

---

# SciLifeLab Bioinformatics Platform

National Bioinformatics Infrastructure Sweden (NBIS)

***Niclas Jareborg***

*NGS course*

*Göteborg 2017-05-19*

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

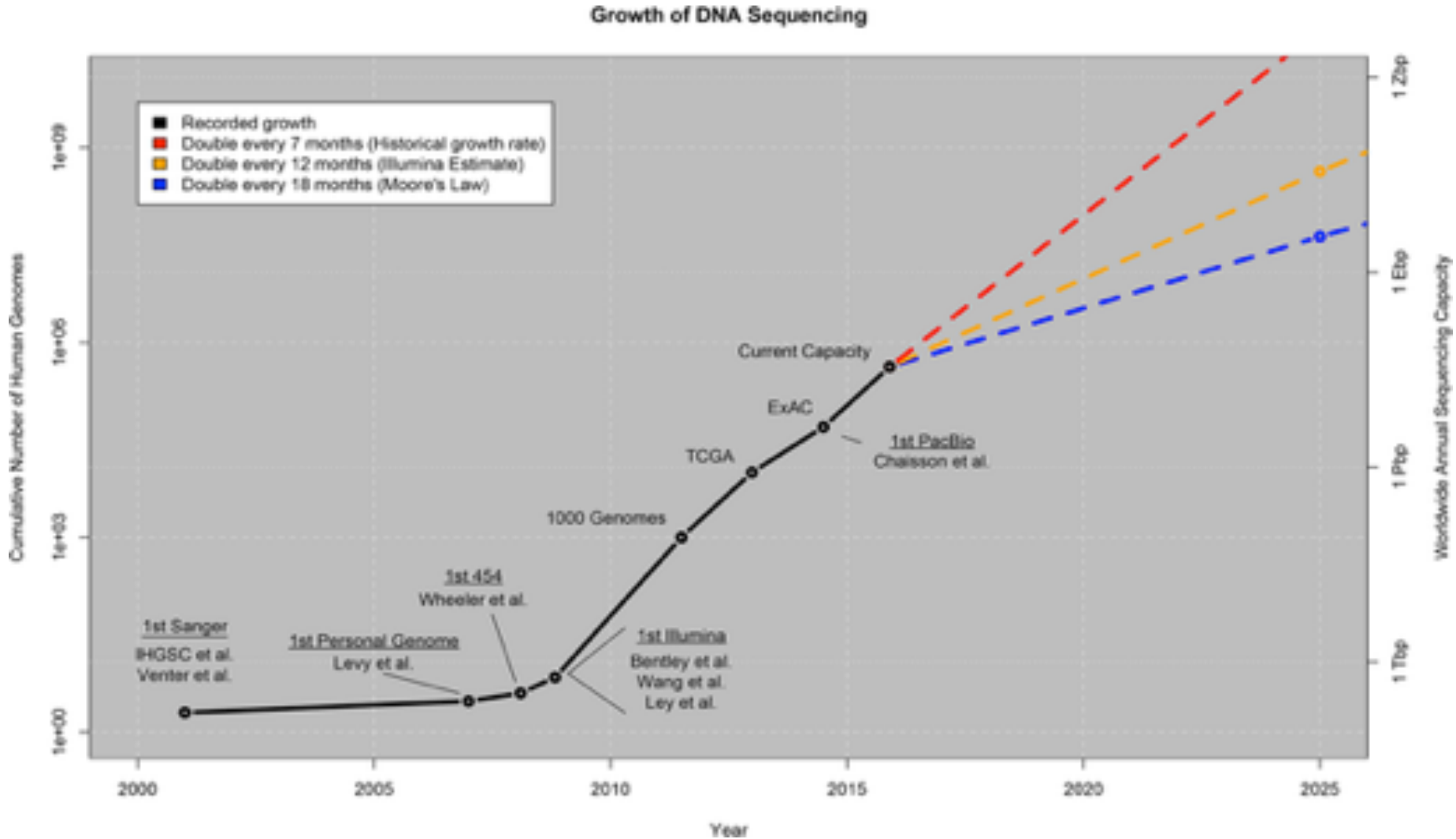


Computer resources free for Swedish researchers



## Bioinformatics as infrastructure

# Data growth



## NBIS activities

# Support, tools and training



Support



Tools



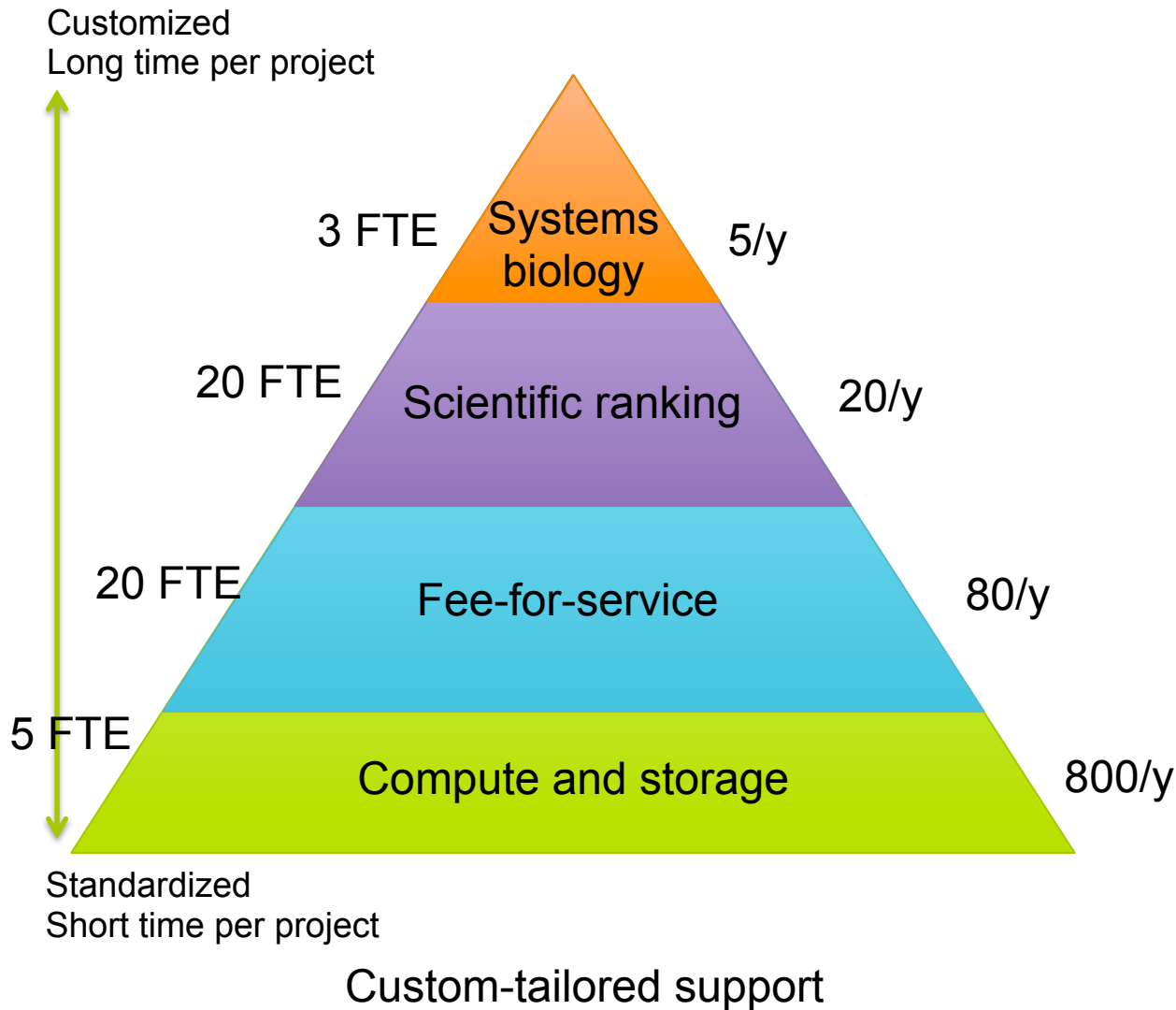
Training





- **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





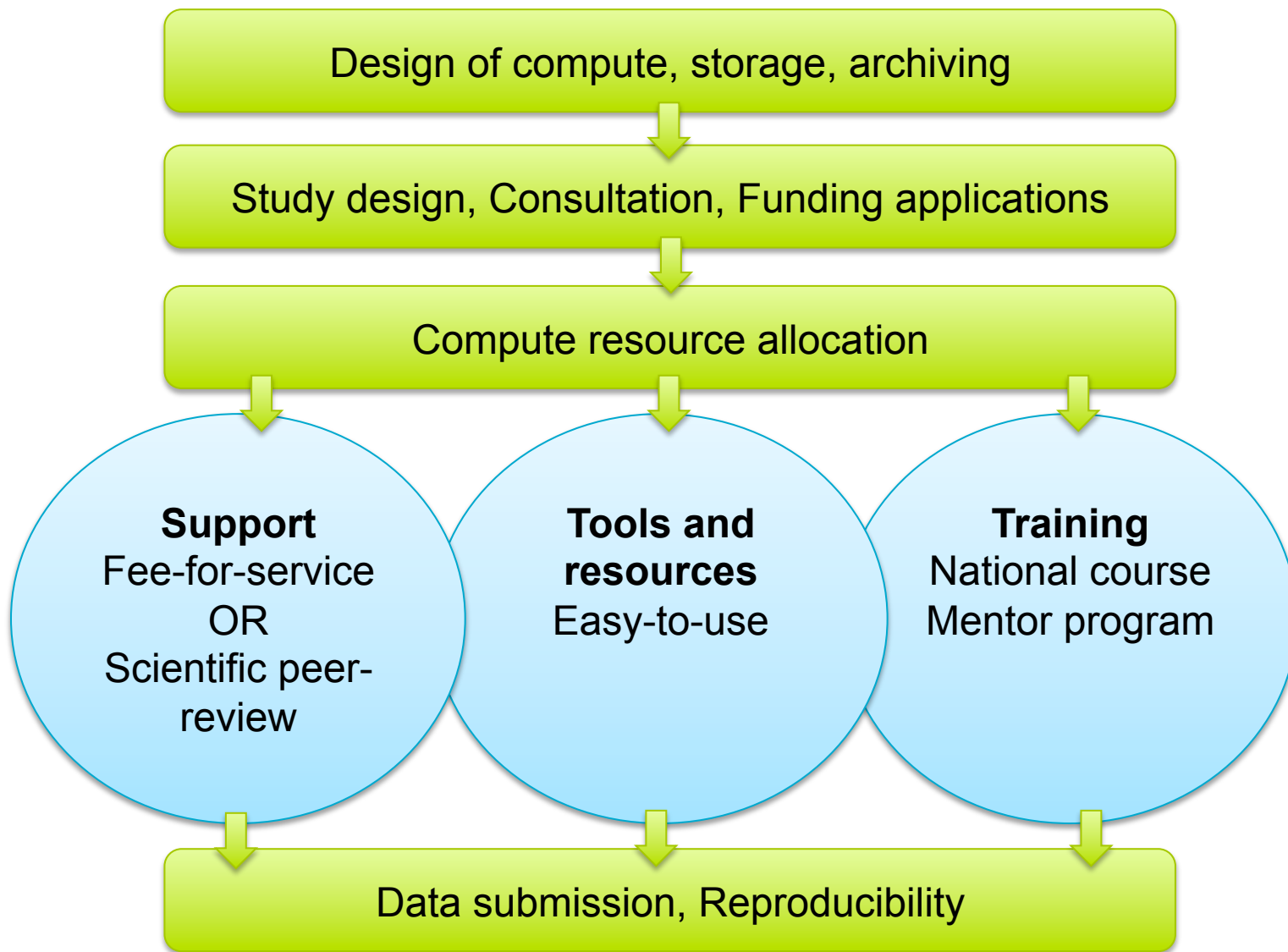
Training



5 FTE  
Systems development



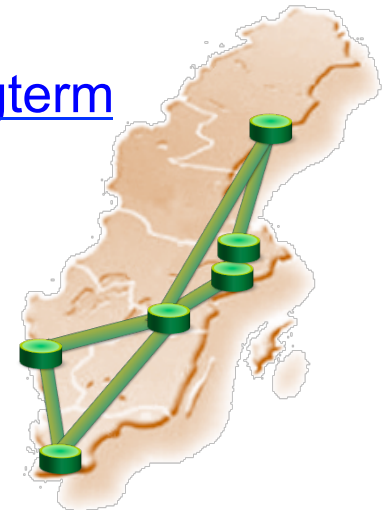
1 FTE  
Data management





- 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)

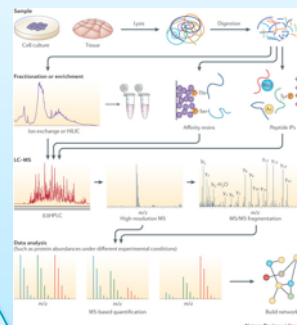
[www.scilifelab.se/platforms/bioinformatics/](http://www.scilifelab.se/platforms/bioinformatics/)  
[www.nbis.se](http://www.nbis.se)



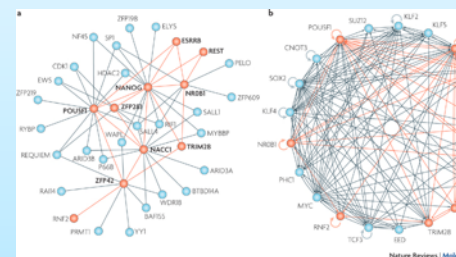
# Bioinformatics support



Genomics



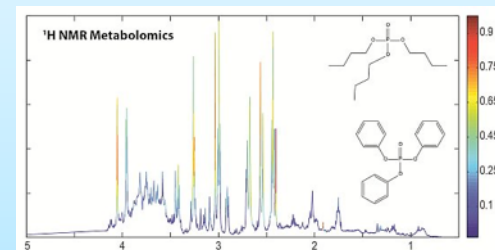
Proteomics



Systems biology



Biostatistics



Metabolomics

## 2 tracks!

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

The screenshot shows a web browser window displaying the nbis.se website. The browser's address bar shows 'nbis.se'. The website's navigation menu is open, showing the following options: Support, Infrastructure, Training, News, and About. The 'Support' option is highlighted with a red oval. Below the navigation menu, the website features a large banner with the text 'NBIS NATIONAL BIOINFORMATICS INFRASTRUCTURE SWEDEN' and a sub-header 'NBIS is a distributed national bioinformatics infrastructure, supporting life sciences in Sweden'. Below the banner, there are three green buttons with icons representing people, a gear, and a graduation cap. The browser's taskbar at the bottom shows several open tabs, including 'nbis.se/support' and two PDF files.

The screenshot shows a web browser window with the URL `nbis.se/support/support.html`. The page features a dark navigation bar with the NBIS logo on the left and the SciLifeLab logo on the right. The navigation menu includes links for Support, Infrastructure, Training, News, and About. The main content area is titled "Support - help with analyses" and contains several paragraphs of text.

**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 will have to pay for extended support, you will also need to sign a document where you agree to...



ny metod som ers: X safex.se X Sprängmedel Dyn X Sidan kunde inte laddas X stenspräckaren - S X Stenspräckning oc X NBIS support form X Bengt

nbis.se/support/supportform/index.php

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

**NBIS** Support ▾ Infrastructure ▾ Training ▾ News About ▾ SciLifeLab

**CONSULTATION & SUPPORT** **LONG-TERM 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

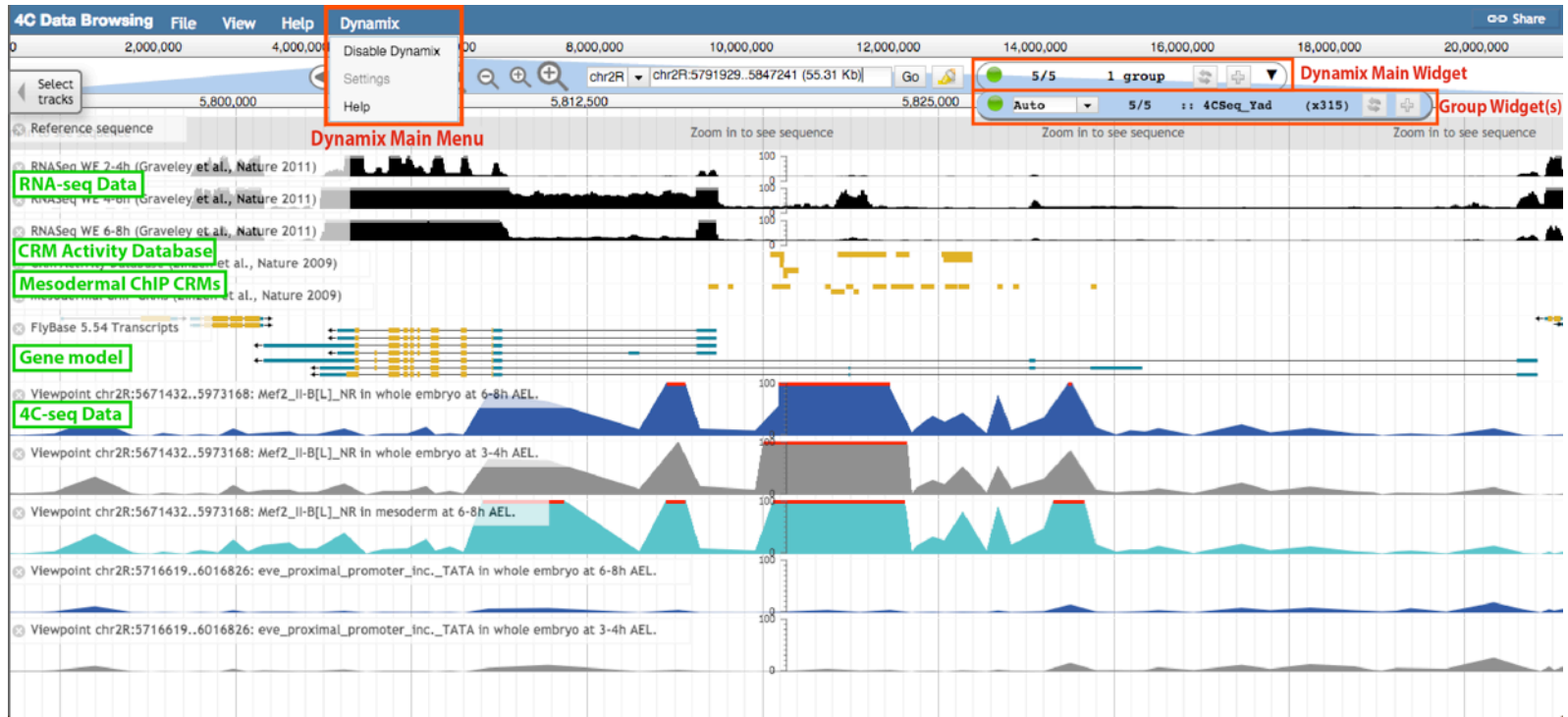
e-mail

PI

PI e-mail

Research organization

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



- 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

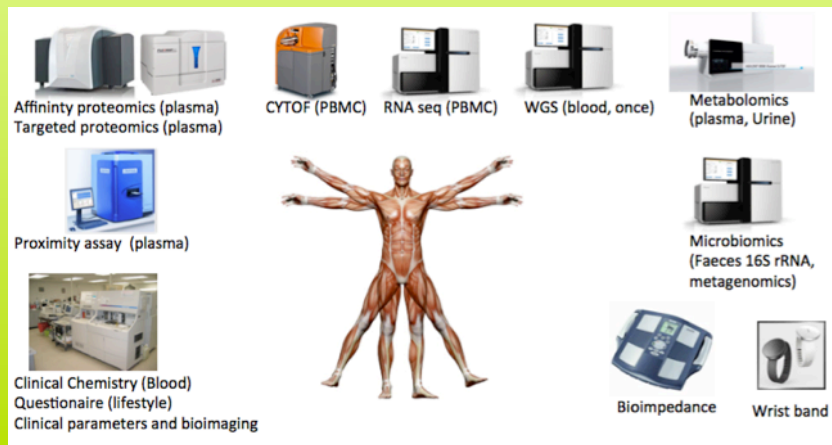
### Combine data from SciLifeLab platforms

- 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





## Compute and storage of sensitive data

- Local EGA
- ePouta integration pilot
- microMosler
- 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** [http://bit.ly/nbis\\_bets\\_board](http://bit.ly/nbis_bets_board)

*NBIS should enable world-class life science by providing expert knowledge, creative data integration, advanced training, efficient data publication and analysis methods.*

| NOW  |  |     |         |      |     | NEXT |   |     |    |    |                             |
|--|--|-----|---------|------|-----|------|---|-----|----|----|-----------------------------|
| <i>We have done the experimenting, we have learned, and we're ready to build it for the world. It is fully funded, plans are clear and we can predict with reasonable confidence what impact the Bet may have.</i> |  |     |         |      |     |      |   |     |    |    |                             |
| Rank   | Bet/Project  | Ext | PO      | PL   | PM  | Rank | Bet/Project                                 | Ext | PO | PL | PM                          |
| 1  | <a href="#">Local EGA</a>                                    | x   | NJ?     | ? 6? |     | 1    | <a href="#">Tools and guides collection</a> | PE  | ?  | <1 | <a href="#">Impute GV</a>   |
| 2  | <a href="#">SweFreq</a>                                      | x   | Adam A  | NJ   | <1  | 2    |   |     |    |    | <a href="#">Permutatio</a>  |
| 3  | <a href="#">ePouta integration pilot</a>                     | x   | Artti P | JH?  | 1   | 3    |   |     |    |    | <a href="#">Developme</a>   |
| 4  | <a href="#">microMosler</a>                                  |     | NJ?     | JH   | 1?  | 4    |   |     |    |    | <a href="#">Improve bi</a>  |
| 5  | <a href="#">WGS structural variation WF</a>                  |     | BN      | PO   | 1   | 5    |   |     |    |    | <a href="#">Single-cell</a> |
| 6  | <a href="#">Falcon on milou</a>                              |     | HL      | JB   | <1  | 6    |   |     |    |    | <a href="#">Chip-seq p</a>  |
| 7  | <a href="#">Pouta Blueprints</a>                             |     | OS      | JH   | 2   |      |   |     |    |    | <a href="#">CWL for N</a>   |
| 8  | <a href="#">web-servers with EGI cloud vo.NBIS.se</a>        |     | MB      | NS   | 2   |      |   |     |    |    | <a href="#">OpenCB</a>      |
| 9  | <a href="#">Human Metabolic Atlas (HMA)</a>                  |     | TS      | LH   | 1-6 |      |   |     |    |    | <a href="#">Local Ense</a>  |
| 10   | <a href="#">SSC HPC2N region dynamic resource allocation</a> |     | MB      | NS   | 2   |      |   |     |    |    | <a href="#">NBIS Tools</a>  |
| 11   | <a href="#">Database for SciLife/AZ proj</a>                 | x   | BPn?    | JV?  | 4?  |      |   |     |    |    | <a href="#">IgDiscover</a>  |

Ongoing dev projects not resourced by the dev team

|  |   |  |    |     |   |
|--|---|--|----|-----|---|
|  | <a href="#">WGS somatic variant calling WF</a>    |  | BN | (S) | 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

Features Business Explore Pricing This organisation Search Sign in or Sign up

**NBIS** NBIS -- National Bioinformatics Infrastructure Sweden  
Sweden <https://nbis.se>

Repositories People

Pinned repositories

**MrBayes**

MrBayes is a program for Bayesian inference and model choice across a wide range of phylogenetic and evolutionary models. MrBayes uses Markov chain Monte Carlo (MCMC) methods to estimate the poster...

11 3 5

**wgs-structural**

Whole Genome Sequencing Structural Variation Pipelines

1 3 5

**swefreq**

...

1 3

**ENA-submission**

Tool for easing the submission of data to the ENA

1 1

**LocalEGA**

NBIS repository for the Local EGA code

1 1

**development**

...

3

Search repositories... Type: All Language: All

**LEGA-ingestion**

Local EGA ingestion microservice

Python Updated 20 minutes ago

**Top languages**

Python Java Shell Perl HTML

**Most used topics**

bioinformatics  
bioinformatics-pipeline

**swefreq**

Python 3 1 Updated 2 hours ago

**homo-horreum**

Python 1 Updated 4 hours ago

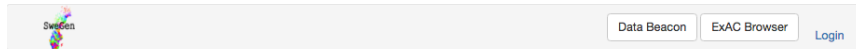
Public code repositories on GitHub

# SweFreq – Swedish frequency database

- **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)

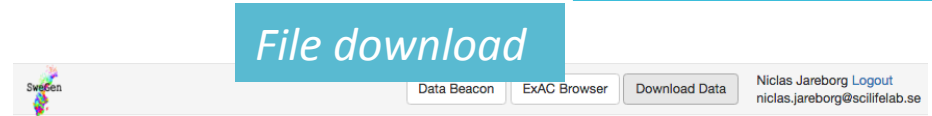
Funding: SciLifeLab  
 Sequencing: NGI  
 Variant calling: NGI  
 QC: NBIS  
 Data access interface: NBIS



## 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.

## 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 persequeris mel no. Stet everti sea ne. Utamur laoreet interpretaris vix et, ex qui solum abhorreat. Vis detracto liberavisse te, sea odio possim instructor te, vidit scripta ex his.

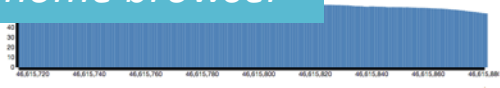
I hereby consent to the agreement:

[Download](#)

## Genome browser

Region: 22 / 46615715 / 46615880

Coverage summary



Genes

Variants

| Variant                        | Chrom | Position | Consequence | Filter | Annotation | Page | Allele Count | Allele Number | Number of Homozygotes | Allele Frequency |
|--------------------------------|-------|----------|-------------|--------|------------|------|--------------|---------------|-----------------------|------------------|
| 22-46615732 A > G (rs1387874)  | 22    | 46615732 | p.Arg1702v  | PASS   | missense   | 1    | 1988         | 0             | 0.000000              |                  |
| 22-46615822 A > G (rs14814126) | 22    | 46615822 | p.Lys2092v  | PASS   | missense   | 1    | 1988         | 0             | 0.000000              |                  |
| 22-46615880 T > C (rs1982234)  | 22    | 46615880 | p.Val2274v  | PASS   | missense   | 7    | 1988         | 0             | 0.003521              |                  |

## Beacon

SweFreq

22 46615880 C hg19 SweGen

[Is it there?](#) [Help](#)

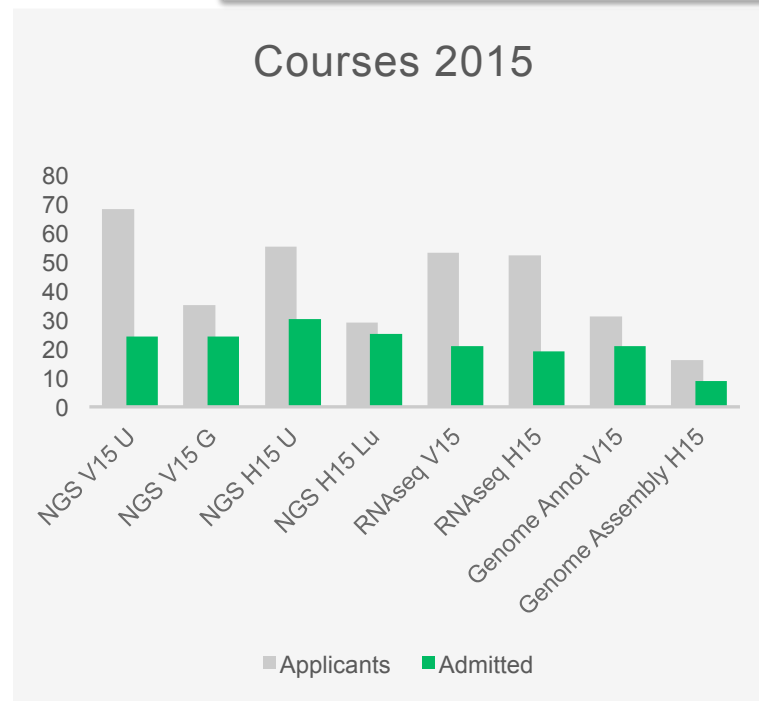
Yes

```
{ "query": { "position": "46541534", "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

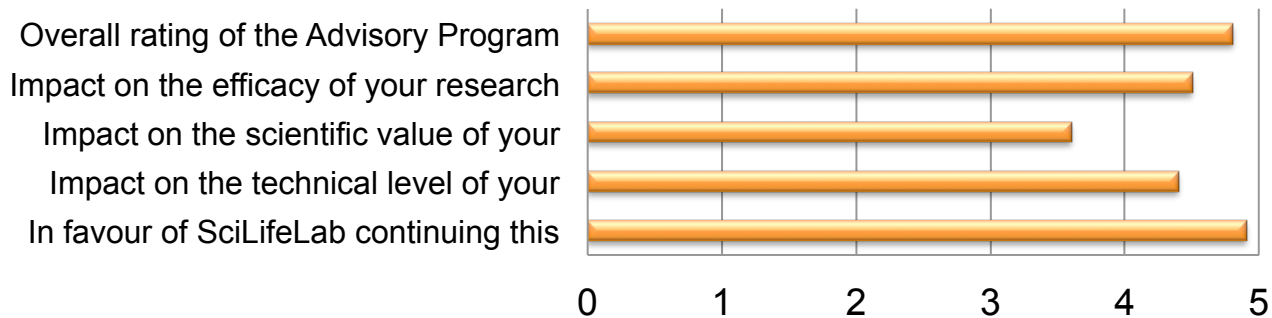
PhD students get a senior bioinformatician as a personal advisor during 2 years of their PhD. Monthly project meetings + two grand meetings per year to aid networking and knowledge transfer.

[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



Elixir

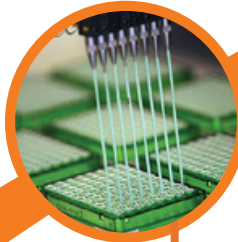
# Why ELIXIR?

- 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*



environment



bioindustries



agriculture



medicine

*ELIXIR underpins life science research – across academia and industry*



# 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**

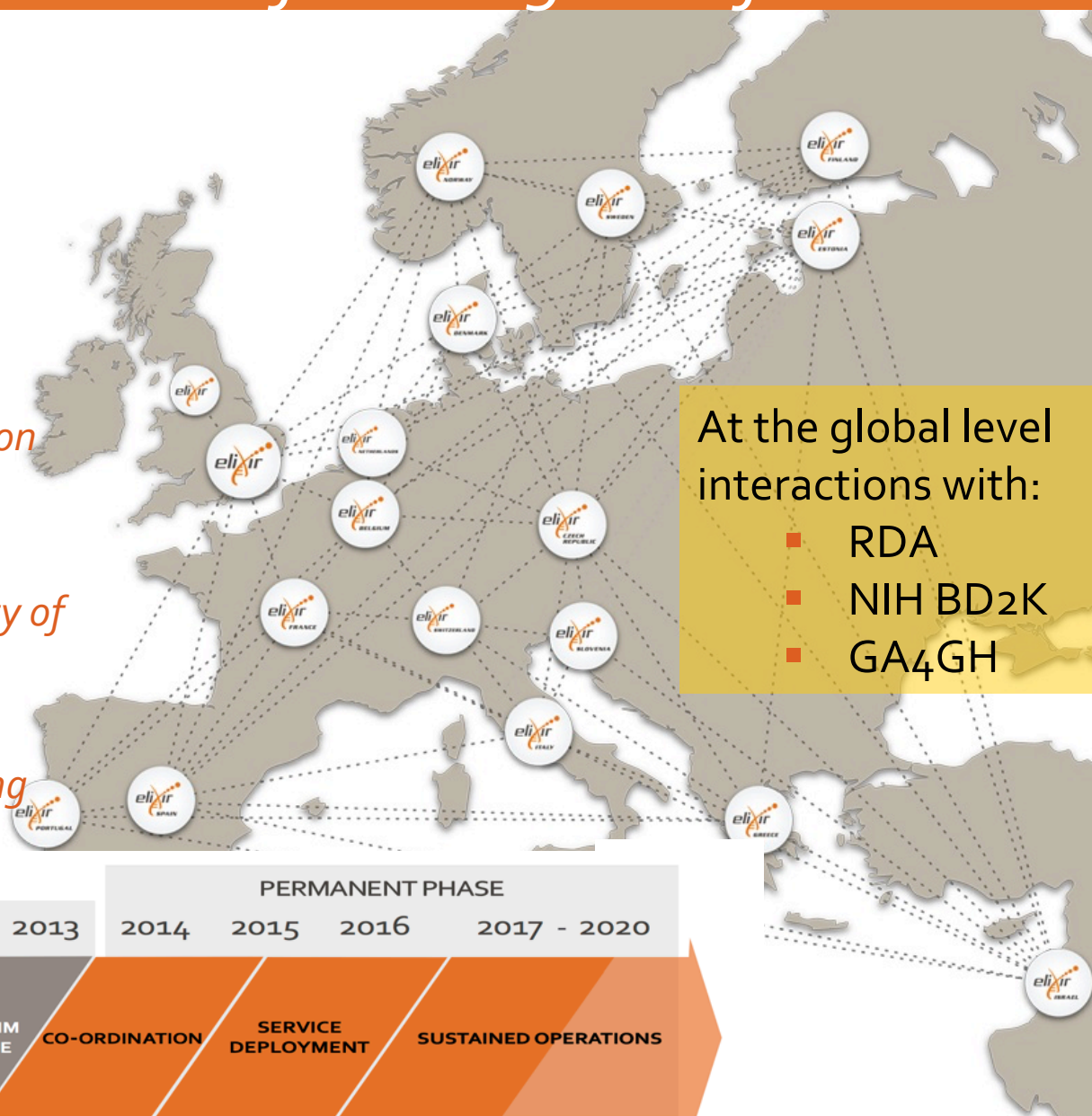
*Access, Exchange & Compute on sensitive data*

- **Standards**

*Integration and interoperability of data and services.*

- **Training**

*Professional skills for managing and exploiting data*



At the global level interactions with:

- RDA
- NIH BD2K
- GA4GH

2007 - 2012

PREPARATORY PHASE

2013

INTERIM PHASE

2014

CO-ORDINATION

2015

SERVICE DEPLOYMENT

2016

SERVICE DEPLOYMENT

2017 - 2020

SUSTAINED OPERATIONS

PERMANENT PHASE

The image shows a screenshot of a web browser displaying the nbis.se website. The browser's address bar shows 'nbis.se'. The website has a dark blue header with the 'NBIS' logo in large orange letters and 'NATIONAL BIOINFORMATICS INFRASTRUCTURE SWEDEN' in green below it. A navigation menu is open, listing: Support, Infrastructure, Training, News, and About. A dropdown menu under 'Support' lists: General information, Consultation, Support, Long-term Support, Genome Assembly and Annotation, and Apply here!. Below the header, a banner reads: 'NBIS is a distributed national bioinformatics infrastructure, supporting life sciences in Sweden'. At the bottom of the screenshot, there are two green buttons with icons: a person icon and a gear icon. A small group photo of people is visible in the bottom right corner of the screenshot. The browser's taskbar at the bottom shows two open PDF files: 'Royex\_gen2\_brosc...pdf' and 'royex\_cartridges\_t...pdf'.

