
SciLifeLab Bioinformatics Platform

National Bioinformatics Infrastructure Sweden (NBIS)

Björn Nystedt
www.nbis.se

RNA course

Uppsala 13.03.2017

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

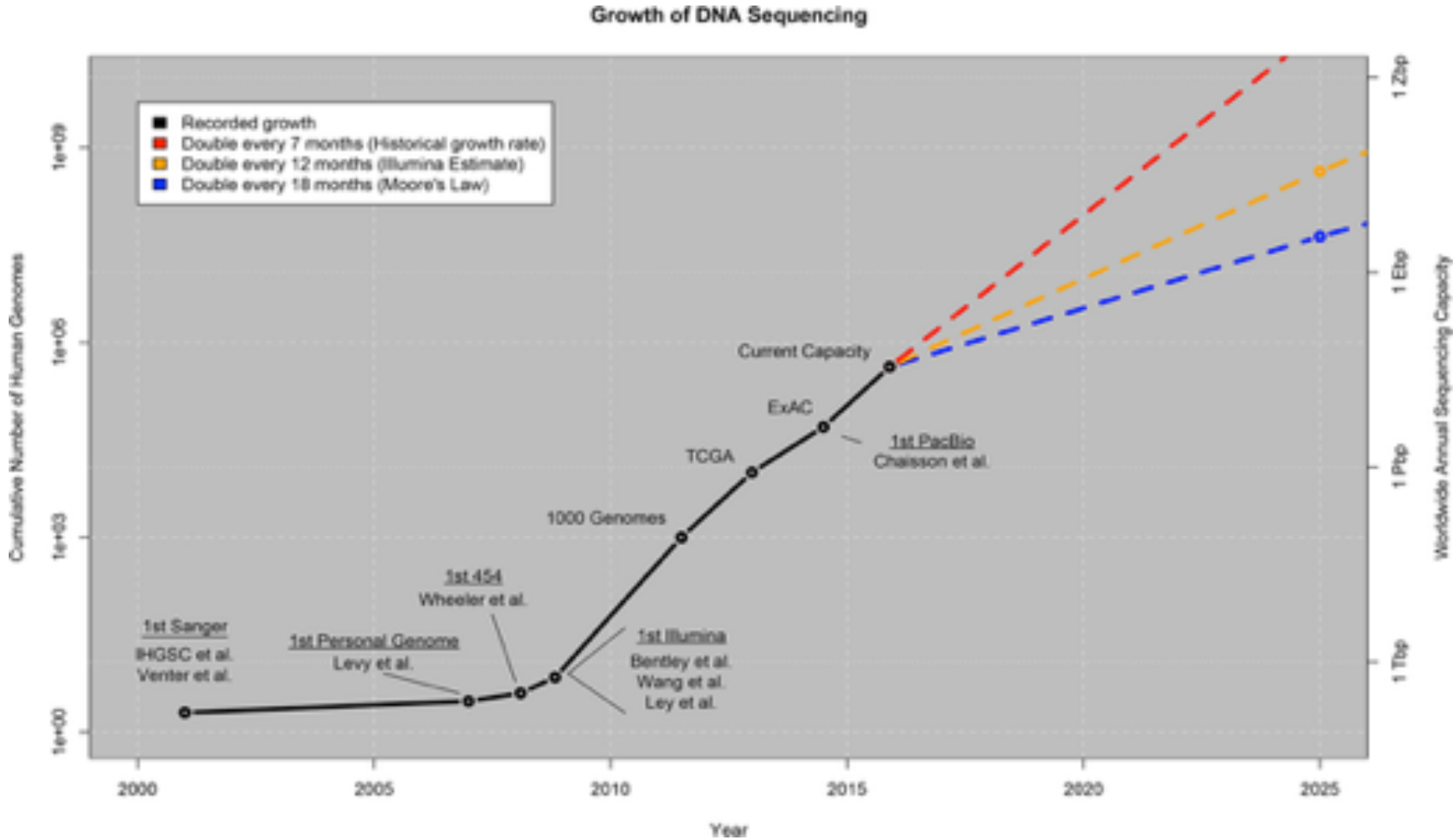
SNIC



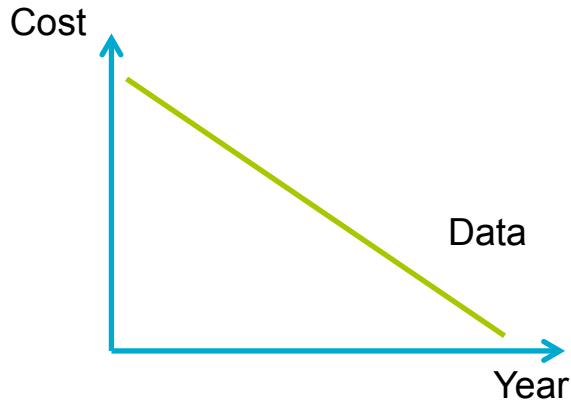
Computer resources free for Swedish researchers

Bioinformatics as infrastructure

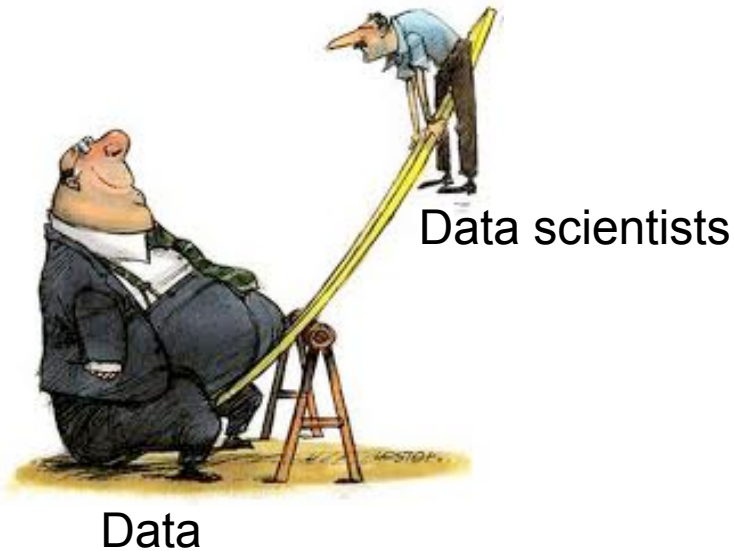
Data growth



Production is cheap, analysis is not



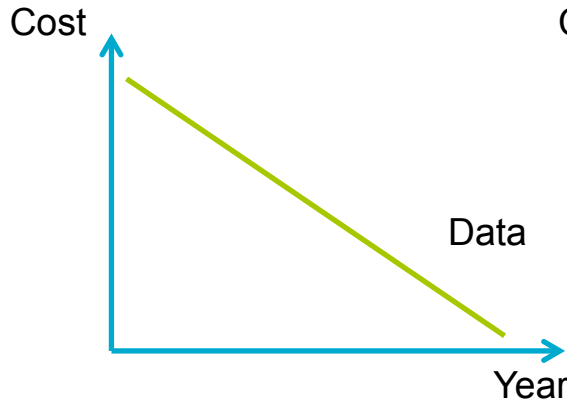
“Per base”



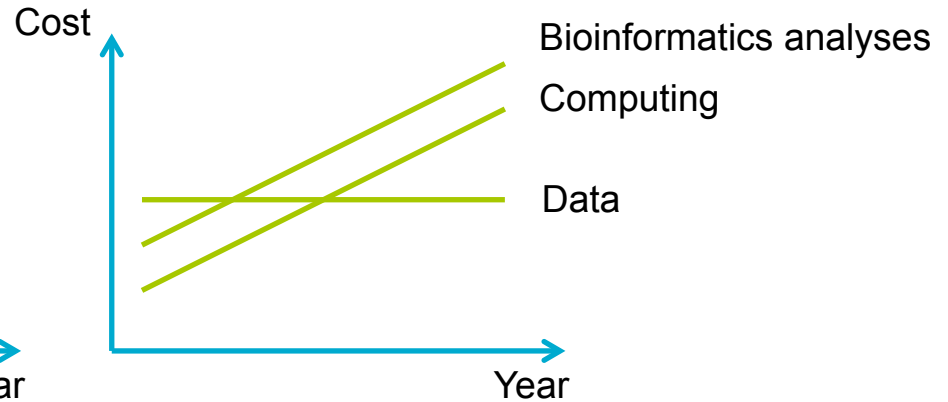
Our role

We want to help the Swedish Life Science community to **build knowledge** in large-scale data analysis, and to make bioinformatics **easily accessible** for all.

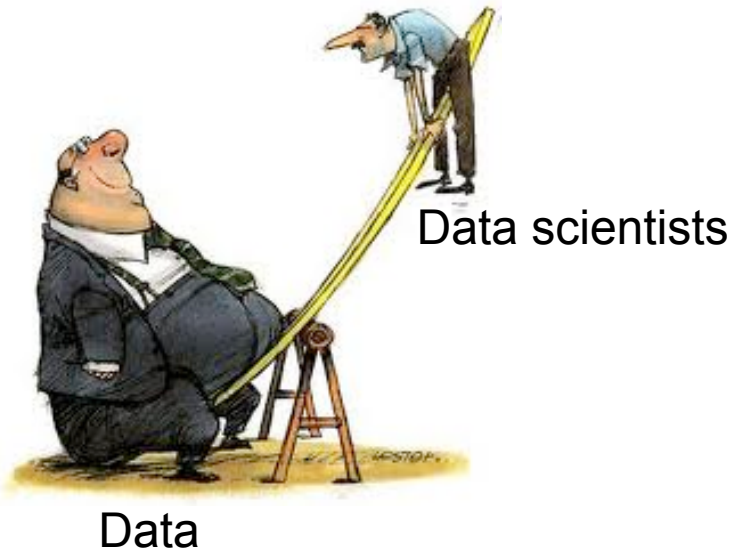
Production is cheap, analysis is not



“Per base”



“Per project”



Our role
 We want to help the Swedish Life Science community to **build knowledge** in large-scale data analysis, and to make bioinformatics **easily accessible** for all.

NBIS activities

Support, tools and training



Support



Tools



Training



4 facilities, ~60 FTEs

- **Support and Infrastructure**

Wide competence in bioinformatics, Assembly/Annotation, SysDev

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

Large collaborative projects selected by scientific ranking

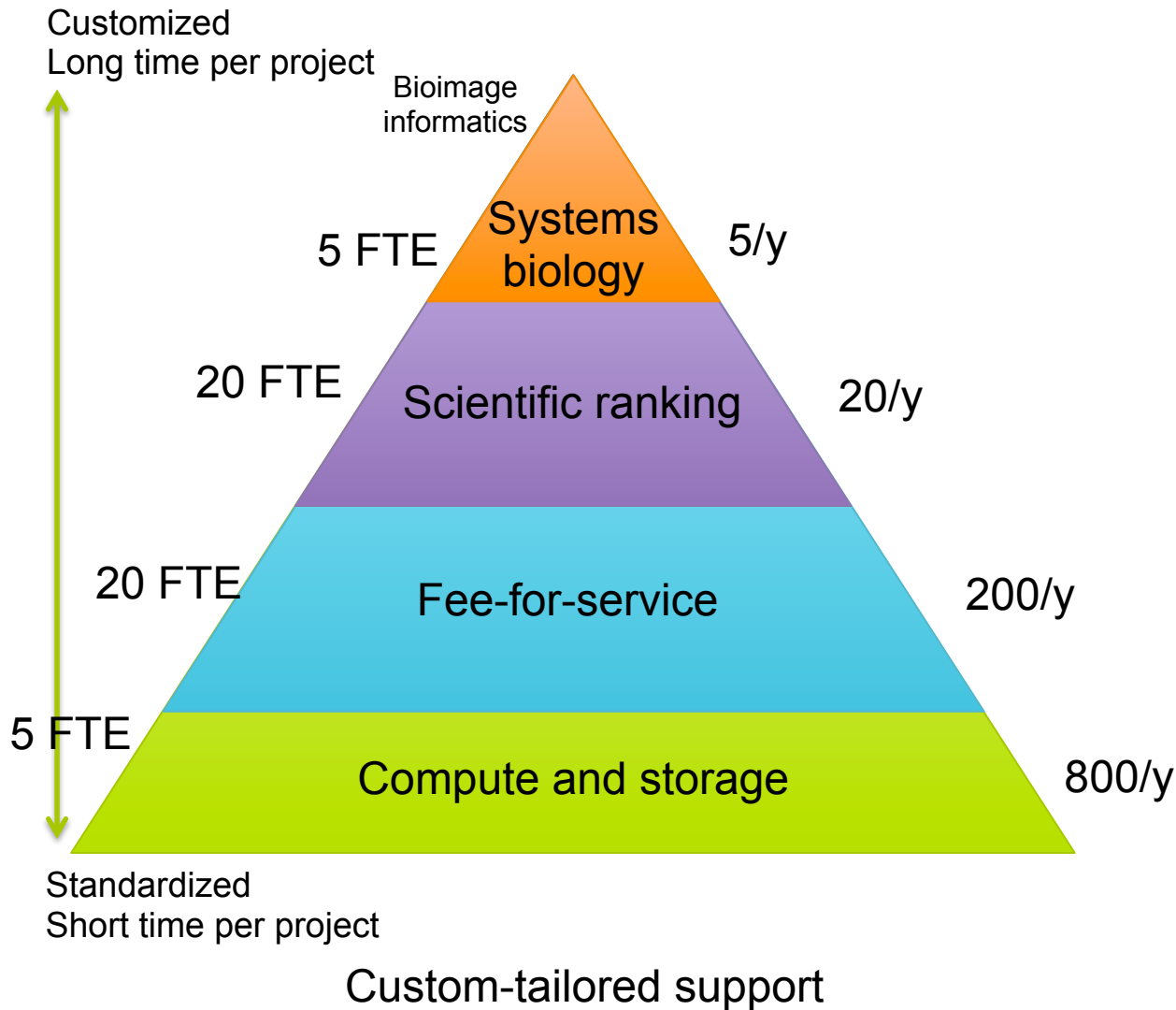
- **Systems biology**

Network analyses and Integrative bioinformatics

- **Compute and storage**

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





Training

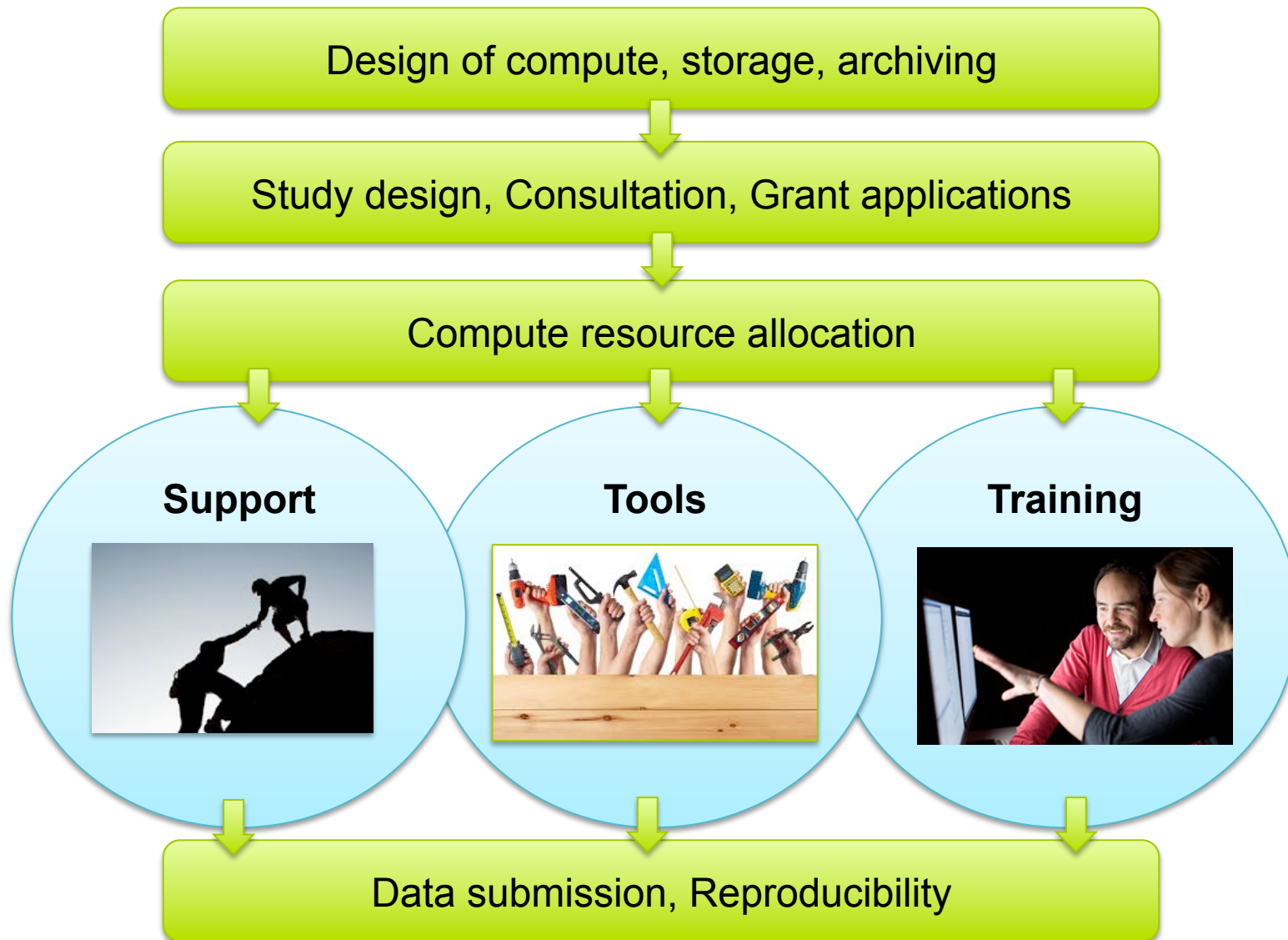


5 FTE
Systems development



1 FTE
Data management

User benefits

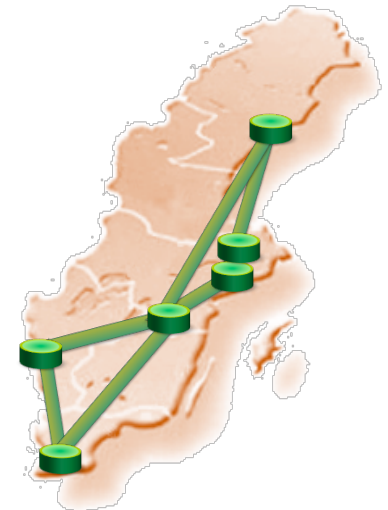




Custom-tailored support

www.scilifelab.se/platforms/bioinformatics/
www.nbis.se

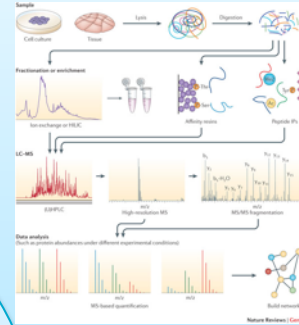
- Study design consultation (free)
www.nbis.se/support/supportform/index.php
+ drop-in sessions every week @ 6 sites
- Short- and Medium-term support (User fee 800 kr/h)
www.nbis.se/support/supportform/index.php
- Long-term support and systems biology
(500h, free, scientific evaluation)
www.nbis.se/support/supportform/index.php?form=longterm



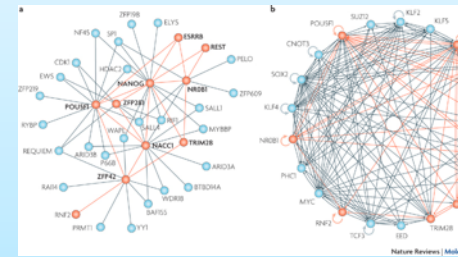
Bioinformatics support



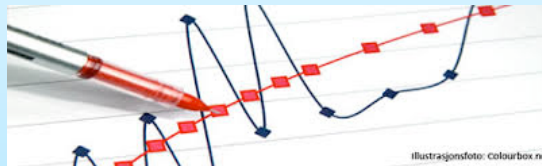
Genomics



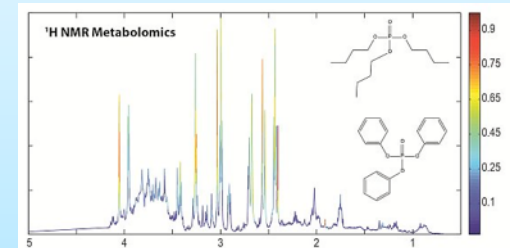
Proteomics



Systems biology



Biostatistics



Metabolomics

2 tracks!

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

The image shows a screenshot of a web browser 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 main content area features a large 'NBIS' logo in orange and green, with the text 'NATIONAL BIOINFORMATICS INFRASTRUCTURE SWEDEN' below it. A banner below the logo reads 'NBIS is a distributed national bioinformatics infrastructure, supporting life sciences in Sweden'. At the bottom of the page, 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 'Royex_gen2_brosc...pdf' and 'royex_cartridges_t...pdf'.

The screenshot shows a web browser window with the URL `nbis.se/support/support.html`. The page header includes the NBIS logo and a navigation menu with items: Support, Infrastructure, Training, News, and About. The SciLifeLab logo is also present in the top right of the page header.

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...

At the bottom of the browser window, there are two PDF files open in the taskbar: `Royex_gen2_brosch...pdf` and `royex_cartridges_t...pdf`. A "Visa alla" button is visible in the bottom right corner of the browser window.

CONSULTATION

SUPPORT

LONG-TERM SUPPORT

NBIS offers bioinformatics consultations as a free service.

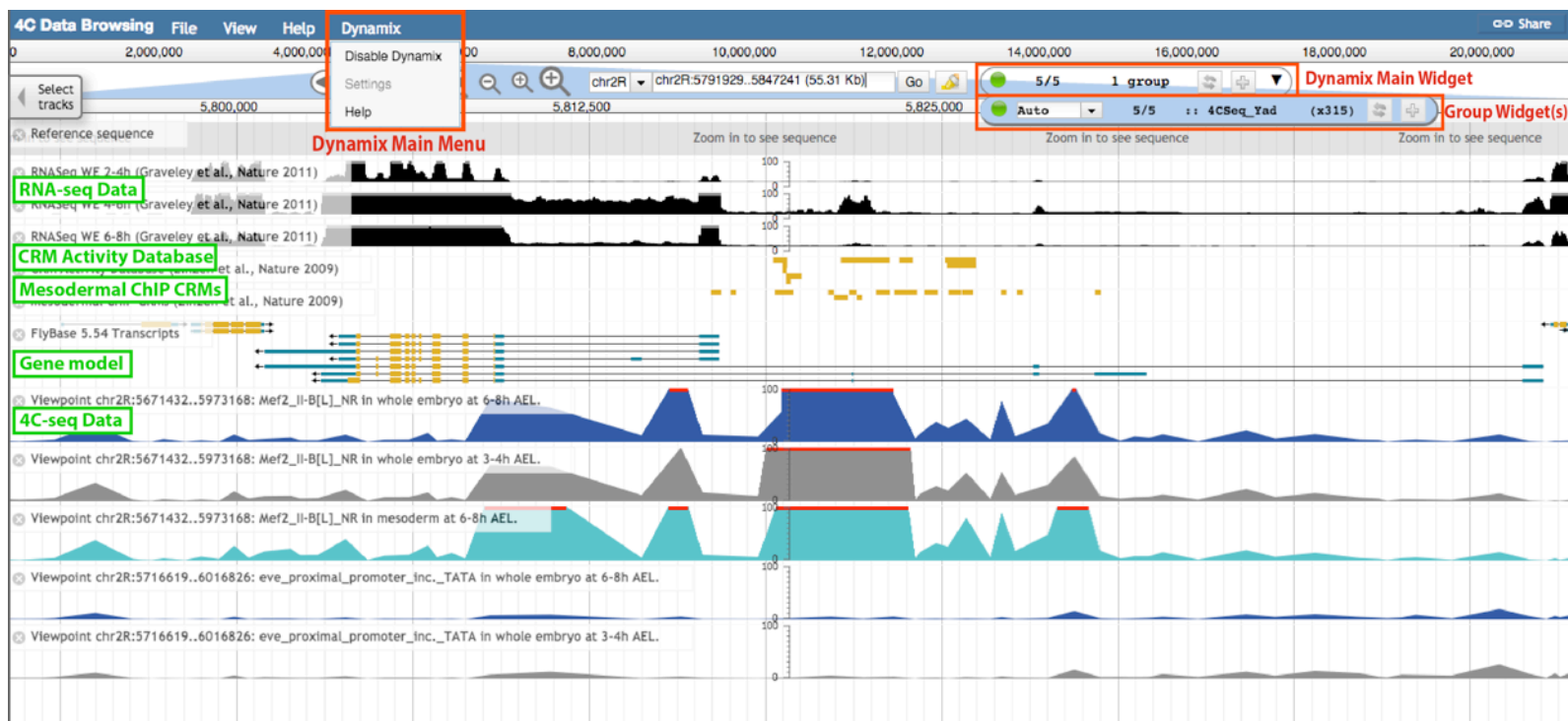
[Click here for further information about the NBIS consultation service.](#)

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

By applying for NBIS Support you also agree to our [user agreement](#).

NBIS Consultation 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	ALF ▾
Project Title	Please include descriptive keywords.
Subject	NGS ▾
Description	Brief project description.
Starting time	ASAP ▾



- 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

Cost effective with high quality!

BigData/Integrative omics

4 FTE, joint effort by Long-term Support and Systems Biology

Projects apply in the regular Long-term Support calls

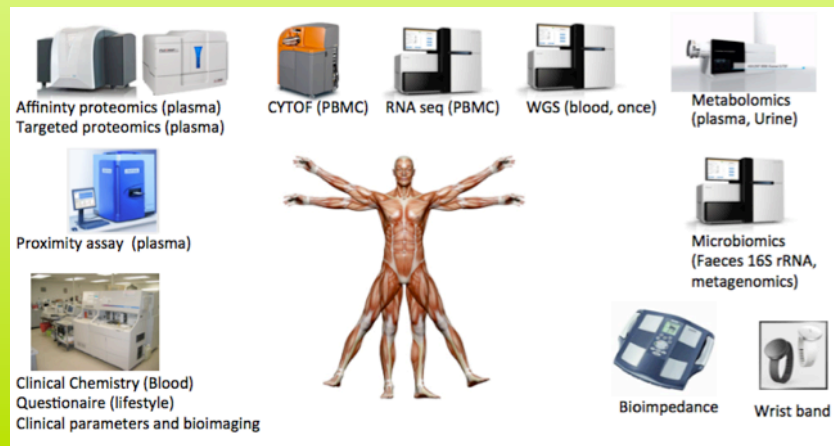
Combine data from SciLifeLab platforms

- 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

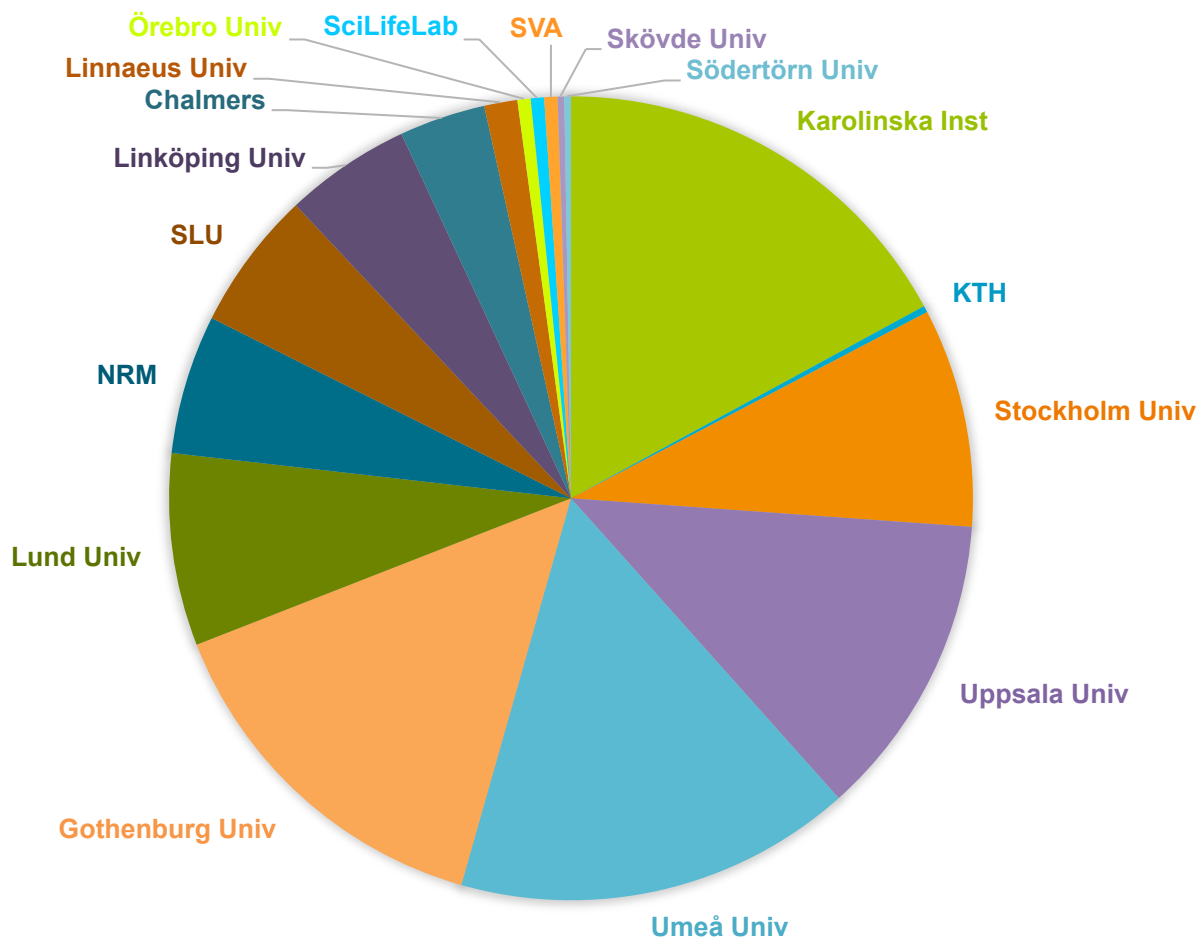
Support track for integrative projects

First call Feb 2016; First few projects initiated

Involves extensive integration of data



Geographical Distribution of Projects 2015





<https://docs.google.com/spreadsheets/d/1PrehKn2eb0ymfaFtCfvbLrOSKtpTL3qLcWZ2YwoXOIU/edit#gid=0>

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 Prioritization date: 2016-09-02

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>						<i>We're experimenting with different ways to create value with an opportunity, we've identified so we can build the right thing, the right way, at the right time.</i>					
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	Impute GV
2	SweFreg	x	Adam A	NJ	<1	2					Permutatio
3	ePouta integration pilot	x	Antti P	JH?	1	3					Developme
4	microMosler		NJ?	JH	1?	4					Improve bi
5	WGS structural variation WF		BN	PO	1	5					Single-cell
6	Falcon on milou		HL	JB	<1	6					Chip-seq p
7	Pouta Blueprints		OS	JH	2						CWL for NI
8	web-servers with EGI cloud vo.NBIS.se		MB	NS	2						OpenCB
9	Human Metabolic Atlas (HMA)		TS	LH	1-6						Local Ense
10	SSC HPC2N region dynamic resource allocation		MB	NS	2						NBIS Tools
11	Database for SciLife/AZ proj	x	BPn?	JV?	4?						IgDiscover

Ongoing dev projects not resourced by the dev team

	WGS somatic variant calling WF		BN	(SJ)	0
	Haloplex variant calling pipeline		PE	ML	0
	WhatsHap: Genomic phasing tool		PE	MM	0

Open prioritization and background descriptions

Tools and development projects needs to be much more visible!
Work in progress...

SweGen Variant Frequency Database

- 950 twin registry + 50 Northern Sweden
- Deep coverage WGS (30X)
- ExAC browser interface
- Data Beacon
- Full SNP frequency table download



<https://swefreq.nbis.se/#/>

1st release October 2016!

Variant: 22:46615880 T / C

Filter Status: PASS
dbSNP: rs1800234
Allele Frequency: 0.0035
Allele Count: 7 / 2000
UCSC: 22-46615880-T-C
ClinVar: [Click to search for variant in Clinvar](#)

Genotype Quality Metrics
Site Quality Metrics

Annotations

This variant falls on 7 transcripts in 1 genes:

missense
• PPARA Transcripts

intron
• PPARA - ENST00000434345

non coding transcript exon
• PPARA - ENST00000493286

Note: This list may not include additional transcripts in the same gene that the variant does not overlap.

Population Frequencies

Population	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
SweGen	7	2000	0	0.0035
Total	7	2000	0	0.0035

Read Data

This interactive IGV.js visualization shows reads that went into calling this variant.

Note: Read data is not available for this variant.

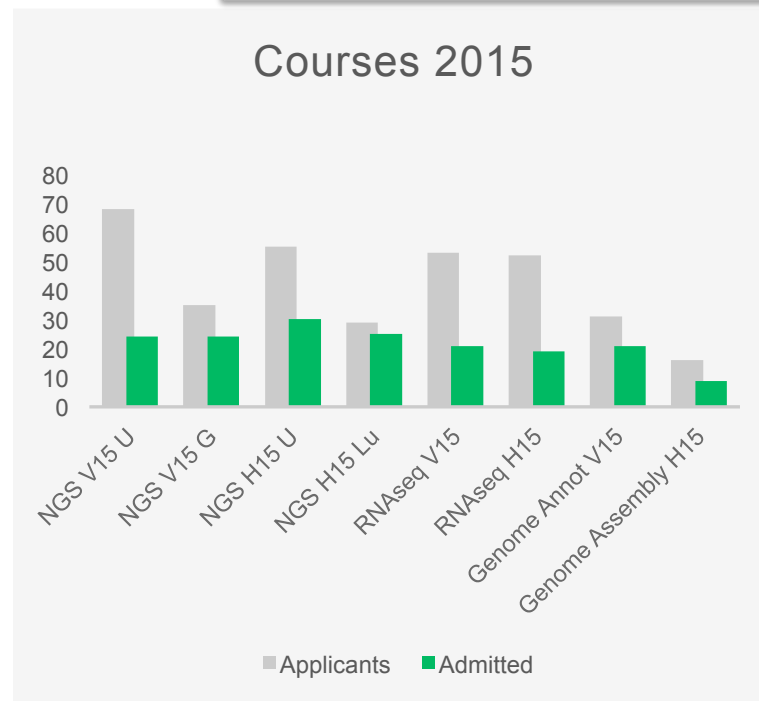
chr22:46,615,730-46,616,030 301 bp

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



- 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/
www.nbis.se/training/events.html

The Swedish Bioinformatics Advisory Program

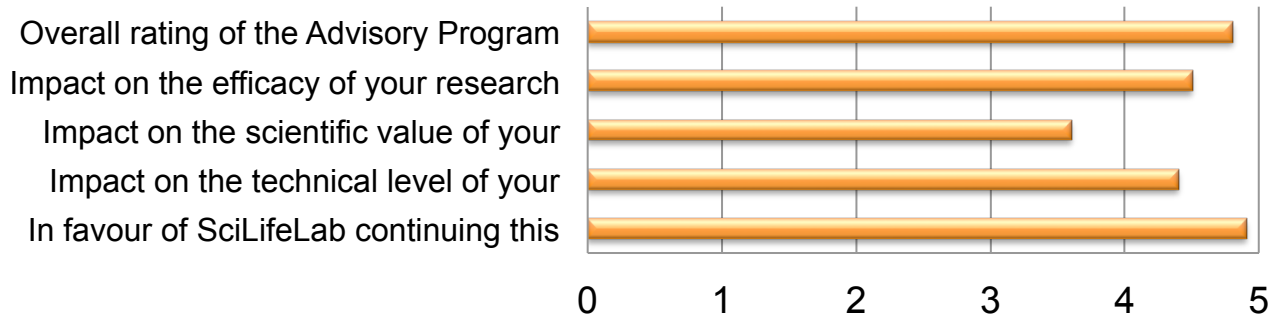
PhD students get a senior bioinformatician as a personal advisor during 2 years of their PhD. Monthly project meetings + two grand meetings per year to aid networking and knowledge transfer.

www.scilifelab.se/education/mentorship/the-swedish-bioinformatics-advisory-program/

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

The Swedish Bioinformatics Advisory Program

Student evaluation, June 2015



Elixir

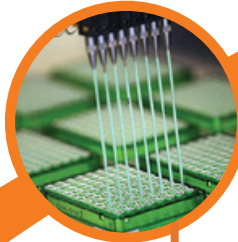
Why ELIXIR?

- 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



The screenshot shows a web browser window with the URL nbis.se. The browser's address bar and tabs are visible at the top. The website's main navigation menu includes links for Support, Infrastructure, Training, News, and About. A dropdown menu is open under 'Support', listing the following options: General information, Consultation, Support, Long-term Support, Genome Assembly and Annotation, and Apply here! Below the navigation, a large banner features the NBIS logo (orange 'NBIS' with a green and blue DNA double helix) and the text 'NATIONAL BIOINFORMATICS INFRASTRUCTURE SWEDEN'. A dark grey bar below the banner contains the text: 'NBIS is a distributed national bioinformatics infrastructure, supporting life sciences in Sweden'. At the bottom of the page, there are two green buttons with icons: one with a person icon and another with a gear icon. The browser's address bar shows the URL nbis.se/support. Two PDF files are open in the browser's tabs: 'Royex_gen2_brosc...pdf' and 'royex_cartridges_t...pdf'.

