Annual Report 2020

Introduction

NBIS (National Bioinformatics Infrastructure Sweden) is a distributed national research infrastructure, hosted by Uppsala University. NBIS constitutes the SciLifeLab Bioinformatics Platform. NBIS is formed by 11 partners: Uppsala universitet (hosting NBIS), Chalmers tekniska högskola AB, Göteborgs universitet, Karolinska Institutet, Kungl. Tekniska högskolan, Linköpings universitet, Lunds universitet, Naturhistoriska riksmuseet, Stockholms universitet, Sveriges lantbruksuniversitet, and Umeå universitet

Funding is provided from the Swedish Research Council (VR), Knut and Alice Wallenberg Foundation, Science for Life Laboratory (SciLifeLab), Swedish universities, and user fees. In addition, we have funding from NordForsk (NeIC), ELIXIR and EU.

The corona virus pandemics has forced NBIS to work much more remotely with many digital meetings, something that we have already used much in our activities since our staff is at geographically different places. Our drop-in sessions have been moved entirely to the digital world for most of 2020 and continuing into 2021. We also transformed the retreat to a digital retreat in order to still be able to discuss and brainstorm even though at physical distance. Looking back on 2020, we see no major drops in number of users or number of support projects, but some activities, e.g. our summer school, have been paused. Also the international work within ELIXIR has been predominantly run via digital meeting systems, including the kick-off of a new EU project.

Vision and Mission

Enable world-class life science research and maximise scientific and societal impact of collected data by:
- Providing expert knowledge, innovative data integration, advanced training, efficient data publication for open science, and access to high-performance data analysis methods
- Coordinating bioinformatics support within Sweden and making bioinformatics easily accessible for life science researchers
- Swiftly responding to changes in support needs as new techniques are developed and utilised
- Forming the Swedish ELIXIR node and participating in relevant international projects

Support

One of the major activities is support, where our staff helps researchers with bioinformatics tasks in various projects. Currently, NBIS provides expertise in many areas within bioinformatics: genome assembly, genome annotation, genetic variation, comparative genomics, phylogenomics, transcriptomics, proteomics, metabolomics, systems biology, single-cell biology, biostatistics, and multi-omics integration.

In the new funding landscape, the universities need to take a larger part of infrastructure funding, in proportion to the utilisation of NBIS by the respective university. This is reflected in the increased university contributions from 2018 and onwards. NBIS obtains user fees on direct project-related support, with the exception of the KAW-funded LTS support (cf. below). It is not feasible to apply user fees on the infrastructure, outreach and training activities, staff education and project management.

NBIS serves users at all major universities. In 2020, NBIS provided support to **267 PIs** (104 female, 163 male). The distribution of the supported PIs is shown in Figure 1. The top three universities using NBIS are Karolinska Institutet, Uppsala University and Lund University. The total number of active support projects were 212, and in addition to those NBIS has also provided 104 study design consultations.
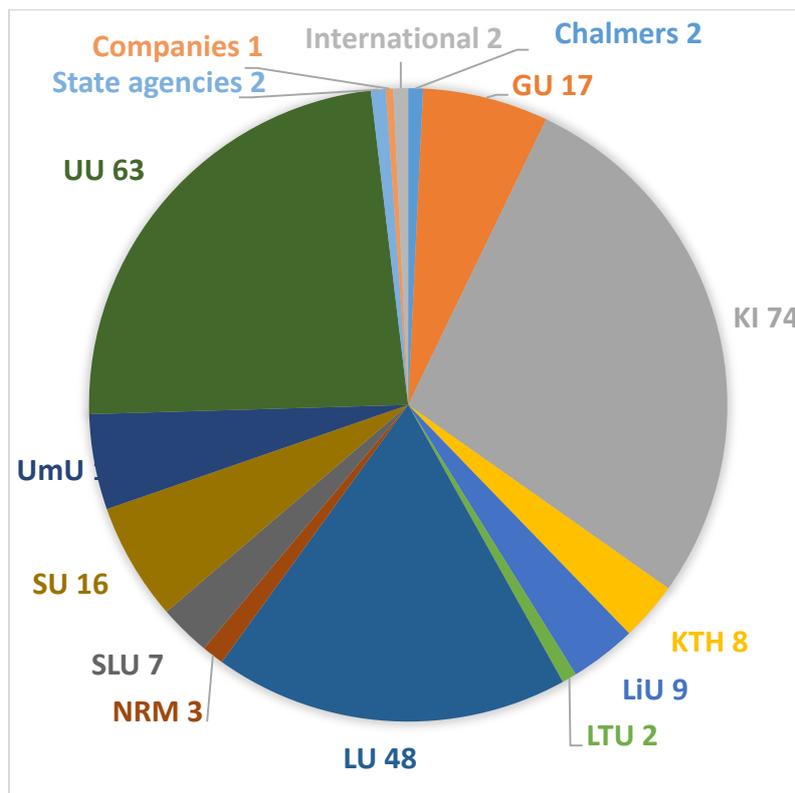

Figure 1 Distribution of PIs that have received support during 2020

In 2020, short- and medium-term support (SMS) was provided to 115 projects and long-term support (LTS) to 97 projects. A wide variety of projects were supported, with gene expression studies being the single biggest category. About half of all support is provided from one site to another, emphasising the importance of a national infrastructure.

Analysed data were from a range of sequencing-based methods, and also other large-scale data, such as proteomics, metabolomics and cell imaging. Projects including single-cell RNA sequencing (scRNA-seq) and human whole genome sequencing data remain in high demand of support, while emerging areas include ancient DNA (aDNA) and cell image analysis.

Examples of important achievements from NBIS support projects

NBIS has a total of 47 publications in 2020 from our direct project support work (excluding Compute & Storage), listed in the Annex. A few publications are shortly highlighted here.

Human whole-genome sequencing (WGS) is emerging as a powerful and affordable method in both research and diagnostics, and NBIS has been an active driver in this field in Sweden, including involvement in both resources such as the Swedish reference cohort SweGen and improved bioinformatics workflows used by many national and international research groups (Garcia, M. et al.: Sarek: A portable workflow for whole-genome sequencing analysis of germline and somatic variants. *F1000Research 9*:63 (2020); https://doi.org/10.12688/f1000research.16665.2).

NBIS has also been involved in many WGS research projects, e.g. to further elucidating the genetic architecture of schizophrenia (the single largest Swedish Human WGS effort published to date), concluding that ultra-rare non-coding structural variants at the boundaries of topologically associated domains (TADs) increase risk for schizophrenia (Halvorsen, M. et al.: Increased burden of ultra-rare structural variants localizing to boundaries of topologically associated domains in schizophrenia *Nat Commun 11*, 1842 (2020); https://doi.org/10.1038/s41467-020-15707-w).

An increasing number of studies approach NBIS for assistance to combine analyses of multiple types of omics data, and in a uniquely broad multi-omics study, NBIS helped analyse a longitudinal wellness cohort with 100 healthy individuals with respect to blood molecular profiles including proteomics, transcriptomics, lipidomics, metabolomics, autoantibodies and immune cell profiling, complemented with gut microbiota composition and routine clinical chemistry, emphasising the importance of an individual-based definition of health (Tebani, A. et al.: Integration of molecular profiles in a longitudinal wellness profiling cohort. *Nat Commun 11*, 4487 (2020); https://doi.org/10.1038/s41467-020-18148-7).

We have also applied Machine Learning to address complex predictive modelling of multi-omics data of clinical relevance (Fernandez, C. et al.: Plasma Lipidome and Prediction of Type 2 Diabetes in the Population-Based Malmö Diet and Cancer Cohort. *Diabetes Care 43*, 366–373 (2020); https://doi.org/10.2337/dc19-1199).

Many researchers turn to NBIS for support in RNA-seq based gene expression studies. Radiotherapy, used in cancer treatment, is known to cause side-effects in forms of long-lasting cognitive impairments. Zanni et al. examined the effect of lithium treatment on the neural stem and progenitor cells after brain irradiation. The study showed that lithium treatment can reverse irradiation-induced loss of cognitive impairment, and more importantly, suggested a model of treatment. NBIS contributed by analysing and interpreting the gene expression data (Zanni et al.: Lithium treatment reverses irradiation-induced changes in rodent neural progenitors and rescues cognition. *Molecular Psychiatry 26,* 322–340 (2021); https://www.nature.com/articles/s41380-019-0584-0).

The ability of NBIS to rapidly develop expertise and provide support for new data types has enabled several research groups to address new questions using emerging technologies such as single-cell RNA sequencing (scRNA-seq) and Spatial Transcriptomics (ST). For example, a recent scRNA-study revealed tissue-specific transcriptional profiles and heterogeneity of innate lymphoid cell across four human tissues of importance for immune response in humans (Mazzurana, L. et al.: Tissue-specific transcriptional imprinting and heterogeneity in human innate lymphoid cells revealed by full-length single-cell RNA-sequencing. *Cell Res.* (2021) Epub ahead of print; https://doi.org/10.1038/s41422-020-00445-x).

NBIS also comprise strong competence in ecology and evolution, exemplified e.g. in involvement in studies of methylation patterns in chicken domestication (Höglund, A. et al.: The methylation landscape and its role in domestication and gene regulation in the chicken. *Nature Ecology & Evolution 4,* 1713–1724 (2020); https://www.nature.com/articles/s41559-020-01310-1), and ancient-DNA demographic studies of the extinct Woolly Rhinoceros (Lord, E. et al.: Pre-extinction Demographic Stability and Genomic Signatures of Adaptation in the Woolly Rhinoceros. *Current Biology 30*, 3871–3879 (2020); https://doi.org/10.1016/j.cub.2020.07.046).

Reconstruction of the birth of a male sex chromosome present in Atlantic herring. Sex chromosome evolution is known as the queen of problems in evolutionary biology. It has a crucial role in sex determination and there are diverse systems discovered in many different species. An international team led by researchers from Uppsala university for the first-time discovered the sex determination system in the Atlantic herring. The male-specific region is a small segment containing only three genes which was validated by a multi-omics approach (genomics, transcriptomics, proteomics, and functional validation). NBIS was involved in the annotation of the identified region (Rafati, N. et al.: Reconstruction of the birth of a male sex chromosome present in Atlantic herring. *PNAS 117*, 24359–24368; https://doi.org/10.1073/pnas.2009925117)

Dysregulation in Akt/mTOR/HIF-1 signaling identified by proteo-transcriptomics of SARS-CoV-2 infected cells. This publication sheds a light on how SARS-CoV-2 engages cellular host pathways and innate immunity in infected cells. Four pathways, ErbB, HIF-1, mTOR and TNF signaling were markedly modulated during the course of the SARS-CoV-2 infection.
NBIS assisted with the creation of the transcriptomic and proteomic networks, performed their biological characterization at community level, and their integration. The approach used made it possible to recapitulate common biological functions in the two networks, despite widely different coverages of RNAseq and proteomic data. (Appelberg, S. et al.: Dysregulation in Akt/mTOR/HIF-1 signaling identified by proteo-transcriptomics of SARS-CoV-2 infected cells. *Emerging Microbes & Infections 9*, 1748–1760; https://doi.org/10.1080/22221751.2020.1799723).

Translating GWAS-identified loci for cardiac rhythm and rate using an in vivo image- and CRISPR/Cas9-based approach. Image- and CRISPR/Cas9-data from zebrafish were used to characterize genes involved in Heart Rate Variability (HRV). A novel gene involved in HRV was identified, and the roles of already known genes was confirmed. The results can be used to identify novel drug targets to prevent cardiac death. NBIS staff designed and implemented a pipeline for quality control of sequence data and variant calling (von der Heyde, B. et al.: Translating GWAS-identified loci for cardiac rhythm and rate using an in vivo image- and CRISPR/Cas9-based approach. *Sci. Rep. 10,* 11831; https://www.nature.com/articles/s41598-020-68567-1).

User fees

NBIS has four support tracks:
- Study design consultation (≤ 3h)
- Short- and Medium-term Support (SMS)
- Partner Projects (PP)
- Long-term Support (LTS)

For SMS and PP, NBIS charges user fees, while LTS is provided for free according to the funding requirements by the Knut and Alice Wallenberg foundation. In recent years, the total income from user fees has increased and in 2020, the amount was 7.5 MSEK, indicating better awareness of costs for data management and advanced bioinformatics support.

Infrastructure

A fundamental part of NBIS is the formation of a sustainable bioinformatics infrastructure for life sciences, consisting of access to tools and data. The infrastructure is typically constructed as domain-specific supporting layers utilising resources from the national e-infrastructure providers SNIC and SUNET. Our computational and storage needs are set up in close collaboration with these infrastructures, in order to avoid duplication of efforts and to benefit from the competences within SNIC and SUNET.

Data management

During 2020, NBIS has expanded its data management efforts by hiring three new data stewards to form a data management team currently consisting of four data stewards and one data manager. A clearer Vision & Mission for the data management activities at NBIS has been established.

The overarching vision is that Swedish life science researchers apply good data management practices so that the research outputs produced are available to the global research community, and to society at large, according to the principles of Open Science, Reproducible Research, and FAIR (Findable, Accessible, Interoperable, Reusable).

The mission of the data management team to work towards this vision is to ensure that Swedish life science researchers have access to data management **support** and **training** to be able to manage their research data according to international best practice, and that these data management practices, solutions and training for Swedish life science are established in **collaboration** with relevant national and international stakeholders. The primary collaborator in this regard is the SciLifeLab Data Centre, with which the collaboration has been intensified during this year as new initiatives have been initiated, particularly in response to the COVID-19 pandemic. The Data management team has contributed to the establishment of the national Swedish COVID-19 data portal set up by the Data Centre. In relation to this, joint work has also been initiated to establish support routines for submitting COVID-19 related data to core deposition databases, which also contributes to the NeIC PaRI project.

NBIS has previously, in close collaboration with the SciLifeLab Data Centre, established an online resource for users to create Data Management Plans (DMP). This tool, called the Data Stewardship Wizard (DSW), originating from the ELIXIR collaboration, is provided as a service to the

community. In the DSW, we have continued to improve and deepen the life science-focused DMP template model that adheres to the national guidelines. The plans to arrange Bring Your Own DMP consultation workshops unfortunately had to be cancelled due to the COVID-19 pandemic.

A key factor in good management of research data is to make generated datasets discoverable and accessible to society, and the research community in particular. This is done by submitting the datasets to trusted international public repositories, especially the core Deposition Databases identified by ELIXIR. We have continued supporting users in making such submissions. As the FAIRification of research data is one of the main objectives of the NBIS data management team, one focus activity has been to work out how this can be achieved for the projects that engage the NBIS bioinformatics analysis support. Working together with the NBIS bioinformaticians, the data management team has been developing procedures aiming at promoting good data management practice in NBIS support projects, e.g. that datasets in these projects are published in suitable international repositories whenever possible.

The ethics and laws around personal data presents challenges when it comes to making datasets containing e.g. human genetic data available to the research community. The strategy taken is to provide a solution for this under the umbrella of the European Genome-phenome Archive (EGA), and in collaboration with other ELIXIR nodes to establish a Federated EGA landscape. The NBIS systems development team has made considerable progress in establishing an operational Swedish EGA node, that should become available as a service for archiving human genetic and phenotypic research data for the Swedish biomedical research community in the near future. The Data management team has during 2020 initiated work to establish a user-facing submission support process, to facilitate high quality data and metadata data depositions, aiming at high FAIRness of the data deposited.

Data Management training efforts have been focused on providing guidelines documentation, and on laying the groundwork for an entry level course on data management practices. We have extended and improved on the content of the data management information resource *SciLifeLab Data Guidelines* (https://scilifelab-data-guidelines.readthedocs.io/en/latest/docs/index.html). Of note is the addition of a section on covid-19 related data, based on available international recommendations and standards. There is currently a lack of data management competence and know-how in the Swedish research community. To meet this, we during 2020 constructed a practical course *Introduction to Data Management Practices*, targeted at researchers to take their first steps towards good and reproducible data management practices. Due to the COVID-19 pandemic, the course had to be postponed to spring 2021.

Much effort has been spent on taking on a leadership role in the pan-European ELIXIR-CONVERGE project focused on data management. ELIXIR-CONVERGE aims to connect and align ELIXIR Nodes to deliver sustainable FAIR life-science data management services. NBIS as the Swedish ELIXIR node has led the work in work package *WP1 Data Management Expert Network*. A competence network for data management issues has been established across all 23 participating national nodes, with focus on knowledge sharing and capacity building. During 2020, the network has produced the first version of *Data Management best practices guidelines for the Life Sciences*, which is available at the ELIXIR RDMKit web resource (https://rdmkit.elixir-europe.org). The data Management team and other NBIS staff have also actively contributed to activities in other work packages *WP2 Training*, *WP3 RDM Toolkit*, *WP5 Demonstrators*, and *WP7 Federated EGA*.

Compute & Storage

High throughput biomedical science depends on high-performance computers for bioinformatics analysis. The hardware is maintained by SNIC but NBIS provides expertise needed to efficiently access the computational and storage resources. Notably, well over 800 bioinformatics-related software packages are installed and maintained on a general-purpose SNIC compute cluster and the SNIC compute cluster for sensitive data, Bianca. Compute & Storage also administers resource allocations and participate in help-desk support at the major national compute clusters, handling 1927 omics research projects with 895 unique PIs in 2020. These projects have used a total of close to 5.5 million core hours each month and ~9.5 PB of storage.

In collaboration with UPPMAX, NBIS has contributed to a decision by SNIC to procure GPU resources on the sensitive data system, Bianca, in 2021. An increased emphasis on data management as it relates to SNIC resources has resulted in a new process for storage allocation decisions, as well as contributions to a new SNIC storage policy and data removal guidelines. Additionally, a significant effort is spent on supporting users in navigating the judicial issues related to sensitive data processing, e.g. performing Data Protection Impact Assessments and writing Personal Data Assistant Agreements.

Systems Development & Tools

Our systems development team provides support in deploying tools so that they can be used by the entire life science community, and not just by bioinformaticians. We also facilitate large scale analyses by the development of stable and user-friendly pipelines. NBIS works strategically with Reproducible Research, actively promoting and deploying stable workflow systems and reproducible compute environments, both for internal tools development and for supported research projects.

The combination of bioinformatics experts, who can appraise the scientific value and usefulness of tools, with developers who have the capability of making tools stable and accessible, allows for development projects driven by user needs. The development team also assists in internal development projects, *e.g.* for integrating data services with other initiatives. The team is also active in the global alliance for genomic and health (GA4GH), working with technical standards and frameworks.

The system development teams also serve NBIS infrastructure internally. During the year, this included an extension of the genome annotation cluster. There we provided expertise during the hardware procurement, set up hardware and configuration management systems, and our developers created environments for the pipelines used in data analysis.

NBIS maintains a software repository for internal purposes and also public repositories for programs and scripts that are developed by our staff in order to get them to spread to the scientific community. For the latter, we have an account on GitHub. We enter relevant Swedish tools in the ELIXIR Tools Registry.

NBIS has worked on several projects during the year, and the largest project is the ELIXIR- and Tryggve-funded Federated EGA. In this project, several of our developers work together with developers from other ELIXIR nodes.

NBIS is contributing to the "Human Metabolic Atlas" that is used for management of genome-scale metabolic models, browsing and visualising metabolic networks in the context of quantitative data from the Human Protein Atlas project.

NBIS is also developing a range of pipelines for other data types. For example, our pipelines for single-cell RNA-seq and metagenomics are in advanced stages of development, openly available and already applied in multiple projects (see http://www.nbis.se/infrastructure/tools/ and http://www.nbis.se/support/ltsprojects.html).

Examples of important infrastructure achievements during 2020

Federated EGA
NBIS has in our capacity as the Swedish ELIXIR node worked in the EU project ELIXIR-EXCELERATE and the Nordic NeIC project Tryggve2 to establish a system for handling sensitive human data across the ELIXIR nodes: the federated EGA (European Genome-phenome Archive).

The federated EGA is instrumental to enable the **European 1+ Million Genomes initiative (1+MG)** [1]. In April 2018, a joint declaration of cooperation regarding giving access to at least 1 million sequenced genomes in the European Union by 2022 was signed by Sweden and 12 other member states; now a total of 24 member states have signed. To fulfil this vision in the short time-span outlined, it is necessary to build on existing research infrastructures. This involves transfer of existing and emerging technology and know-how between the research and diagnostic areas. Prime examples of technology to support this European effort are the Beacon and federated EGA activities that NBIS is already engaged in.

The federated EGA enables a sustainable, secure and legally correct storage of sensitive human genetic data. This **facilitates data publication and open science** (but with controlled access since the data are sensitive) according to the FAIR principles, which in turn enables data sharing of benefit for scientists, healthcare providers and industry.

During 2020, NBIS has provided knowledge in systems development in order to create important building blocks for data submission at the local EGA nodes. The EGA-SE node is technically ready and legal agreements are negotiated between the universities. In collaboration with our Nordic partners we have added functionality to run a standalone sensitive data archive.

Encima web app
One ongoing support project in the system development group is the creation of the web app encima: The Encyclopedia of Cancer Immune Microenvironment. The PIs of this project are Artur Mezheyeuski and Patrick Micke at IGP in Uppsala University. The web app enables users to explore their data base of histological and multiplex fluorescence pictures from cancer patients by comparing user-selected subsets of the data based on features such as cancer stage and tumour type and to see Kaplan-Meier survival plots. A prototype of the page is up and running on https://encima.one. A proper release is planned to the beginning of 2021.

Web-based herd management system
During 2020, the NBIS development team has supported the non-profit organisation "Föreningen Gotlandskaninen" in developing a web-based herd management system. The project is funded by the Swedish Board of Agriculture, as part of Sweden's "Action Plan for Animal Genetic

---

[1] https://ec.europa.eu/digital-single-market/en/european-1-million-genomes-initiative

Resources". The system allows gene bank managers to keep track of breeding and genealogy for their entire breeds, and herd owners to easily maintain their herds.

This system will be a huge boon for the 455 members and 299 herds of the breeding organisation, but also mark a milestone for NBIS. As part of NBIS's strategic goal to provide support to Swedish research, finding ways to support non-academic groups is a serious challenge. The system is built so it easily can adapt to fit the needs of some of the 20 other organisations helping out in the goals of conserving the genetic variation of about 70 different breeds of domesticated animals. The system is open-source, licensed under the GNU GPL3.

Training

NBIS staff is involved in many training activities, both in our own courses and as invited teachers/speakers in courses arranged by others. During 2020, we, as many other training providers, needed to speed up the shift from onsite to online training. NBIS successfully shifted all our courses to be delivered online. This will continue in 2021 and also be a part of our course programme beyond as we believe by doing online training we make the training more inclusive for the community. For the 2020 courses, we had over 400 participants in our advanced courses, tailored to graduate students and post-docs; the courses are listed below. This is a slight decrease compared to previous years. However, the majority of NBIS educational efforts were delivered during fall as courses were cancelled and/or postponed during spring due to the pandemic outburst in February/March of 2020. NBIS successfully has built a collaborative effort internationally with the Swiss Institute of Bioinformatics (SIB) Training group. and for the years 2019/2020 delivered a highly appreciated seasonal school in the topic of single cell omics analysis, which is one of the research fields where we see a high demand for training.

In addition to our national and international training efforts, NBIS provides individual training of researchers in our project-related support in order to teach them new bioinformatics tools and to help them utilise bioinformatics tools more efficiently. The training activities are also an efficient way to increase the flow of projects by helping scientists to be able to perform parts of the bioinformatics analyses themselves. NBIS runs a very much appreciated PhD advisory mentor programme in bioinformatics, and we are supporting the VR-funded research school in bioinformatics with training activities.

Course	From date	To date	Place
Courses organised by NBIS			
Single cell RNA sequencing data analysis	27.01.2020	30.01.2020	Stockholm
Introduction to Bioinformatics using NGS data	27.01.2020	31.01.2020	Uppsala
Omics integration and Systems Biology	05.10.2020	09.10.2020	Lund
Snakemake BYOC (bring-your-own-code) workshop	05.10.2020	07.10.2020	Online
Python programming with applications to bioinformatics (NBIS/ELIXIR-SE)	19.10.2020	23.10.2020	Online
Python programming with applications to bioinformatics (NBIS/ELIXIR-SE)	19.10.2020	23.10.2020	Umeå
R foundations for Life Science	02.11.2020	06.11.2020	Online
Introduction to Bioinformatics using NGS data	09.11.2020	13.11.2020	Online

Introduction to biostatistics and machine learning	15.11.2020	19.11.2020	Online
Tools for Reproducible Research	23.11.2020	27.11.2020	Online
Epigenomics Data Analysis: from bulk to single cell	23.11.2020	27.11.2020	Online
RNA-seq data analysis	30.11.2020	04.12.2020	Online
Courses co-organised by NBIS			
Workshop on Genomics	06.01.2020	19.01.2020	Czech Republic
NBIS/SIB Seasonal school: single cell omics analysis	23.08.2020	28.08.2020	Switzerland

In order to encourage and facilitate reproducible research, NBIS apply methods for reproducibility within an increasing number of projects. To share our lessons from this endeavour, we have created a course entitled "Tools for Reproducible Research", which is offered twice since 2019 (during 2020, 2021 using an online format). We have also made reproducibility a recurring theme in our other data analysis and computing courses. All courses in the NBIS catalogue are aimed to be a reoccurring in their delivery.

A new collaborative training effort started during 2019 between SIB and NBIS and this was continued in 2020 and beyond. We see an international demand and gap with regards to single-cell omics training and knowledge transfer. Experts from our two organisations collaborated to design and set-up the single-call omics analysis course delivered to an international audience. The joint SIB/NBIS autumn school took place in Leyson, Switzerland, September 2020.

Outreach

Outreach activities have proven important to inform the scientific community about the support that NBIS can provide, to increase collaborations and to increase the number of users. These consist of an annual symposium and user meeting, which was canceled due to COVID-19 this year; regular presentations at different universities, providing the possibility to meet staff representing our wide variety of competences; and weekly Bioinformatics Drop-in sessions enabling face-to-face contacts between researchers and our experts, which many times is the first contact in a support case. Due to COVID-19, the Drop-in sessions were moved online and occurred simultaneously across all sites. This not only enabled increased direct contact between users to all NBIS bioinformaticians, but also between NBIS bioinformaticians. This is a format we will keep, in addition to local drop-ins resuming post-COVID. During 2020, we arranged approximately 60 local drop-in sessions at the six sites Uppsala, Stockholm, Göteborg, Linköping, Lund and Umeå, and 18 site-wide zoom drop-ins starting from week 33.

In 2020 NBIS strategically assigned a community coordinator to oversee and plan NBIS outreach activities and two local site coordinators. The work to become more engaged locally nationally started immediately with a joint effort between NBIS and NGI to reach out and let life scientists know that we can help strengthen their grant applications using our bioinformatics, sequencing and data management expertise during VR grant season. NBIS also participates in local bioinformatics networks, such as GOTBIN in Gothenburg.

In the SciLifeLab Talkshow, an intra-SciLifeLab outreach and community effort, NBIS was involved both in planning and represented in the content.

During 2020, SciLifeLab BiG Talks! continued with an invited international seminar holder; "Petabase-scale sequence alignment catalyzes discovery of novel Coronaviruses" by Artem Babaian from University of British Columbia in Canada in fall 2020. BiG Talks! is a collaborative effort between NBIS and other SciLifeLab platforms to increase the networking possibilities for staff-scientists and affiliated researchers. The seminars are given by internationally renowned speakers within the fields of Bioinformatics and Genomics. The seminars are live broadcasted via the SciLifeLab Youtube channel in order to enable all SciLifeLab sites to participate.

Collaboration with industry

NBIS continues our efforts to increase contacts with life science companies, which can utilise our resources at a full-cost basis, enabling them access to tools, expertise, and training.

Contributions by the Swedish ELIXIR node (ELIXIR-SE)

- ELIXIR-SE has continued maintaining and updating the Human Protein Atlas – which now is officially named as an ELIXIR Core Data Resource – in the ELIXIR landscape.
- ELIXIR-SE has been very active in the systems development for Federated EGA in collaboration with other Nordic ELIXIR nodes, ELIXIR-Spain and ELIXIR-EBI.
- Two new EU-funded projects started 2020 – ELIXIR-CONVERGE for data management and B1MG for the European 1+ Million Genome project.

Staff

The table shows the number of staff in FTE (full time equivalents) during 2020 for the different functions in NBIS. Gender balance is 73% male and 27% female.

NBIS	Staff (FTE)
1 Central functions, incl. management	9.2
2A Support (SMS)	16.8
2B Support (LTS)	14.4
2C Support (Sysbio)	1.7
3 Data management	2.8
4 Systems development	9.3
5 Compute & Storage	4.7
6 Training	8.8
7 ELIXIR	4.8
Other	0.3
TOTAL	**72.7**

Economy report for 2020

NBIS	2020	Result	Budget
Incomes	VR	21 000 000	21 000 000
	SciLifeLab National	18 700 000	20 000 000
	Universities + SciLifeLab SFO	14 070 000	14 300 000
	KAW	20 797 306	19 399 000
	NordForsk (Tryggve)	1 787 059	2 000 000
	User fees	7 564 581	7 000 000
	SUM Incomes	**83 918 946**	**83 699 000**
Expenses	Personnel	60 151 777	62 115 600
	Equipment	1 930 163	
	Travel	410 084	
	Other costs	893 519	
	Office space	4 492 814	21 824 400
	Indirect costs	12 776 605	
	SUM Expenses	**80 654 962**	**83 940 000**
	Surplus forwarded to 2021	3 263 984	

The surplus from 2020 is due to late received co-funding, parental leaves and delayed recruitments; the surplus will be used for additional staff recruited early 2021.

**Distribution of expenses on the different areas**
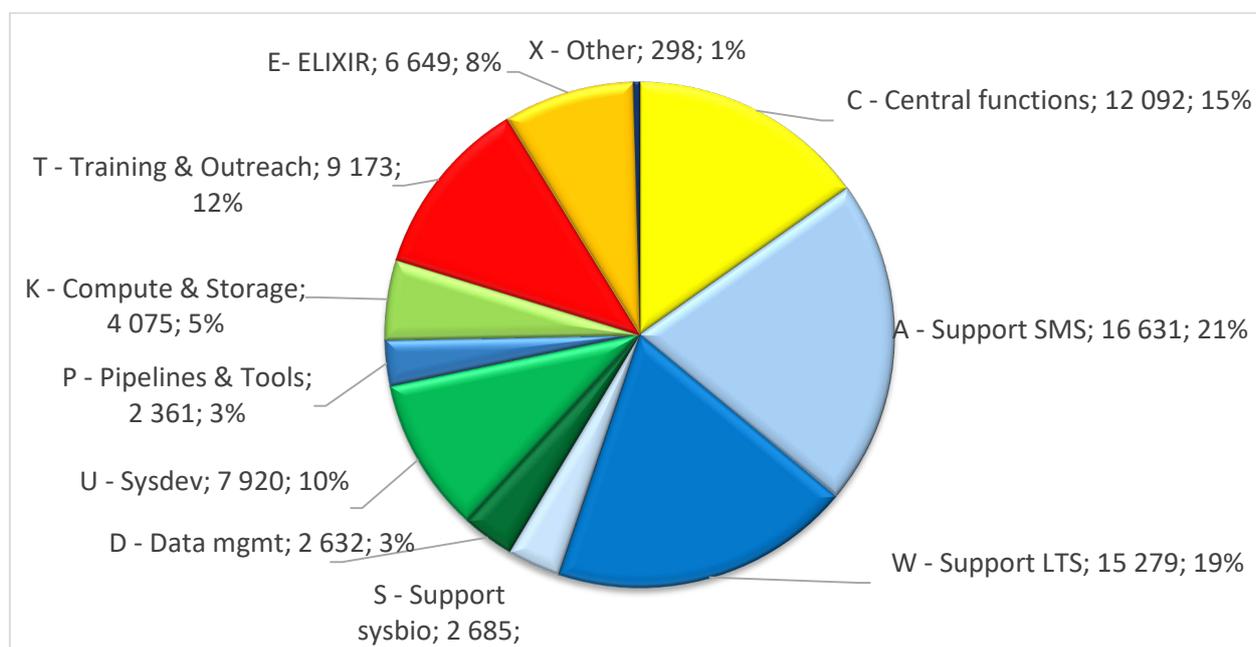
Annex – Key performance indicators

1. Number of projects

During 2020, we have worked on 212 support projects and provided 104 consultations.
The 212 support projects involve 181 unique PIs (68 female, 113 male).

2. Number of PIs distributed on universities

Univ	# PI
Chalmers	1
GU	6
KI	55
KTH	4
LiU	6
LU	32
NRM	2
SLU	3
SU	13
UmU	8
UU	48
Other	3
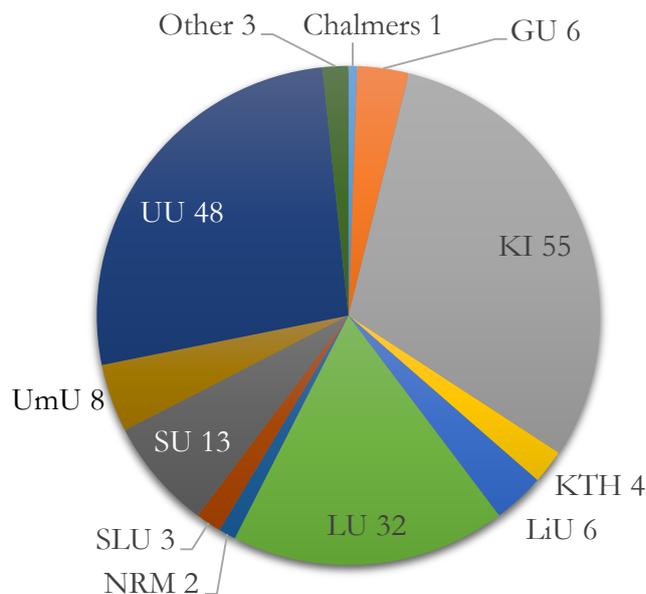
3. Number of projects distributed on SCB codes

SCB code and subject	# projects
102 Computer and Information Sciences	4
105 Earth and Related Environmental Sciences	1
106 Biological Sciences	78
209 Industrial Biotechnology	1
301 Basic Medicine	67
302 Clinical Medicine	43
303 Health Sciences	5
304 Medical Biotechnology	6
305 Other Medical and Health Sciences	2
401 Agricultural, Forestry and Fisheries	4
501 Psychology	1

4. Gender balance

Total 267 PIs that have received support, of which 104 female (39%) and 163 male (61%).

5. Publications

A total of 48 publications in 2020 from our direct project support work (excluding Compute & Storage); DOIs provided in the table below.

10.1093/jn/nxaa391	10.1111/1365-2745.13550	10.1111/1758-2229.12880
10.1038/s41467-020-15707-w	10.1038/s41598-020-76137-8	10.1111/jsr.13041
10.1038/s42003-020-1028-8	10.1101/2020.11.03.20225524	10.1016/j.cub.2020.07.046
10.1038/s41598-020-58025-3	10.1038/s41598-020-74798-z	10.1073/pnas.1920469117
10.1038/s41467-019-14201-2	10.1111/1755-0998.13252	10.1371/journal.pcbi.1007976
10.1080/22221751.2020.1799723	10.1530/eje-19-0522	10.1080/15592294.2020.1788325
10.1038/s41598-020-68567-1	10.1371/journal.pone.0239495	10.1002/1878-0261.12663
10.1038/s41467-020-16225-5	10.1038/s41559-020-01310-1	10.1073/pnas.1919535117
10.1186/s43008-020-00045-9	10.1002/dvdy.253	10.31235/osf.io/7qeht
10.1038/s41525-020-00152-x	10.1073/pnas.2009925117	10.1080/15592294.2020.1741758
10.1186/s12915-020-00808-1	10.1101/2020.09.16.299966	10.1126/scisignal.aaz1482
10.1186/s13059-020-02035-x	10.1038/s41467-020-18148-7	10.1073/pnas.1918232117
10.3390/jcm9123877	10.1002/ijc.33270	10.12688/f1000research.16665.2
10.1055/s-0040-1720980	10.1038/s42003-020-01214-7	10.1038/s41467-019-13869-w
10.1080/17453674.2020.1846956	10.1126/sciadv.aba8196	10.1186/s13148-019-0803-1
10.1128/msystems.01006-20	10.3389/fmicb.2020.02066	10.1126/science.aay5947

6. Number of users that have applied for access but not being prioritised

Here we show numbers from the WABI part of NBIS, since other users have the option of paying user fees as long as their project is technically feasible and NBIS has capacity. In 2020, a total of 74 applicants of which 20 were granted LTS support (27%). Female: 39 applicants of which 9 were granted (27%). Male: 35 applicants of which 11 were granted (31%).

In addition, we can mention that during consultations or drop-in sessions we identify if a project is technically suitable. If a project is not technically sound, that project never enters our support system.