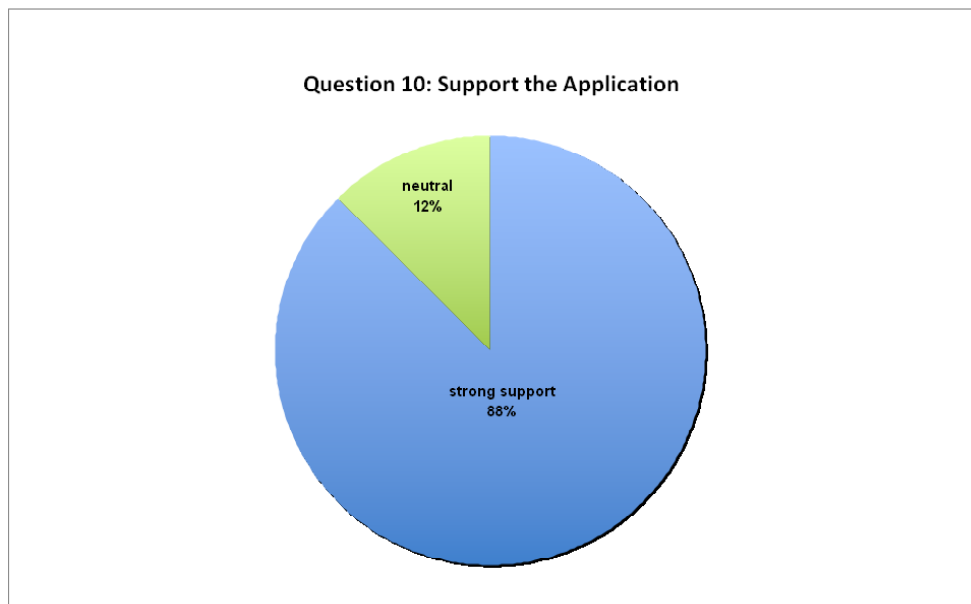
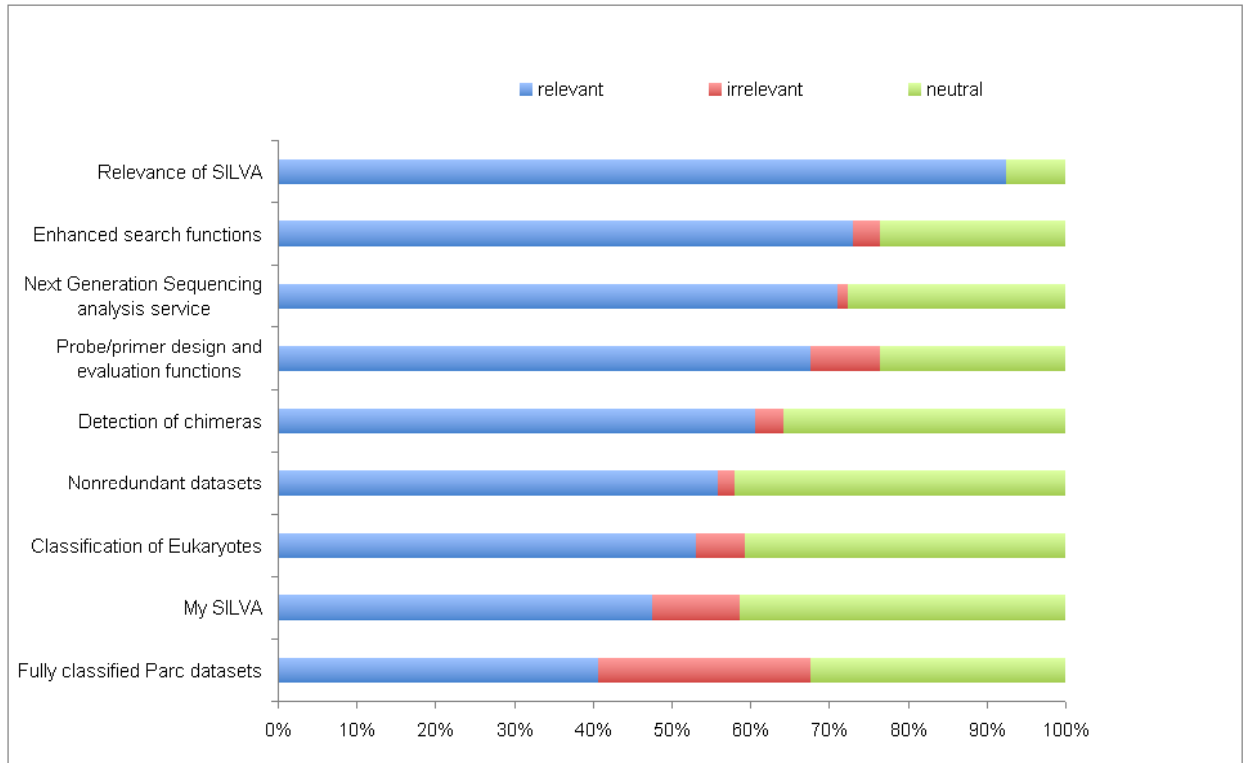


## SILVA Survey: Building the next generation databases for ribosomal RNAs

Start of survey: 25.03.2010, end of survey: 09.04.2010

Total responses: 145

Number of different countries: 26



## Acknowledgements

- The SILVA database project is the only group that is providing a frequently updated source of high quality DNA reference sequences. It has been an invaluable resource, especially for 18S rRNA research groups, since the RDP has not updated their 18S database for many years. SILVA is the new gold standard for rRNA analyses.
- I have found SILVA a fantastic service for myself and the students that I work with. It has made ARB accessible to a wide variety of researchers and has accelerated our phylogenetic analyses. Thanks guys!
- I appreciate the smaller NR data set that can be loaded into ARB. This has allowed me to actually work with the data set on my computer (that has limited memory). Highlighting key species in red or blue has been incredibly helpful for building trees. This is fantastic. It would be really useful to be able to create a small dataset that contains just the type species. This would be particularly helpful for making smaller phylogenetic trees. I have found the sequence alignments from Silva to be the most representative and contain the fewest errors and has become our institute's preferred alignment. Overall Silva is an incredible resource that is crucial to my research and my understanding of the microbial ecology and systematics. Thank you.
  - Please have a look at the "All Species Living Tree Project" <http://www.arb-silva.de/projects/living-tree/> which contains only typespecies.
- Congratulation to the new SILVA beta version to all authors, features like downloading sequences from a specific publication, and filtration based on the taxonomy are very helpful. Also the SSUref 102 NR is helpful because it works much better on slower computers. ARB and SILVA are very helpful for my study and I totally support further developments/implementation for 454 data (where I currently have to use other resources for).
- This a great service to the community that you provide and you should be commended for it!
- SILVA war eine der nützlichsten Erfindungen für Umweltforscher die Molekularbiologie benutzen. So einfach war Bäume basteln noch nie! Sie müssen unbedingt weiter machen! Was bei einer Projektpause passiert ist gut an RDP zu erkennen (wo man sich redlich bemüht, wirklich!) Liebe DFG: Besser kann man sein Geld kaum anlegen!
- I would like to state that arb project, and arb-silva.de web site was a major boost on my daily work. Thank you for you great work!
- Silva is a highly valuable resource for my research! Thus, I strongly support the Silva developers and the continuous maintenance/improvement of the Silva project.
- Of the 3 major online repositories of 16S rRNA multi-species alignments, Silva is my first choice; I only use the others for special tasks not easily accomplished via Silva or my local Arb version of the Silva db. The primary reason: the high quality Arb databases available for download. The metadata available is invaluable, above and beyond the aligned sequences themselves. Secondary reasons: high alignment quality, with access to quality scores for alignment, chimeras, the LTP dataset, and now the nonredundant dataset. The comprehensiveness of the data is nice, but not as important to me as the other features most of the time, I (and many others) need to be going in the opposite direction dereplication! The release of the NR set was a very welcome surprise.

- You can see I fully support your efforts to get as much financial support (and do as much additional work!) as possible. Life is so much easier with well-maintained public databases.
- It is a very useful tool for the whole molecular ecology community. I hope your project is funded, as it is important to continue to evolve with the new sequencing methodologies (e.g. pyrosequencing).
- Keep up the good work! It is unbelievably useful to have so much information available in the ARB framework. Much more useful than the interfaces to other sequence collections (e.g. genbank, kegg, pfam, etc...).
- SILVA and ARB have been the backbone of my research for the last 6 years and I would not have been able to accomplish my work without it. It is an incredible resource for all microbial ecologists and I sincerely hope that the developing/working group gets all the funding it needs to continue the improvements!

### **Eukaryotes**

- A Dataset for fungi that covers 18S and ITS might be very helpful - all other eukaryotic sequences are of less interest as they are not so difficult to integrate in the datasets in ARB. What will come with 454 sequencing - I have no idea if SILVA and ARB might handle that amount of data or if some additional support is necessary - if yes, than it is worthwhile to get support from SILVA. Personalized services are not necessary in SILVA if one knows how to handle ARB. For those how don't want to get into ARB, you might provide a service. But I think it's not a primary task of SILVA. The good thing you provide is a reliable dataset of sequences to use in ARB and I think this is the main task to follow. Keep this dataset up to date. Thanks for doing so.
- The SSU alignment of eukaryotes is really poor and should be improved.
- Facilitate alignment of eukaryotes through adapted secondary structure templates.

### **Enhanced Search Functions**

- The online Search feature is hard to work with. For example, I would like ALL values that I fill in to be present in my search (similar to Google). Instead if I put in 'rat' it now seems to return sequences that have either one of these terms, not both of them. Then, if I go to list and return to search, my first search term is gone.
  - I guess its ANY what you are looking for, please try the new beta search functions at: <http://beta.arb-silva.de/search/> and in case it does not work, tell us.
- Of all of my responses listed in the survey, the one that I consider the most important to my research would be to enhance the search functions. I am an experienced ARB user, but seem to have constant difficulty searching SILVA for sequence sets that I know exist. Thank you - great work, Frank et al.
- Maybe I missed it, but I can't use multiple queries in the search fields. Such option would be interesting (select several acc number at once).
  - Searching for several accession numbers or even a range of accession numbers is possible in the new beta-search system please check <http://www.arb-silva.de/documentation/background/search-tutorial-beta/>

## More Online Analysis Tools

- As user, I would like to have my sequence checked before submission to embl by a web-based quality check, e.g. a pintail service. If all users do in the process of assembling the full sequence, the overall quality will improve.
- What about integrating the mothur tool for OTU, rarefaction curves, etc. calculation?
- I normally use the Greengenes webpage to align, detect chimeras etc. I find very useful the classification tool they have with the 5 different taxonomies - is it possible to implement such a 16S -level classification tool in SILVA? I have my new 454 data with euk clusters so I hope a web implementation allowing the alignment against the SILVA database, chimera detection and a MASK column-type of analyses in SILVA becomes available. Keep on the good work!
- I would suggest integrating a function for comparing different sequence libraries (like the LibCompare function at RDP, but for more than 2 datasets).
- It would be nice if there were a way to have a built in classifier for silva.
- The datasets are getting so large now that it is straining the computer resources available to individuals on limited budgets. Allowing on-line analyses would be of great benefit to all users.
- It is a great database however there is a need for more web tools that can be used for analysis of our own sequences.

## Next Generation Sequencing Support

- Next gen sequence analysis is exploding, however because of short read lengths we still must rely on classical library construction to get full sequences for probes, qPCR assays etc. Reliable alignment and mapping of 454 reads (400-500bp) to the ref database is important.
- The recent pyrosequencing revolution in microbial ecology is remarkable because there seems to have been little discussion about the use of short reads... One of the main objections that people had against partial sequence based methods like DGGE. Nevertheless 454 is here to stay! Although there have been very interesting software developments (e.g. FastUnifrac, Mothur, the RDP pyro-pipeline) increased compatibility of the ARB-SILVA platform with pyrosequencing data is high on my wish list. Best wishes; good luck with grant application.
- Silva is an incredible resource for microbiology and I strongly support your effort. I think support for next-gen sequencing results is the most important new challenge since most users will likely use these platforms in the near future.

## Functional Genes

- First of all, thanks for your hard work. If the arb-silva should support the functional gene alignment and classifier, that will be very helpful for us. Since the diversity information is not enough to explain the scientific questions, more researches will focus on functional genes, like nitrogen, carbon, sulfur and phosphorus cycles. Best wishes.
- Dear SILVA team, I think I read it once in the ARB mailing list (if this is not too far from your scope): it would be very helpful, if there are also databases for functional genes, e.g., coxII or nirK. For sure, all the community should help in constructing this (i.e., submit their new sequences and maybe you take care of non-redundancy). All the best!

## **SINA**

- It would be nice if the results from a SINA run could be obtained in multiple formats (.arb / .fasta / metadata) without the need to re-run the whole alignment process. More fundamentally, publishing SINA as an open source tool and a detailed documentation of the algorithms it uses would strongly add to the transparency of sequence processing using SILVA (for the moment, it seems like a secret 'black box') and to its impact by enabling the implementation of local workflows incorporating SILVA data and tools.
- It would be wonderful to be able to somehow have a local copy, either through ARB or by itself, of the SINA aligner.
- Increase sequence upload allowance heavily.

## **Training**

- It would be great to have the possibility to meet a person who is well versed in SILVA/ ARB to discuss special problems. E.g. I worked alone with ARB for some time and didn't know anybody even in my city to talk about. Another option would be to make the training course cheaper. They were too expensive and when I was a diploma student I had to pay them on my own. Best regards!

## **Misc.**

- The nonredundant database is a great idea but because of the human metagenome project it is also very skewed. A nonredundant database of non-human and animals associated microbes would be even better.
- Documentation of SILVA should be maintained/improved concentration on features, which are not available in GenBank anyway.