

# Introduction to Chipster

20.11.2020

Eija Korpelainen



*CSC – Suomalainen tutkimuksen, koulutuksen, kulttuurin ja julkishallinnon ICT-osaamiskeskus*

# Chipster

- User-friendly analysis software for high-throughput data
- Provides an easy access to over 450 analysis tools
- Command line tools
- R/Bioconductor packages
- Free, open source software
  
- What can I do with Chipster?
  - analyze high-throughput data
  - visualize data efficiently
  - share analysis sessions

# Chipster website (<https://chipster.csc.fi/>)



## Chipster

Open source platform for data analysis



- Home
- Getting access
- Analysis tool content
- Screenshots
- Manual
- Tutorial videos
- Course material
- Cite
- FAQ
- Contact
- For developers:
  - Open source project
  - Tool editor

### Welcome to Chipster

Chipster is a user-friendly analysis software for high-throughput data such as RNA-seq and single cell RNA-seq. It contains over 450 analysis tools and a large collection of reference genomes. Up to version 3.x Chipster is a desktop application requiring Java Web Start. Chipster version 4 is a Web application which runs on your browser. Both run the actual analysis in the cloud and the analysis tool content is the same, but **we encourage users to move to the Web App. The Java desktop application will be phased out in November 2020** because many universities don't provide Java Web Start any more due to the change in Oracle's Java licensing policy.

If you would like to use Chipster running on CSC's server, you need a [user account](#). Please note that both the [v3 server](#) and the [v4 server](#) are also available for local installations free of charge.



#### Launch Chipster v4

*Chipster web application, no Java needed (watch introductory video)*



#### Launch Chipster v3.16.3

*...or launch with more memory: 3 GB or 6 GB*

*Java desktop application. If you have trouble launching it, read the [FAQ](#)*

### Training:

- 20.11.2020 [EGI webinar Using and setting up Chipster](#)
- 3.11.2020 [Intro webinar to MOOC Single-cell RNA-seq data analysis using Chipster](#)
- 27.10.2020 [MOOC Single-cell RNA-seq data analysis using Chipster](#)
- 27.-28.10.2020 [Single-cell RNA-seq data analysis using Chipster, University of Luxembourg](#)
- 21.4.2020 [Microbial community analysis of 16S data](#), virtual course
- 05.3.2020 [RNA-seq data analysis](#), CSC
- 10.1.2020 Community analysis of amplicon sequencing data, Ruokavirasto

### News and resources:

- 14.10.2020 [Lecture videos of single cell RNA-seq data analysis with Chipster course](#)

# Chipster user interface

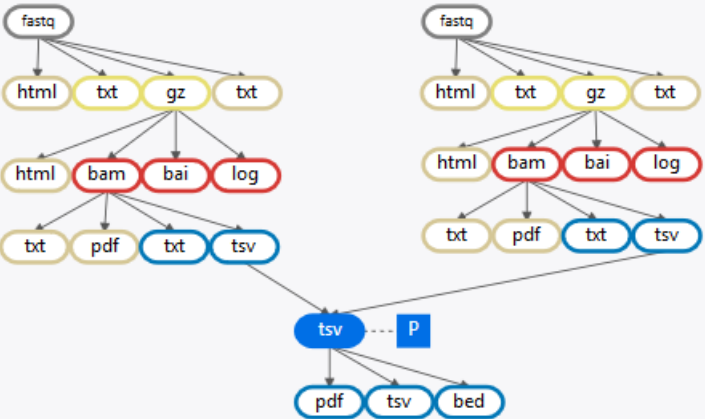


## Files

Workflow List

Find file

Add file



## Tools

NGS Microarray Misc

Find tool

Jobs 0

Category

- Quality control
- Preprocessing
- Utilities
- Matching sets of genomic regions
- Alignment
- Variants
- RNA-seq
- Small RNA-seq
- Single cell RNA-seq
- ChIP- and DNase-seq
- 16S rRNA sequencing
- CNA-seq

Tool

- Read quality with FastQC
- Read quality with MultiQC for many FASTQ files
- Read quality statistics with FASTX
- Read quality statistics with PRINSEQ
- RNA-seq quality metrics with RseQC
- RNA-seq strandedness inference and inner distance estimation using RseQC
- Collect multiple metrics from BAM
- PCA and heatmap of samples with DESeq2
- Check FASTQ file for errors
- Combine reports using MultiQC

Parameters

Run

The tool runs FastQC on multiple FASTQ files, and then combines the reports using MultiQC. Input file is a single Tar package containing all the FASTQ files, which can be gzipped. This tool is based on the FastQC and MultiQC packages. [More info...](#)

## File

ngs-data-table.tsv

Spreadsheet Text Expression Profile Scatter Plot Phenodata Details

First 101 rows of 58396 [View in full screen to see all the rows.](#)

Full Screen

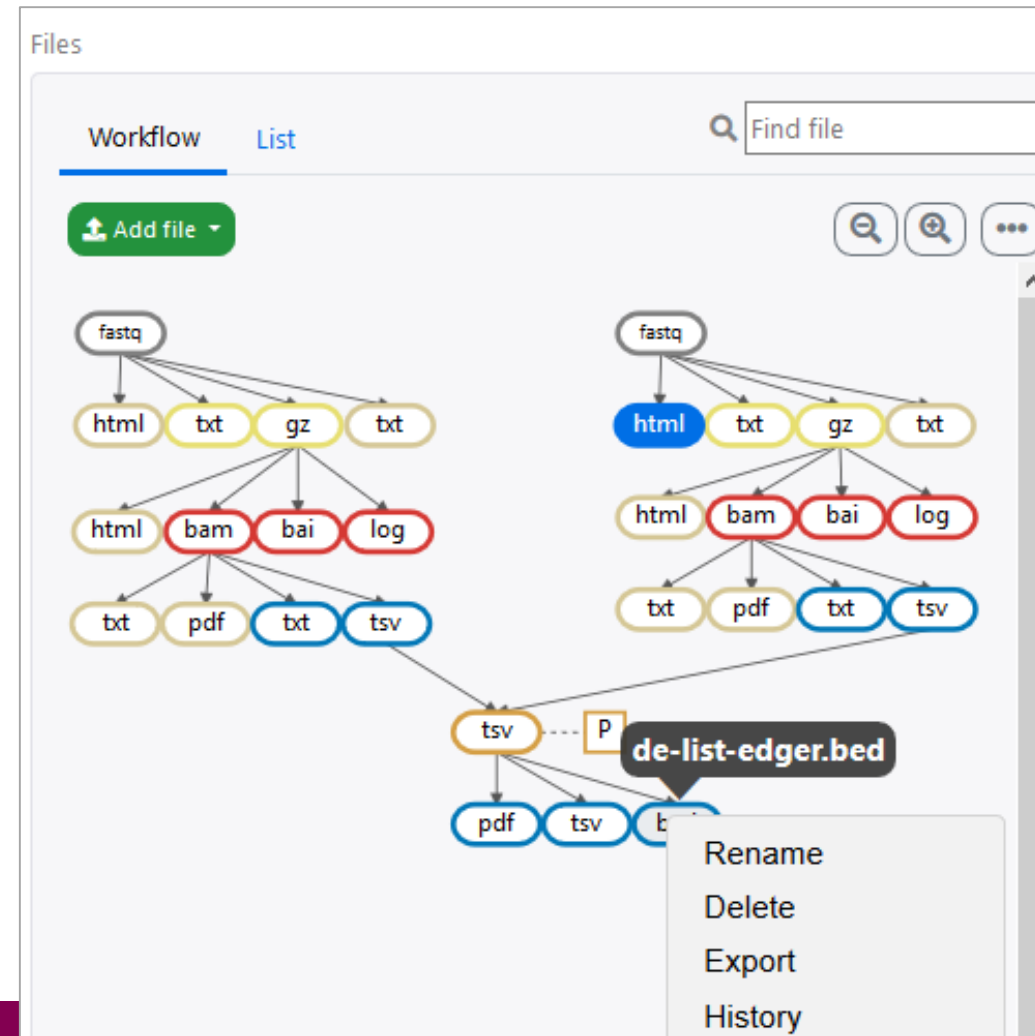
identifier	chr	start	end	length	sequence	chip.sample001.tsv	chip.sample002.tsv
ENSG000000000003	X	100627108	100639991	12883	NA	0	0
ENSG000000000005	X	100584801	100599885	15084	NA	0	0
ENSG000000000419	20	50934866	50958555	23689	NA	0	0

# Analysis sessions

- Your analysis is saved automatically in the cloud
  - Session includes all the files, their relationships and metadata (what tool and parameters were used to produce each file).
  - Session is a single .zip file.
  - Note that cloud sessions are not stored forever! Remember to download the session when ready.
- You can share sessions with other Chipster users
  - You can give either read-only or read-write access
- If your analysis job takes a long time, you don't need to keep Chipster open:
  - Wait that the data transfer to the server has completed (job status = running)
  - Close Chipster
  - Open Chipster later and the results will be there

# Workflow view

- Shows the relationships of the files
- You can move the boxes around, and zoom in and out.
- Several files can be selected by
  - keeping the Ctrl key down
  - Drawing a box around them
- Right click allows you to
  - Download a file ("Export")
  - Delete a file



# Options for importing data to Chipster



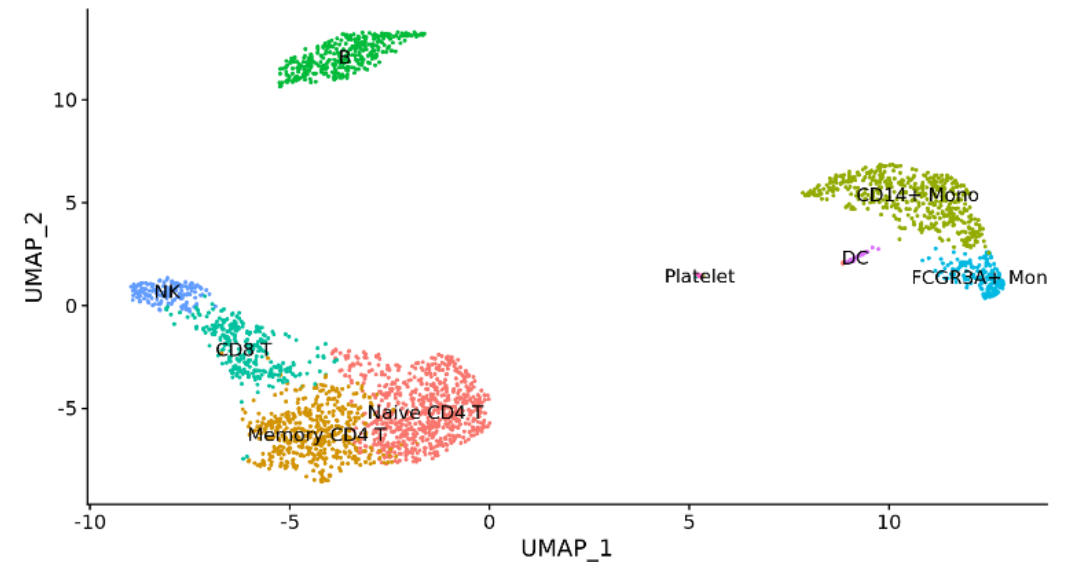
- Add file button
  - Upload files
  - Upload folder
  - Download from URL
- Sessions tab
  - Import session file
- Tools
  - Import from Illumina BaseSpace
    - Utilities / Retrieve data from Illumina BaseSpace
    - Access token needed
  - Import from SRA database
    - Utilities / Retrieve FASTQ or BAM files from SRA
  - Import from Ensembl database
    - Utilities / Retrieve data for a given organism in Ensembl
  - Import from URL
    - Utilities / Download file from URL directly to server

# Example analysis: single cell cRNA-seq

## - clustering cells and finding cluster marker genes



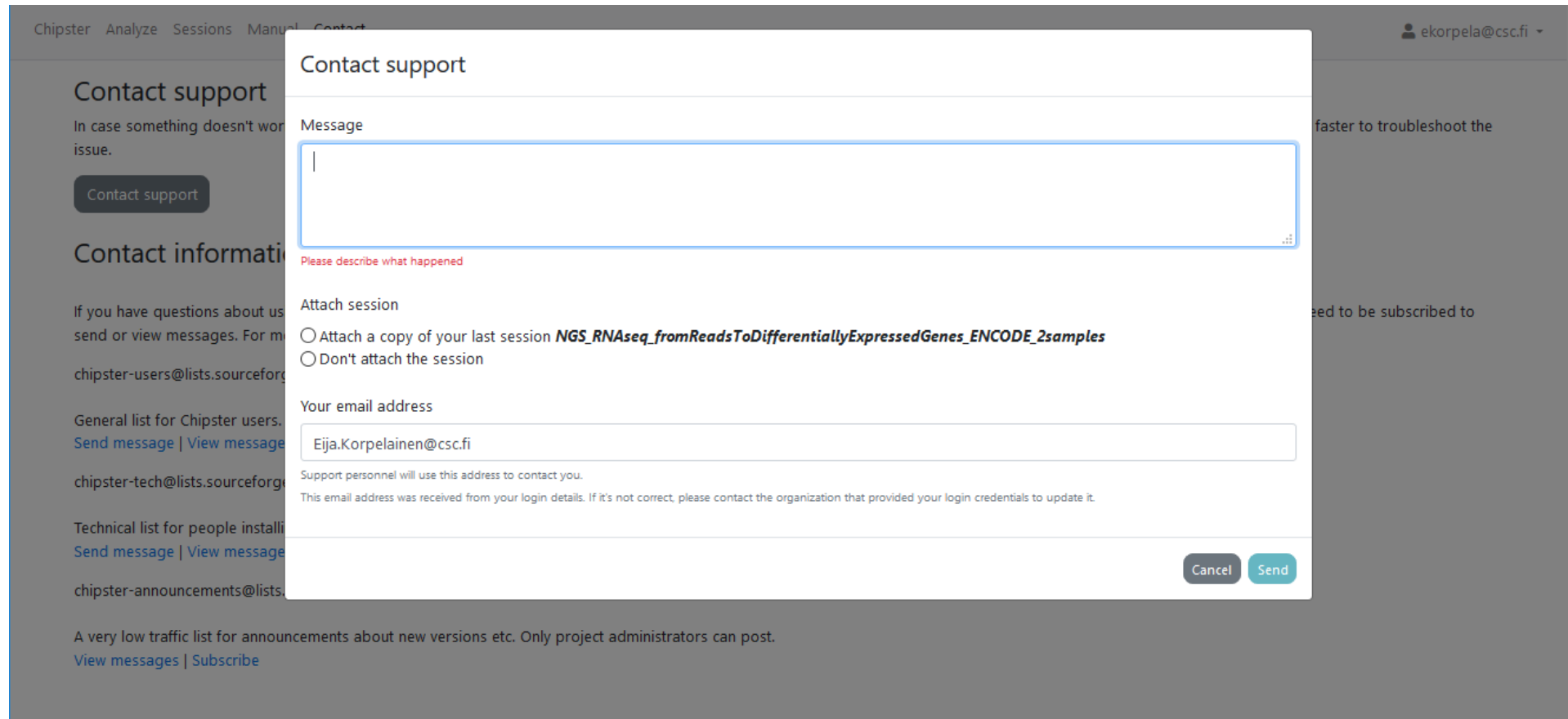
1. Check the quality of cells, filter genes
2. Filter out low quality cells
3. Normalize expression values
4. Identify highly variable genes
5. Scale data, regress out unwanted variation
6. Reduce dimensions using principal component analysis (PCA) on the variable genes
7. Determine significant principal components (PCs)
8. Use the PCs to cluster cells with graph based clustering
9. Visualize clusters with non-linear dimensional reduction (tSNE or UMAP) using the PCs
10. Detect and visualize marker genes for the clusters





# Problems? Send us a support request

-request includes the error message and link to analysis session (optional)



Chipster Analyze Sessions Manual Contact

ekorpela@csc.fi

## Contact support

In case something doesn't work, please contact us. We'll help you faster to troubleshoot the issue.

Contact support

### Contact information

If you have questions about using Chipster, you can send or view messages. For more information, see the following lists:

- chipster-users@lists.sourceforge.net  
General list for Chipster users.  
[Send message](#) | [View message](#)
- chipster-tech@lists.sourceforge.net  
Technical list for people installing Chipster.  
[Send message](#) | [View message](#)
- chipster-announcements@lists.sourceforge.net  
A very low traffic list for announcements about new versions etc. Only project administrators can post.  
[View messages](#) | [Subscribe](#)

### Contact support

Message

Please describe what happened

Attach session

Attach a copy of your last session *NGS\_RNAseq\_fromReadsToDifferentiallyExpressedGenes\_ENCODE\_2samples*

Don't attach the session

Your email address

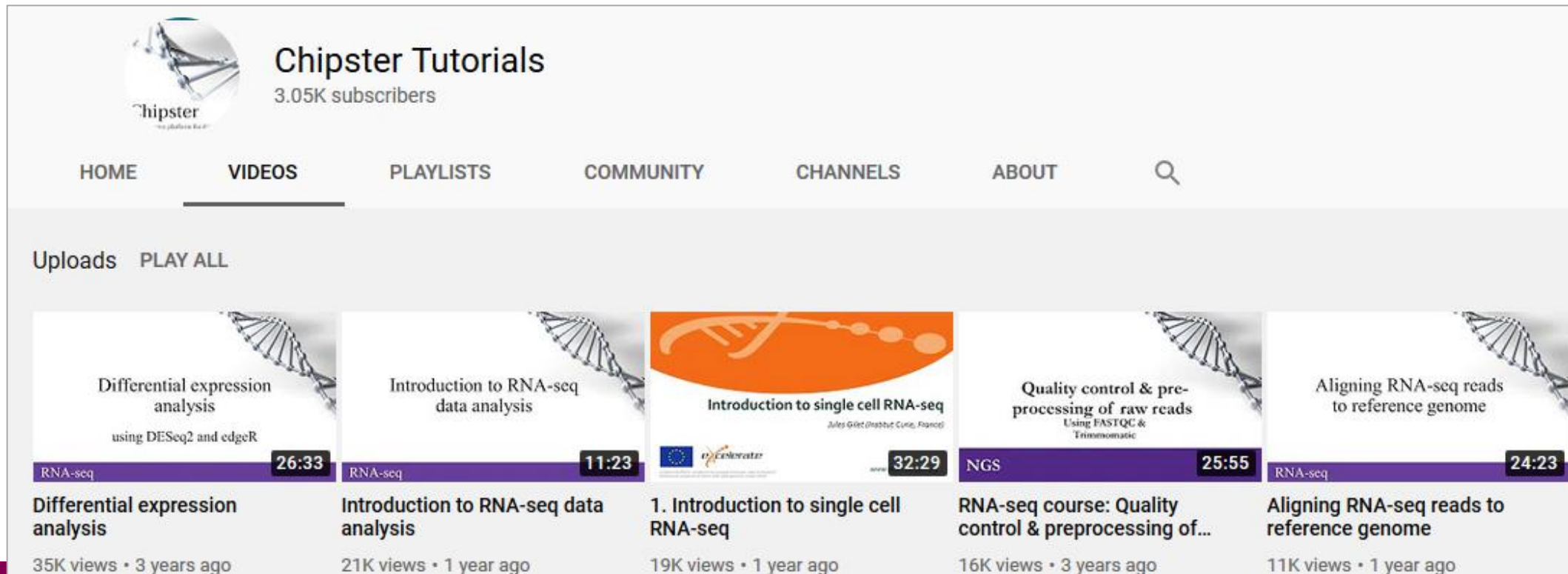
Eija.Korpelainen@csc.fi

Support personnel will use this address to contact you.  
This email address was received from your login details. If it's not correct, please contact the organization that provided your login credentials to update it.

Cancel Send

# More info

- [chipster@csc.fi](mailto:chipster@csc.fi)
- <http://chipster.csc.fi>
- Chipster tutorials in YouTube
- <https://chipster.csc.fi/manual/courses.html>



The screenshot shows the YouTube channel page for "Chipster Tutorials", which has 3.05K subscribers. The navigation menu includes HOME, VIDEOS, PLAYLISTS, COMMUNITY, CHANNELS, and ABOUT. The "Uploads" section is active, displaying a grid of five video thumbnails. Each thumbnail includes the video title, a duration badge, and view statistics.

Video Title	Duration	Views	Time Ago
Differential expression analysis using DESeq2 and edgeR	26:33	35K views	3 years ago
Introduction to RNA-seq data analysis	11:23	21K views	1 year ago
1. Introduction to single cell RNA-seq	32:29	19K views	1 year ago
RNA-seq course: Quality control & preprocessing of...	25:55	16K views	3 years ago
Aligning RNA-seq reads to reference genome	24:23	11K views	1 year ago

# Acknowledgements to Chipster users and contributors



Users' feedback and ideas have helped us to shape the software over the years.  
Let us know what needs to be improved!

