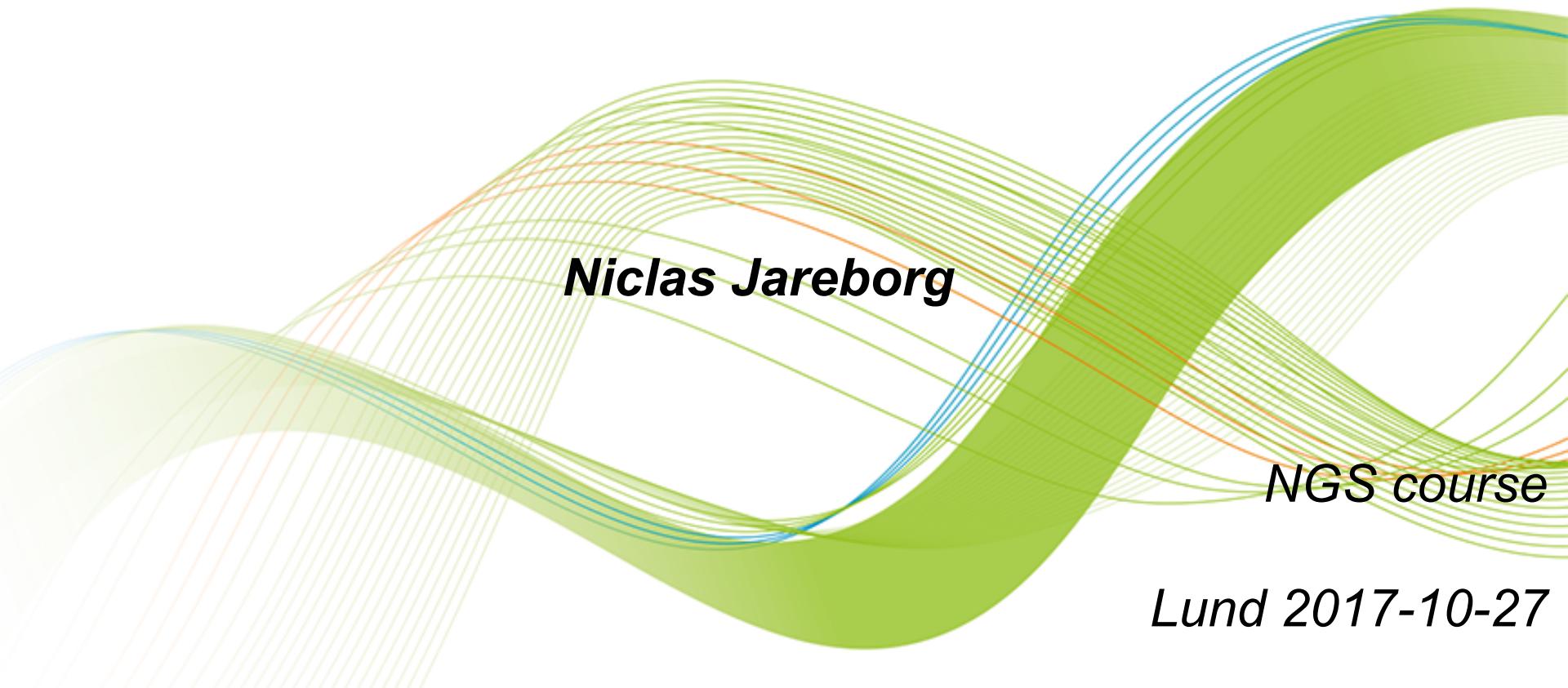


# **SciLifeLab Bioinformatics Platform**

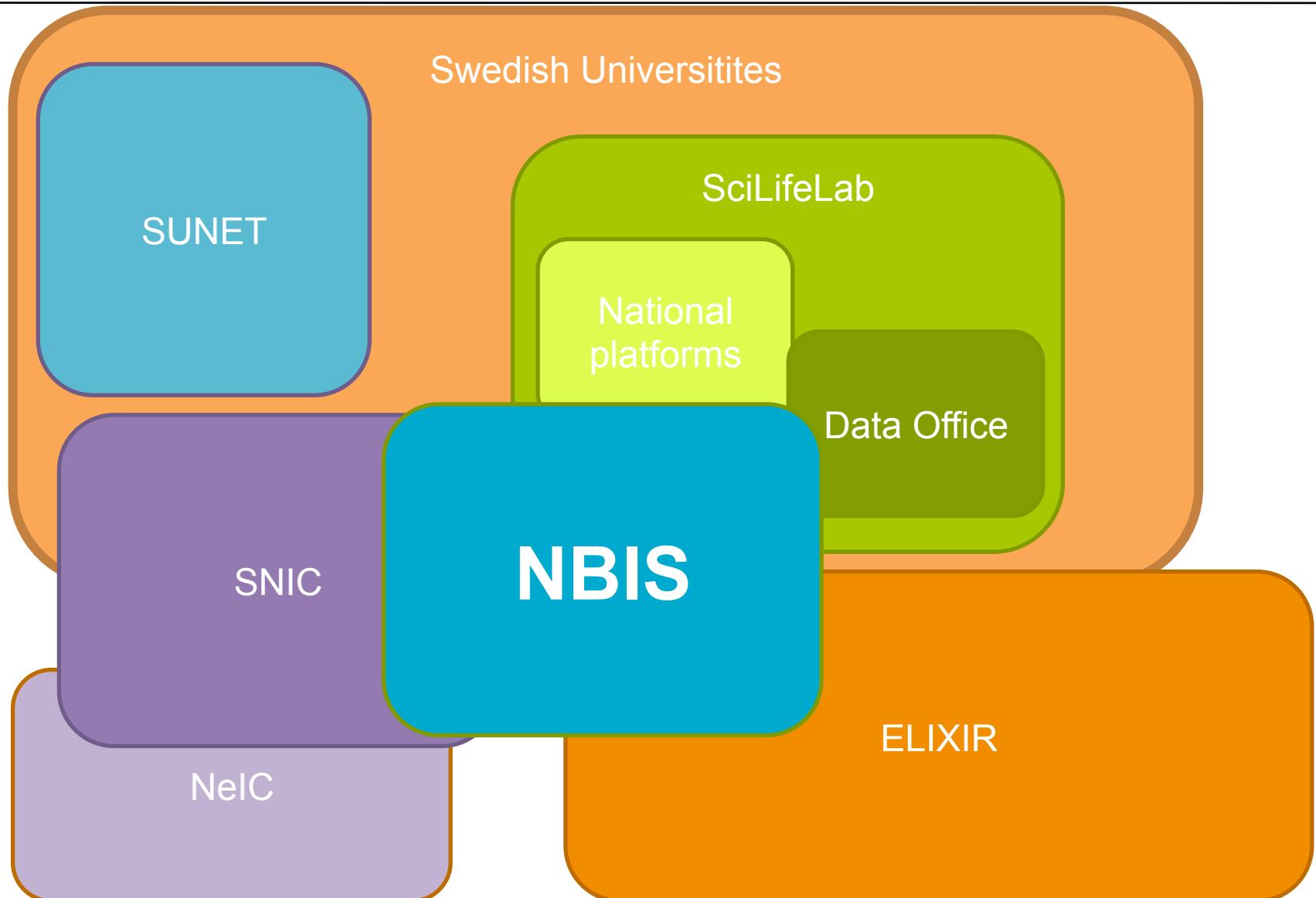
National Bioinformatics Infrastructure Sweden (NBIS)



**Niclas Jareborg**

**NGS course**

*Lund 2017-10-27*



## SciLifeLab

### National service

The Swiss army knife for Swedish Life Science researchers

### Local scientific center



Director: Olli Kallioniemi

Co-director: Lena Claesson-Welsh

#### Vision:

*To be an internationally leading center that develops, uses and provides access to advanced technologies for molecular biosciences with focus on health and environment.*

2010: Strategic research initiative

2013: National resource

2015: New management/chairman

## SciLifeLab platforms

### SciLifeLab national service

National  
Genomics  
Infrastructure

Next  
Generation  
Diagnostics

Single-cell  
omics

National  
Bioinformatics  
Infrastructure  
Sweden

Bengt Persson



Merge of BILS, WABI and  
more; complete 2016.  
National, distributed

VR

SNIC



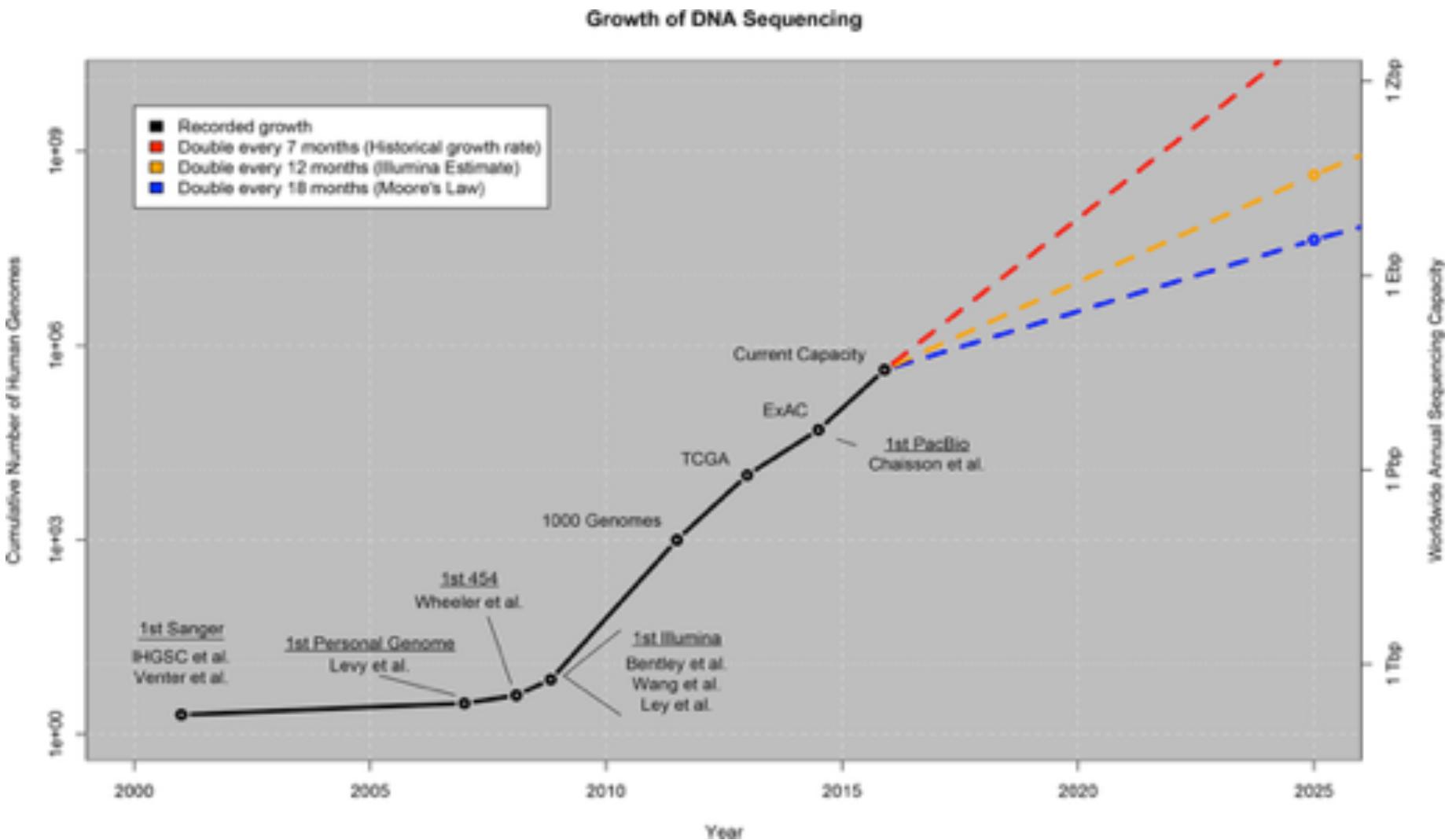
Computer  
resources  
free for  
Swedish  
researchers



---

## Bioinformatics as infrastructure

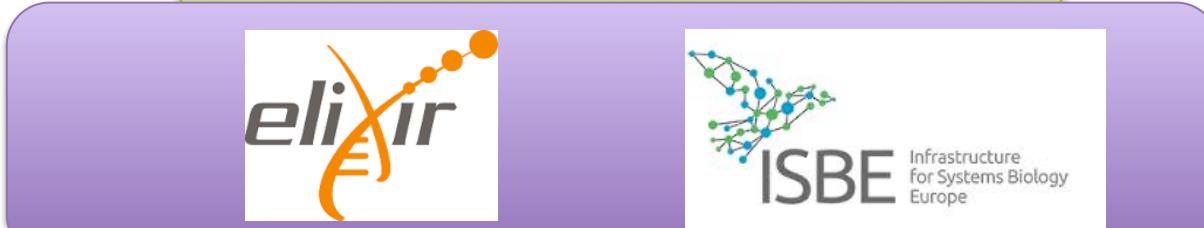
# Data growth



---

## NBIS activities

# Support, tools and training



- **Support and Infrastructure**

Wide competence in bioinformatics, Assembly/Annotation, SysDev

- **Long-term support (WABI)**

Large collaborative projects selected by scientific ranking

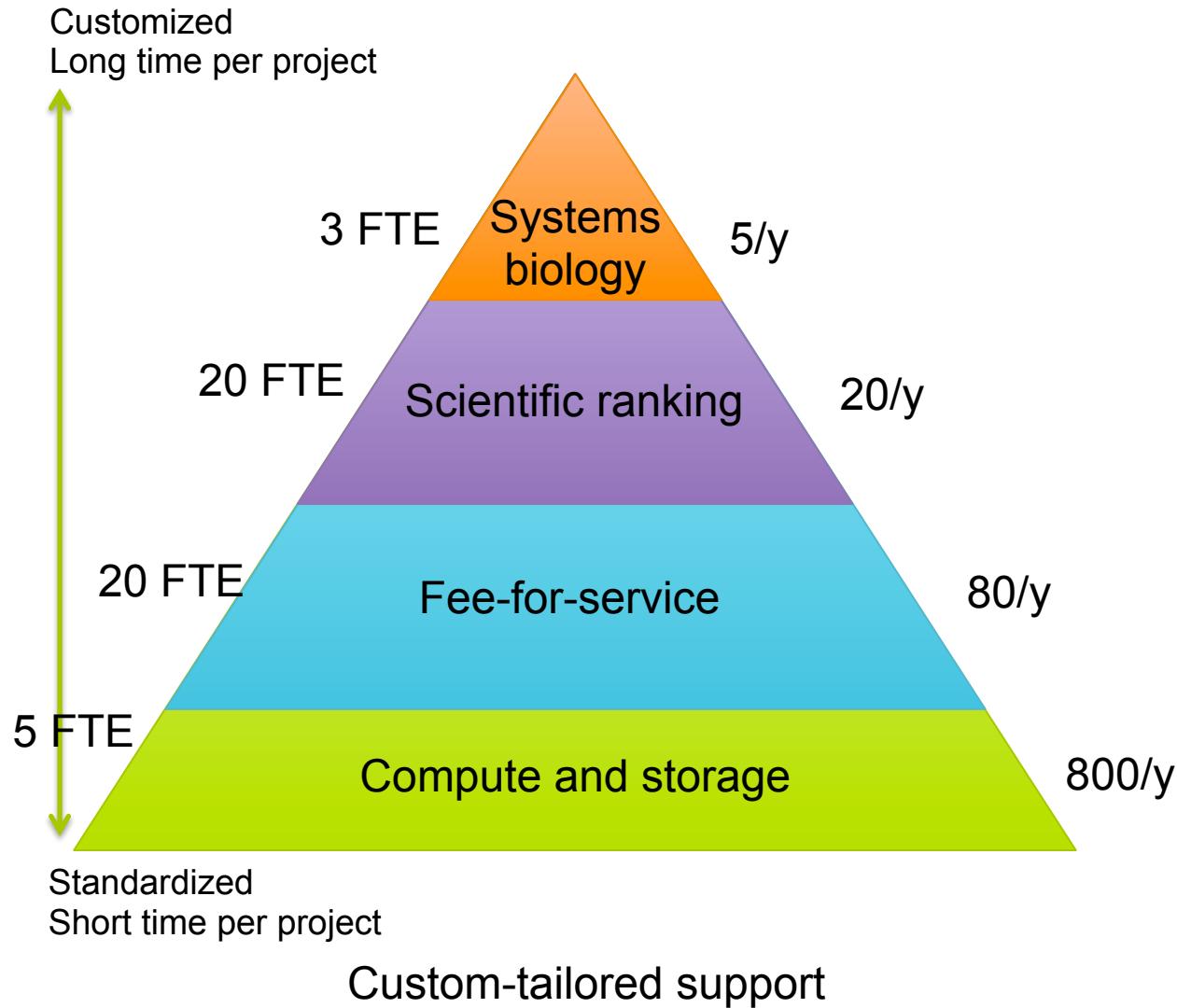
- **Systems biology**

Network analyses and Integrative bioinformatics

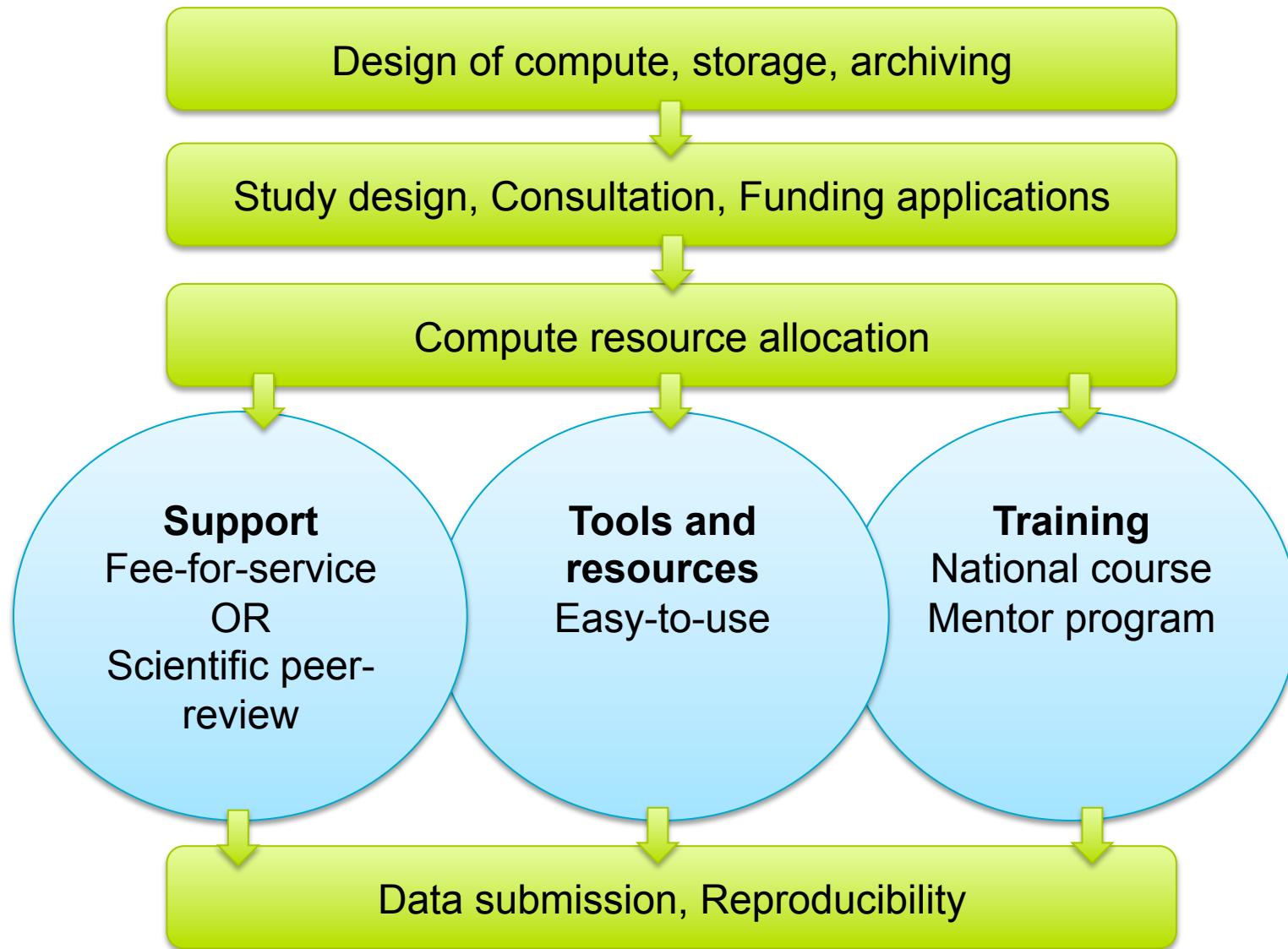
- **Compute and storage**

Computational and storage resources for bioinformatics, especially next-generation sequencing





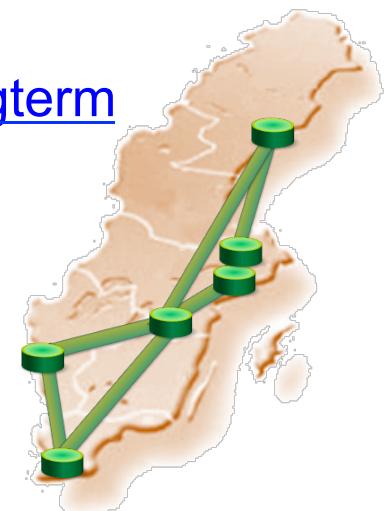
# User benefits





- Study design consultation (free)  
[www.nbis.se/support/supportform/index.php](http://www.nbis.se/support/supportform/index.php)  
+ drop-in sessions every week @ all 6 sites
- Short- and Medium-term support (User fee 800 SEK/h)  
[www.nbis.se/support/supportform/index.php](http://www.nbis.se/support/supportform/index.php)
- Long-term support and systems biology  
(500h, free, scientific evaluation)  
[www.nbis.se/support/supportform/index.php?form=longterm](http://www.nbis.se/support/supportform/index.php?form=longterm)

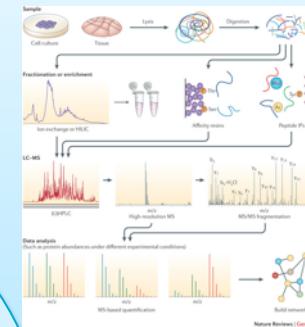
[www.scilifelab.se/platforms/bioinformatics/](http://www.scilifelab.se/platforms/bioinformatics/)  
[www.nbis.se](http://www.nbis.se)



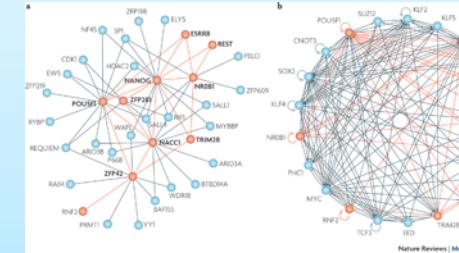
# Bioinformatics support



Genomics



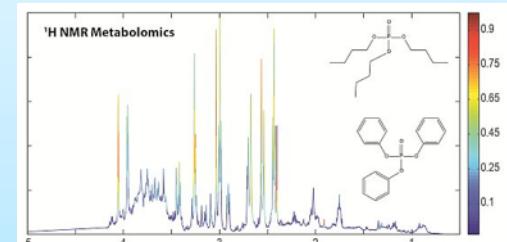
Proteomics



Systems biology



Biostatistics

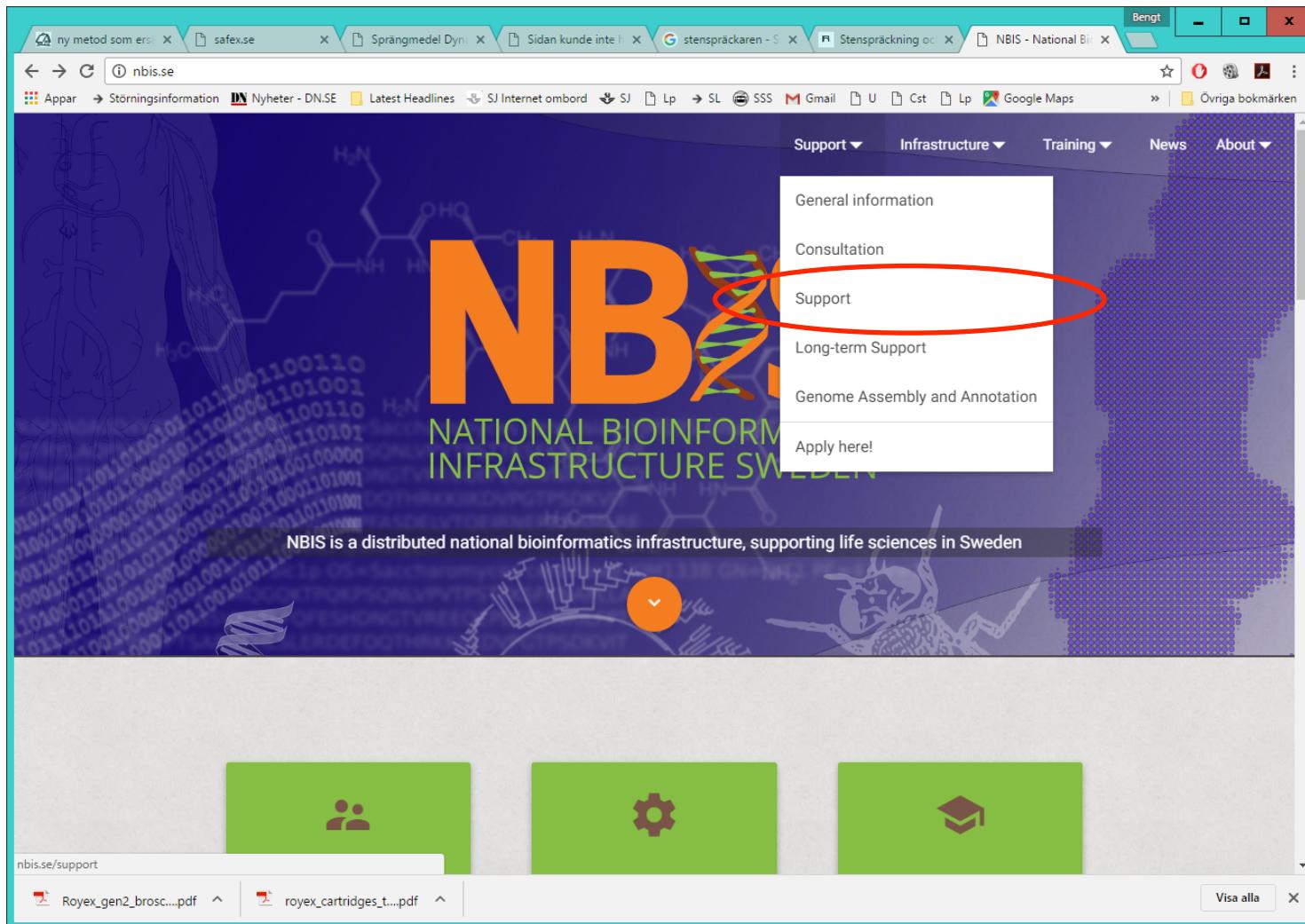


Metabolomics

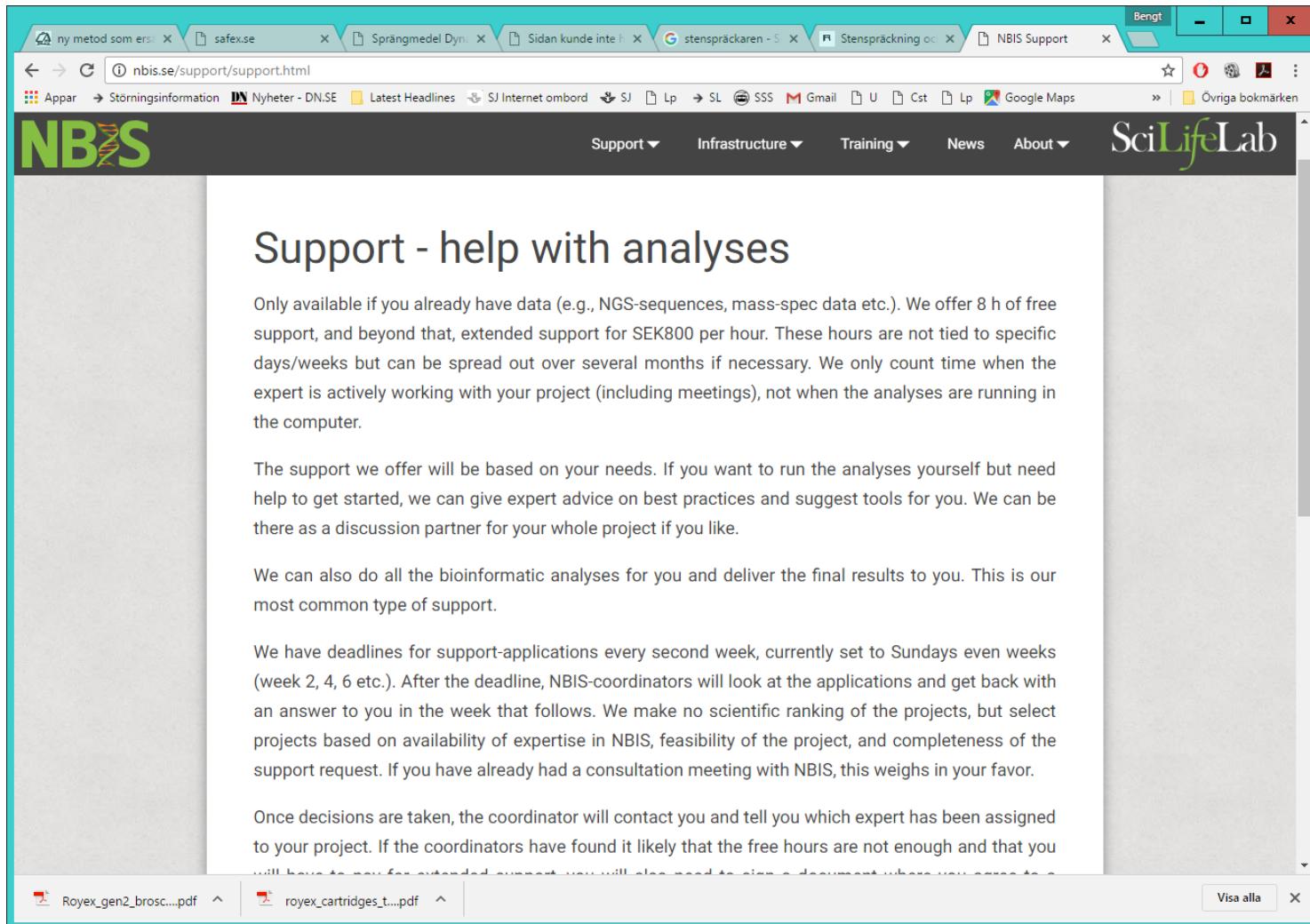
## 2 tracks!

- Fee-for-service (800kr/h)  
Rapid turnaround
- Scientific ranking (free)  
“Long-term Support”  
3 rounds/year

# How to get support nbis.se



The screenshot shows a web browser window with the URL [nbis.se](http://nbis.se) in the address bar. The page has a dark blue background featuring a DNA helix and various scientific illustrations. A navigation bar at the top includes links for Support, Infrastructure, Training, News, and About. A dropdown menu is open under the Support link, listing options: General information, Consultation, Support (which is circled in red), Long-term Support, Genome Assembly and Annotation, and Apply here!. Below the menu, there are three green buttons with icons: a person icon, a gear icon, and a graduation cap icon. At the bottom of the screen, there are two PDF files listed in a sidebar: "Royex\_gen2\_brosc....pdf" and "royex\_cartridges\_t....pdf".



The screenshot shows a web browser window with multiple tabs open at the top. The main content area displays the NBIS support page, which includes sections about support for analyses, deadlines, and contact information. At the bottom, there are links to PDF files for Royex\_gen2\_brosc...pdf and royex\_cartridges\_t...pdf.

**Support - help with analyses**

Only available if you already have data (e.g., NGS-sequences, mass-spec data etc.). We offer 8 h of free support, and beyond that, extended support for SEK800 per hour. These hours are not tied to specific days/weeks but can be spread out over several months if necessary. We only count time when the expert is actively working with your project (including meetings), not when the analyses are running in the computer.

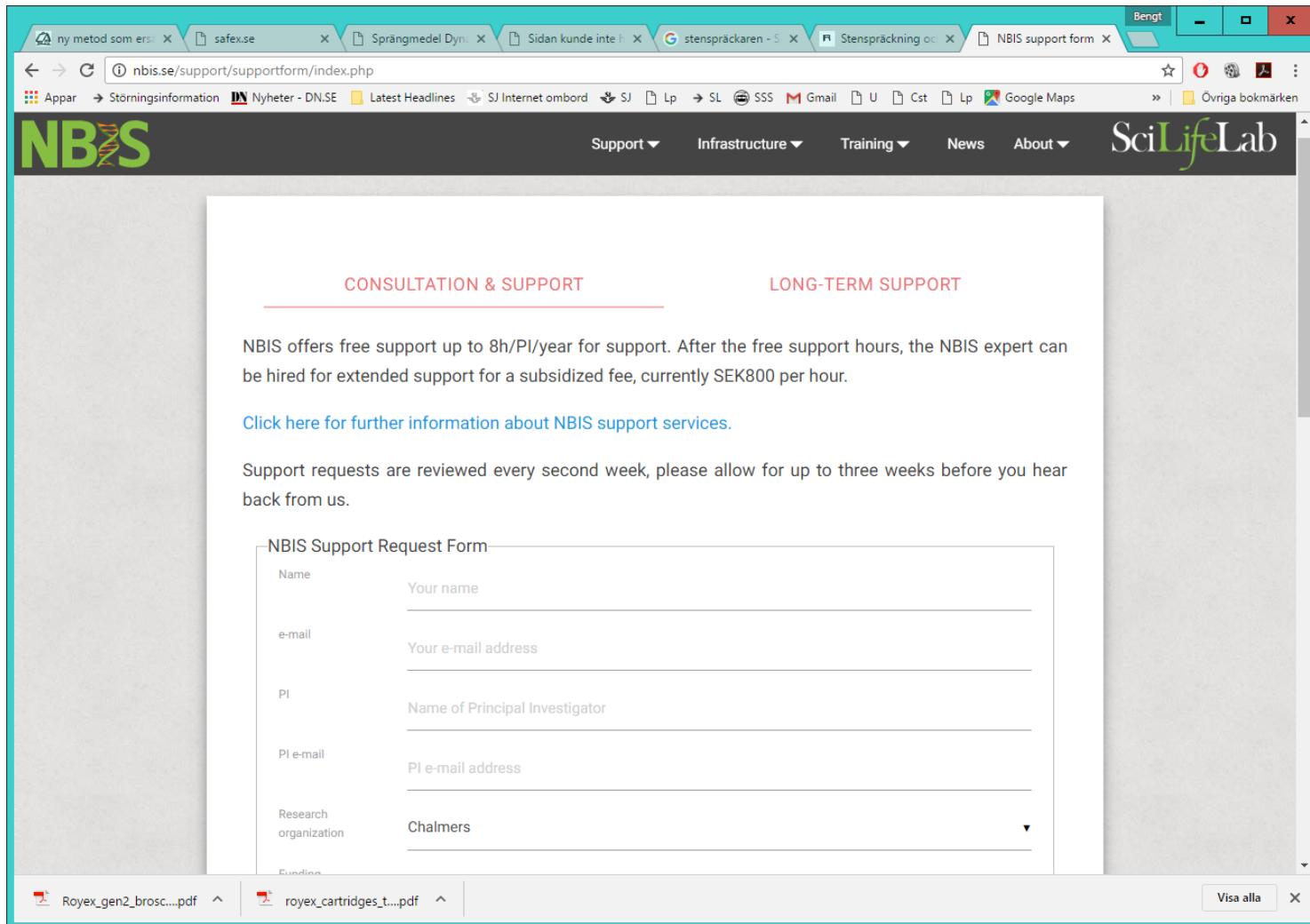
The support we offer will be based on your needs. If you want to run the analyses yourself but need help to get started, we can give expert advice on best practices and suggest tools for you. We can be there as a discussion partner for your whole project if you like.

We can also do all the bioinformatic analyses for you and deliver the final results to you. This is our most common type of support.

We have deadlines for support-applications every second week, currently set to Sundays even weeks (week 2, 4, 6 etc.). After the deadline, NBIS-coordinators will look at the applications and get back with an answer to you in the week that follows. We make no scientific ranking of the projects, but select projects based on availability of expertise in NBIS, feasibility of the project, and completeness of the support request. If you have already had a consultation meeting with NBIS, this weighs in your favor.

Once decisions are taken, the coordinator will contact you and tell you which expert has been assigned to your project. If the coordinators have found it likely that the free hours are not enough and that you

# Support form



The screenshot shows a web browser window with the NBIS support form page open. The page has a header with the NBIS logo and a navigation menu. Below the menu, there are two sections: 'CONSULTATION & SUPPORT' and 'LONG-TERM SUPPORT'. A note states that NBIS offers free support up to 8h/PI/year. A link is provided for further information about NBIS support services. Below this, a note says support requests are reviewed every second week and may take up to three weeks to receive a response. A 'NBIS Support Request Form' is displayed, containing fields for Name, e-mail, PI, PI e-mail, Research organization, and Funding. Two PDF files are shown in the bottom left corner: 'Royex\_gen2\_brosc...pdf' and 'royex\_cartridges\_t...pdf'. A 'Visa alla' button is at the bottom right.

**CONSULTATION & SUPPORT**

NBIS offers free support up to 8h/PI/year for support. After the free support hours, the NBIS expert can be hired for extended support for a subsidized fee, currently SEK800 per hour.

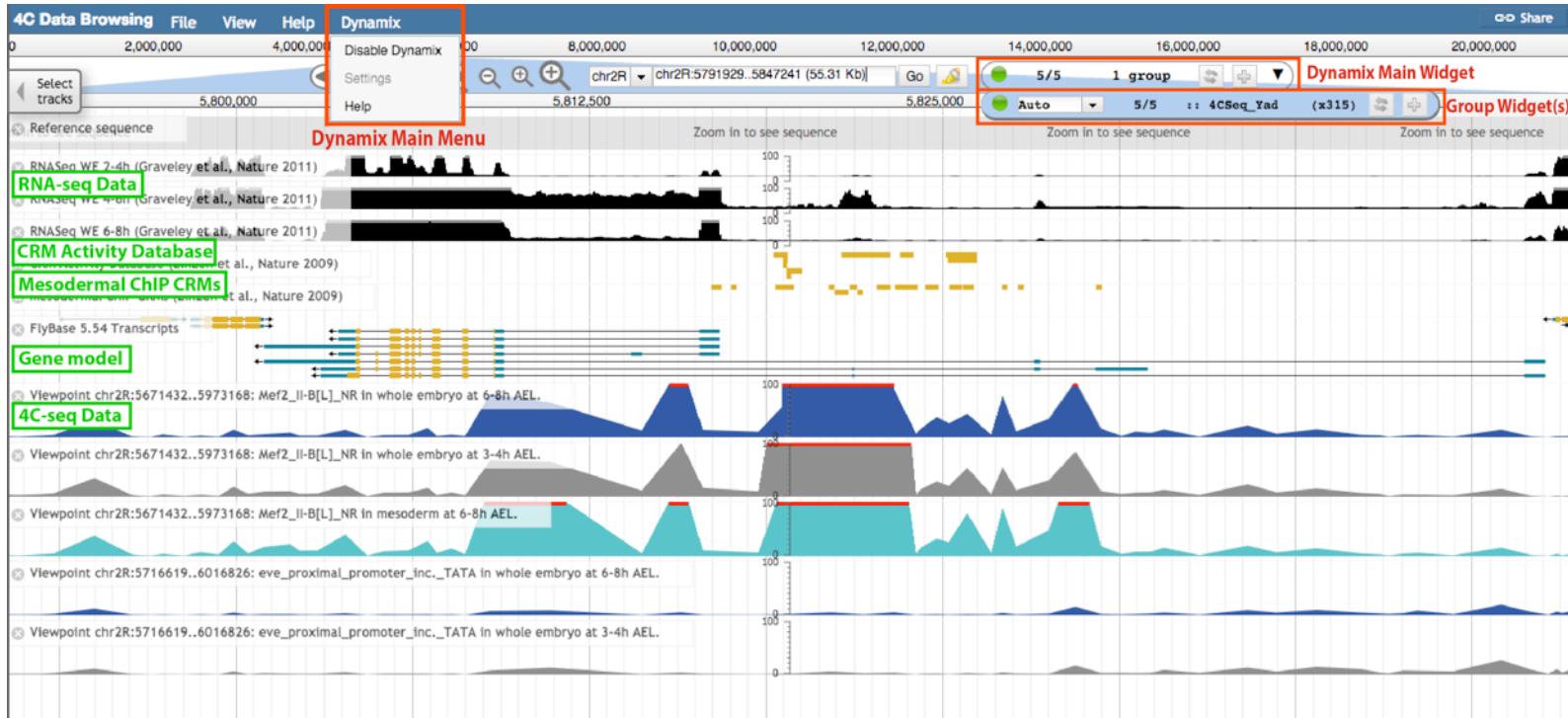
[Click here for further information about NBIS support services.](#)

Support requests are reviewed every second week, please allow for up to three weeks before you hear back from us.

**NBIS Support Request Form**

Name	Your name
e-mail	Your e-mail address
PI	Name of Principal Investigator
PI e-mail	PI e-mail address
Research organization	Chalmers
Funding	

Royex\_gen2\_brosc...pdf    royex\_cartridges\_t...pdf    Visa alla



- 10 - 20 projects per year
- Highly specialized staff and robust pipelines
- Tight user interaction
- Numerous manual and semi-manual QC steps
- Supports ENA submission
- Editable user interface

Henrik Lantz



Cost effective with high quality!

**4 FTE, joint effort by Long-term Support and Systems Biology**  
Projects apply in the regular Long-term Support calls

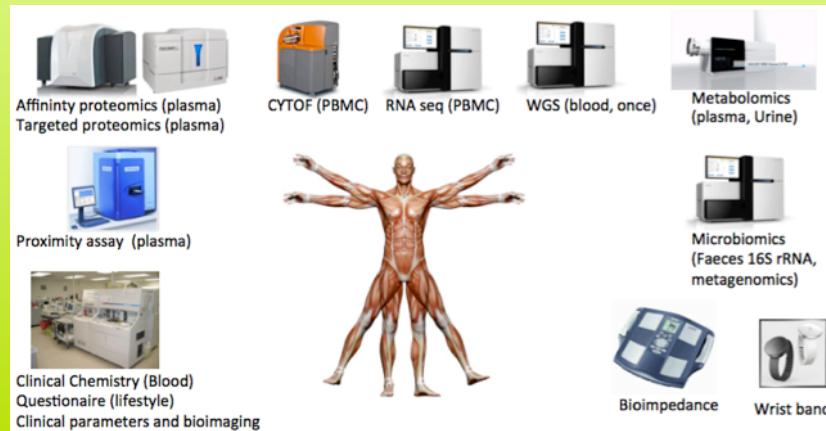
## Combine data from SciLifeLab platforms

- Building **tools** and **resources** for handling very large and/or complex biological data sets
- Typically performed in the context of longer support projects
- State-of-the-art analytical methods for integrating multi-modal biological data sets, eg
  - Machine learning/deep learning
  - Graph-based models
  - Genome-scale metabolic models

### Support track for integrative projects

First call Feb 2016; First few projects initiated

Involves extensive integration of data





# Tools and infrastructure

## Compute and storage of sensitive data

- Local EGA
- ePouta integration pilot
- Mosler
- Pouta Blueprints
- web-servers with EGI cloud vo.NBIS.se

## WGS tools and resources

- SweGen 1000 genomes
- WGS somatic variant calling WF
- WGS structural variation WF

## Software maintenance

- MrBayes
- Structure prediction web services

## Assembly and annotation

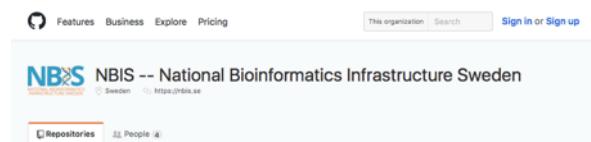
- Falcon on Milou
- ENA submission help

## Other tools and resources

- Human Metabolic Atlas (HMA)
- Haloplex variant calling pipeline
- WhatsHap: Genomic phasing
- IgDiscover: Immunorepertoire

NBIS SYSDEV BETS BOARD						<a href="http://bit.ly/nbis_bets_board">http://bit.ly/nbis_bets_board</a>					
NOW						NEXT					
Rank	Bet/Project	Ext	PO	PL	PM	Rank	Bet/Project	Ext	PO	PL	PM
1	Local EGA	x	NJ?	6?		1	Tools and guides collection	PE	?	<1	
2	SweFreq	x	Adam A	NJ	<1	2					Impute GW
3	ePouta integration pilot	x	Antti P	JH?	1	3					Permutatio
4	microMosler		NJ?	JH	1?	4					Developme
5	WGS structural variation WF		BN	PO	1	5					Improve bi
6	Falcon on milou		HL	JB	<1	6					Single-cell
7	Pouta Blueprints		OS	JH	2						Chip-seq p
8	web-servers with EGI cloud vo.NBIS.se		MB	NS	2						CWL for NI
9	Human Metabolic Atlas (HMA)		TS	LH	1-6						OpenCB
10	SSC HPC2N region dynamic resource allocation		MB	NS	2						Local Ense
11	Database for SciLife/AZ proj	x	BPh?	JV?	4?						NBIS Tools
Ongoing dev projects not resourced by the dev team											
<a href="#">WGS somatic variant calling WF</a> BN (SJ) 0											
<a href="#">Haloplex variant calling pipeline</a> PE ML 0											
<a href="#">WhatsHap: Genomic phasing tool</a> PE MM 0											

Open prioritization  
and background  
descriptions

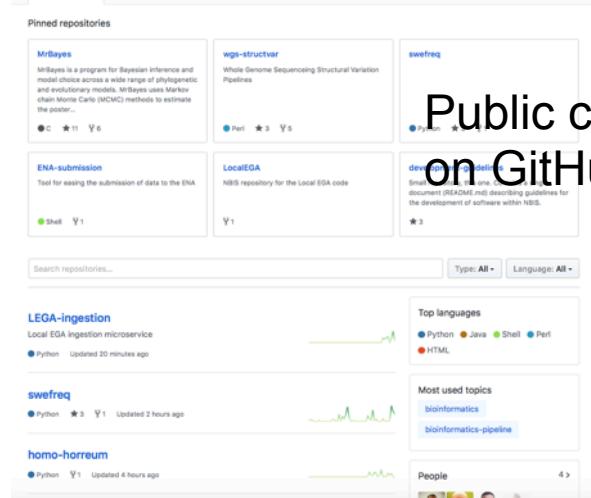


NBIS -- National Bioinformatics Infrastructure Sweden

Repositories People (4)

Pinned repositories

- MrBayes
- wgs-structvar
- swefreq
- ENA-submission
- LocalEGA
- development-tools



Top languages

- Python
- Java
- Shell
- Perl
- HTML

Most used topics

- bioinformatics
- bioinformatics-pipeline

People (4)

Public code repositories  
on GitHub

# SweFreq – Swedish frequency database

SciLifeLab

- Data dissemination – increasing the impact of data generated by SciLifeLab facilities
- Currently: **SweGen WGS Reference Variant Frequency Dataset**
  - 950 twin registry + 50 Northern Sweden
  - Deep coverage WGS (30X)
- More datasets to come

[swefreq.nbis.se](http://swefreq.nbis.se)

Funding: SciLifeLab  
Sequencing: NGI  
Variant calling: NGI  
QC: NBIS  
Data access interface: NBIS

## Swedish Frequency Database

The SweGen dataset contains whole-genome sequences for 1000 individuals sampled from the Swedish population. A detailed description of the project can be found here [add link to BioRxiv article later]. This server hosts the variant frequency data generated within the SweGen project. This data can be used as a reference for the genetics research community and clinical genetics laboratories. Individual positions in the genome can be viewed using either the ExAC Browser or the Data Beacon by clicking the links above.

To access the variant frequency file you need to register.

### Genome browser

Coverage metric: Average

Region: 22 / 46615715 / 46615880

Coverage summary

Position	Coverage
46,615,720	~30
46,615,740	~30
46,615,760	~30
46,615,780	~30
46,615,800	~30
46,615,820	~30
46,615,840	~30
46,615,860	~30
46,615,880	~30

Genes

Variants

Export table to CSV

Variant	Chrom	Position	Consequence	Filter	Annotation	Flags	Alele Count	Alele Number	Number of Homozygotes	Alele Frequency
22:46615732 A/G (n=380714)	22	46615732	p.Arg19Gln	PASS	missense		1	1000	0	0.000000
22:46615822 A/G (n=1614120)	22	46615822	p.Lys239Glu	PASS	missense		1	1000	0	0.000000
22:46615880 T/C (n=190324)	22	46615880	p.Val227Ile	PASS	missense		7	1000	0	0.000217

### File download

Data Beacon ExAC Browser Download Data Niclas Jareborg Logout niclas.jareborg@scilifelab.se

## SweFreq Agreement

In case the agreement is short we can write it right here. Otherwise I can make a link to it.  
Some more text, bla bla Lorem ipsum dolor sit amet, eu ius graeco perfecto, velit delicatissimi an est,  
rebum persequitur mel no. Stet everti sea ne. Utam laoreet interpretaris vix et, ex qui solum  
abhorreant. Vis detracto liberavisse te, sea odio possim instructior te, vidit scripta ex his.  
I hereby consent to the agreement:

[Download](#)

### Beacon

SweFreq Data Beacon ExAC Browser Download Data Niclas Jareborg Logout niclas.jareborg@scilifelab.se

22 46615880 C hg19 SweGen Is it there? Help

Yes

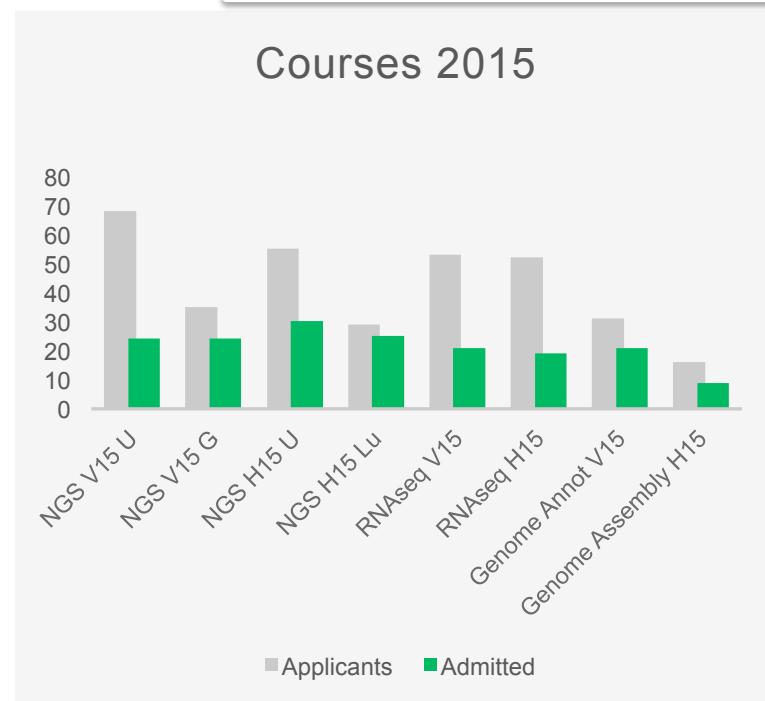
swefreq.nbis.se:4000/query?dataset=SweGen

{"query": {"position": "4661534", "dataset": "SweGen", "reference": "dummy", "chromosome": "22", "allele": "G"}, "beacon": "nbis-beacon", "response": {"observed": 0, "exists": true}}



- Bioinformatics Drop-In
  - Weekly at all sites – initial consultations
- 20-odd courses every year
  - Introduction to Bioinformatics using NGS
  - Introduction to Linux
  - Perl programming
  - Introduction to genome annotation
  - Introduction to multivariate analysis
  - RNA-seq
  - Advanced workshop on NGS data analysis
  - Advanced functional genomics
  - Advanced bioinformatics
- Additional local activities
- Bioinformatics Advisory programme
  - Mentorship in bioinformatics

Gender balance:  
54% female / 46% male



From spring 2017, we plan to double our training efforts to match the increased demands from the scientific community

[www.scilifelab.se/education/courses/](http://www.scilifelab.se/education/courses/)  
[www.nbis.se/training/events.html](http://www.nbis.se/training/events.html)

## The Swedish Bioinformatics Advisory Program

PhD students get a senior bioinformatician as a personal advisor during 2 years of their PhD. Monthly project meetings + two grand meetings per year to aid networking and knowledge transfer.

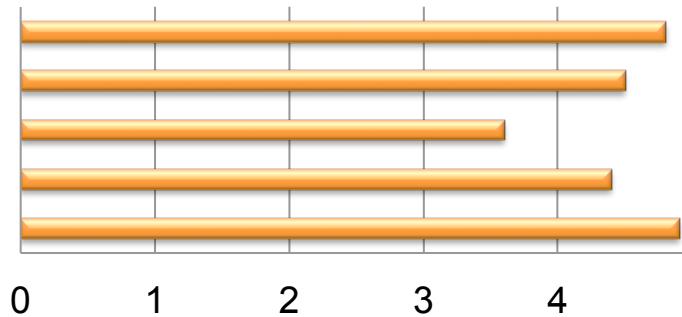
[www.scilifelab.se/education/mentorship/the-swedish-bioinformatics-advisory-program/](http://www.scilifelab.se/education/mentorship/the-swedish-bioinformatics-advisory-program/)

Recent call (2017/2018): 111 applicants for 15 places (!)

### The Swedish Bioinformatics Advisory Program

Student evaluation, June 2015

- Overall rating of the Advisory Program
- Impact on the efficacy of your research
- Impact on the scientific value of your
- Impact on the technical level of your
- In favour of SciLifeLab continuing this

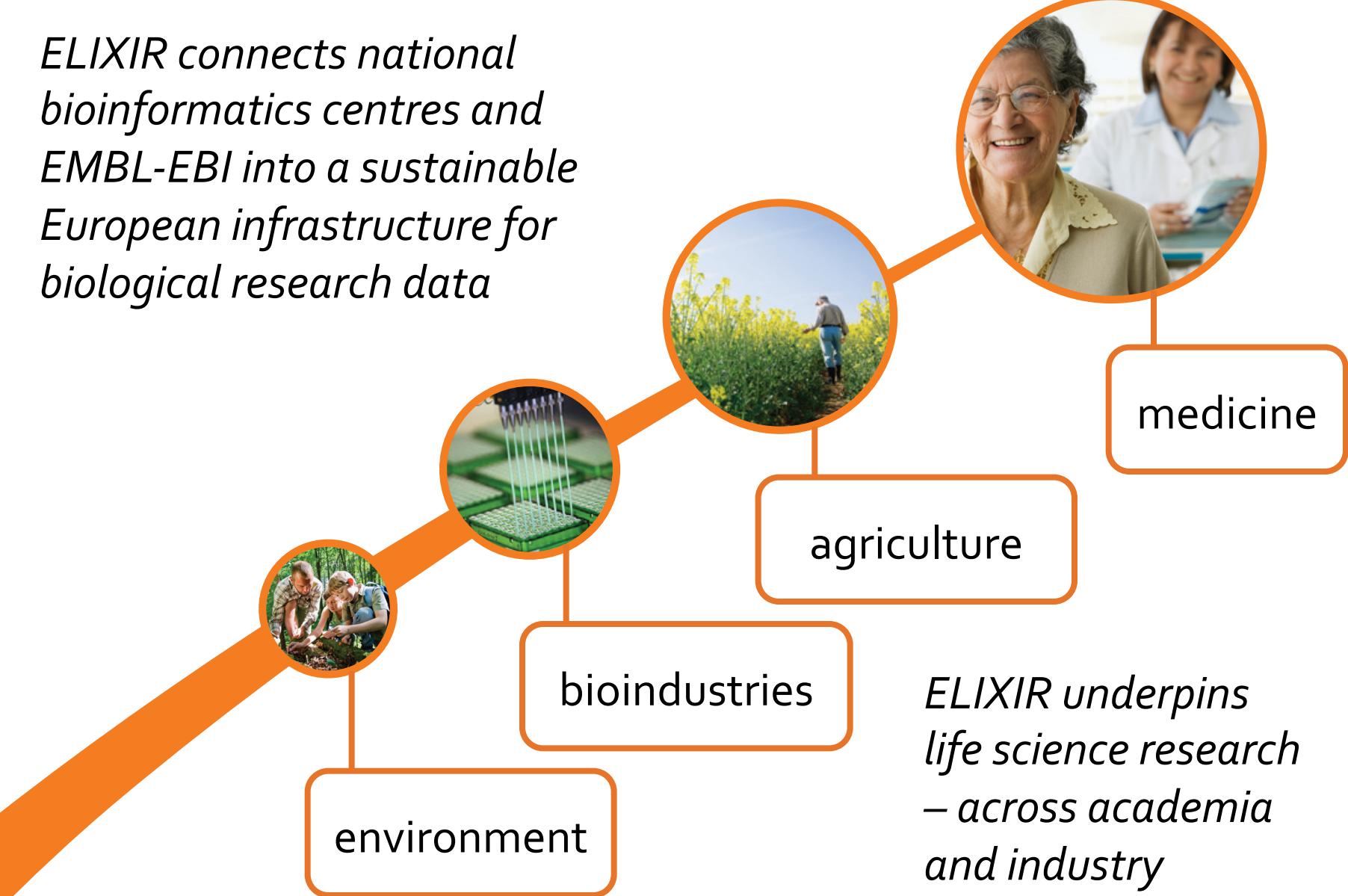


Elixir

# Why ELIXIR?

- Creating a robust infrastructure for biological information is a **bigger** task than any individual organisation or nation can take on alone
- These are issues of such complexity that no single institution or country can tackle alone
- Biology has by far the largest research community:
  - ~3 million life science researchers in Europe
  - >7 million web hits a day at EMBL-EBI alone

*ELIXIR connects national  
bioinformatics centres and  
EMBL-EBI into a sustainable  
European infrastructure for  
biological research data*



# *NBIS coordinates the Swedish node in ELIXIR – the European infrastructure for biological information*

- 

## **Data**

*Sustain core data resources*

- 

## **Tools**

*Services & connectors to drive  
access and exploitation*

- 

## **Compute**

*Access, Exchange & Compute on  
sensitive data*

- 

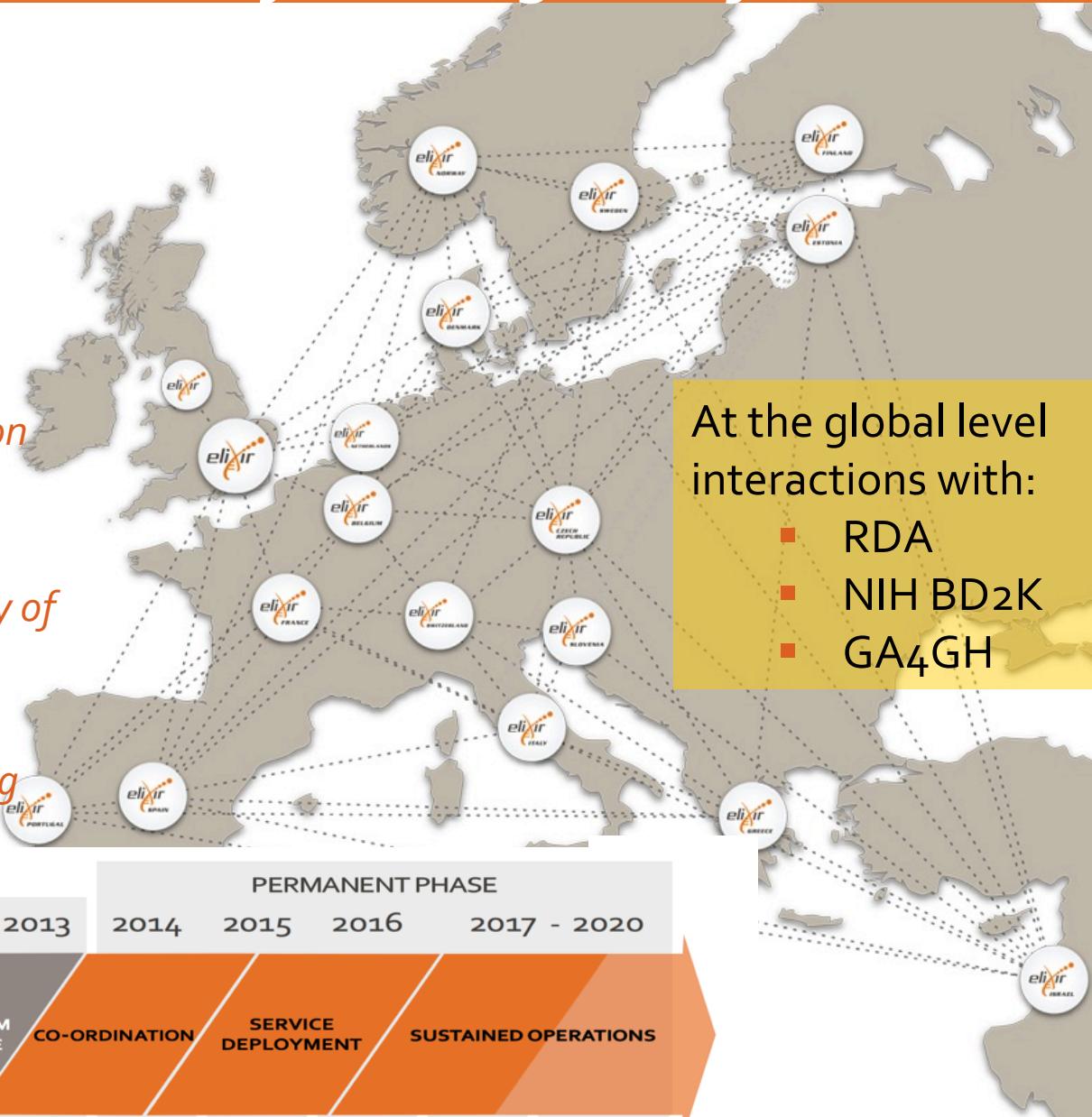
## **Standards**

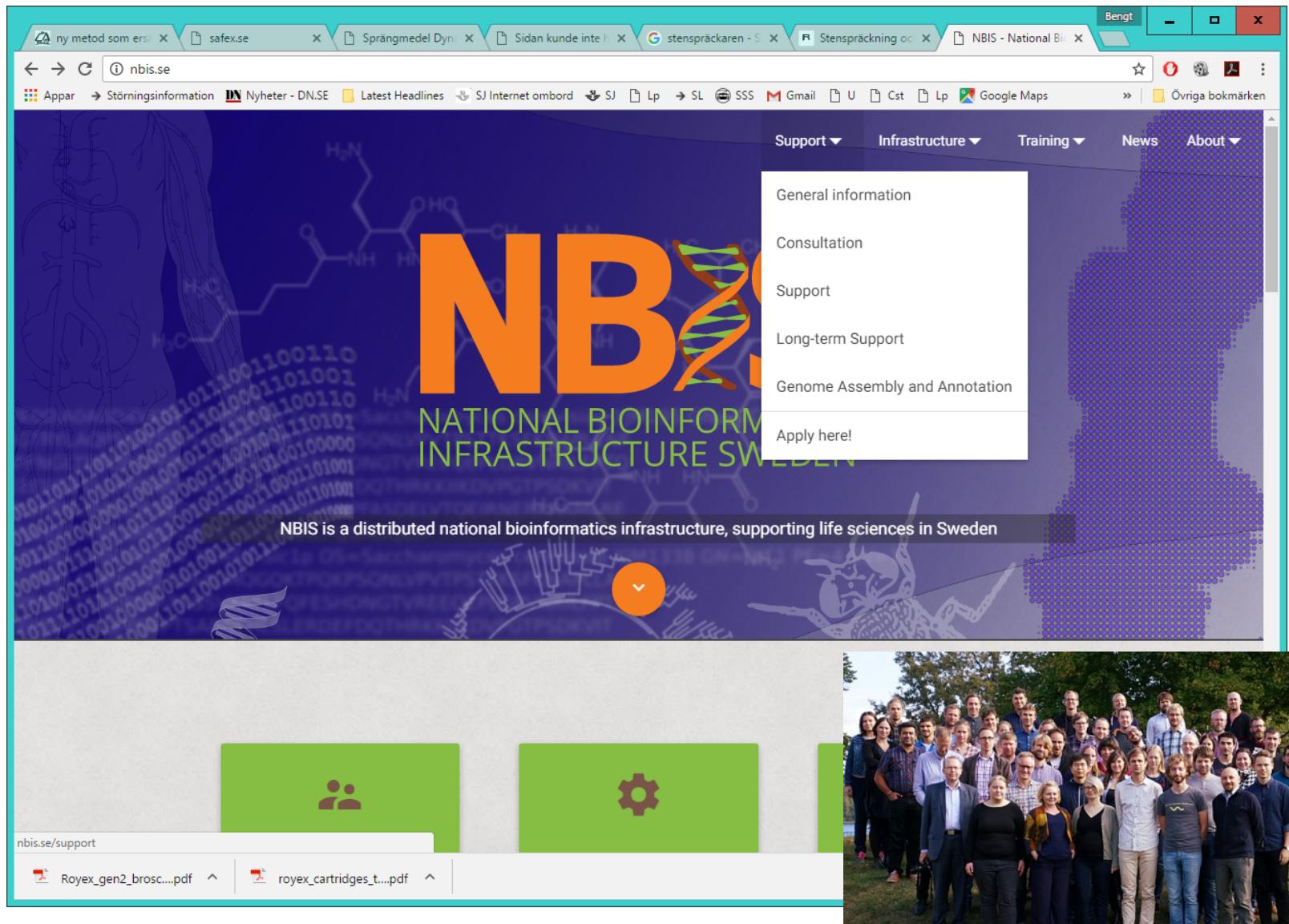
*Integration and interoperability of  
data and services.*

- 

## **Training**

*Professional skills for managing  
and exploiting data*





The screenshot shows a web browser window with the NBIS website loaded. The main header features the NBIS logo and the text "We're here for you! nbis.se". Below the header is a large orange banner with the text "NBIS NATIONAL BIOINFORMATICS INFRASTRUCTURE SWEDEN". A sub-header below the banner reads "NBIS is a distributed national bioinformatics infrastructure, supporting life sciences in Sweden". To the right of the banner is a navigation menu with links to "Support", "Infrastructure", "Training", "News", and "About". A dropdown menu under "Support" is open, listing "General information", "Consultation", "Support", "Long-term Support", "Genome Assembly and Annotation", and "Apply here!". At the bottom of the page, there are two green buttons with icons: one for users and one for a gear. A sidebar on the left contains links to "nbis.se/support" and PDF files named "Royex\_gen2\_brosc....pdf" and "royex\_cartridges\_t....pdf". On the right side of the page is a large group photo of the NBIS team.

ny metod som ers... safex.se Språngmedel Dyn... Sidan kunde inte h... stenspräckaren - S... Stenspräckning oc... NBIS - National Bi... Bengt

nbis.se

Appar Störningsinformation DN Nyheter - DN.SE Latest Headlines SJ Internet ombord SJ Lp SL SSS Gmail U Cst Lp Google Maps Övriga bokmärken

Support Infrastructure Training News About

General information

Consultation

Support

Long-term Support

Genome Assembly and Annotation

Apply here!

NBIS NATIONAL BIOINFORMATICS INFRASTRUCTURE SWEDEN

NBIS is a distributed national bioinformatics infrastructure, supporting life sciences in Sweden

nbis.se/support

Royex\_gen2\_brosc....pdf royex\_cartridges\_t....pdf

