



ELIXIR Node Application Form – Hungary

Part A: Summary of ELIXIR Node

A1: Synopsis

Governance:

Services and tools to be involved in ELIXIR are selected and ranked by their relevance to ELIXIR communities and by the expected impact both nationally and internationally. Potential tools, databases, etc. will be pre-screened and endorsed by the **institutional lead scientists** and the final ranking will be approved by the **Head of Node**.

Services and databases:

The Hungarian ELIXIR Node will provide services in the Data, Tools and Training platforms. The Tools services include **UniTmp** (which is a unified source for transmembrane protein topologies and structures, including five tools and four servers), **IUPred** and **ANCHOR** (characterization of disordered protein regions), **IsoMut** (unique mutation detection in isogenic sample groups), **Kaplan-Meier Plotter** (meta-analysis based biomarker assessment), **G-2-O** (linking mutation or CNV of a gene to survival), **MÉTA** Programme (website encompassing vegetation heritage data of Hungary) and **Kooplex** (virtual research environment infrastructure for collaboration of multicenter research units). The Data services consist of the **National Cancer Registry** (a population-based database that collects all cases with malignancies in Hungary), UniTmp's databases and the **Hungarian Oncogenome Portal** (interactive large scale data depository initiated in 2017 for whole genome and/or exome sequencing data). Involved organizations will provide expert guidance in using the abovementioned tools. To highlight **Training services**, University of Debrecen will provide focused practical courses of 1-2 weeks in a wide spectrum of topics from database mining to basic programming and using open source bioinformatics tools. Pázmány Péter Catholic University will offer the analysis of metagenomic data as well as biomedical images with their **Field Programmable Gate Arrays (FPGA) platforms**.

Organizations involved:

The Node will be hosted by the Institute of Enzymology at the Research Centre for Natural Sciences of the Academy of Sciences (**RNCS HAS**). Institutes of the Academy include the Rényi Institute (**RENYI**), the Centre for Ecological Research (**CER HAS**), and the Centre for Agricultural Research (**CAR HAS**). The Node will cover major universities in Hungary including the Eötvös Loránd University (**ELTE**), University of Debrecen (**UD**), Pázmány Péter Catholic University (**PPCU**), Semmelweis University (**SE**), University of Pécs (**PTE SZKK**) and University of Veterinary Medicine (**UVM**). Extra-university organizations include the National Institute of Oncology (**NIO**) and The National Agricultural Research and Innovation Centre (**NARIC**). The participating organizations are research institutions with active bioinformatics projects.

Impacts:

The Node will strengthen **collaboration** of the diverse and distributed Hungarian bioinformatics community. The Node will enable the organization and coordination of **resources at the national level**. ELIXIR will drive the **international integration** of the national bioinformatics community and will further increase the visibility of research groups involved.

Finance:

Resources dedicated to the Node will include a total of **four positions** (4 FTE) and an annual budget of 25k EUR. Administrative costs will be covered by the **RCNS HAS**.



A2: Abstract

The Hungarian ELIXIR Node will include **twelve different Hungarian universities and research institutes**. The **five major scientific focus** include human genomics (medical genomics and cellular biotechnology), agricultural genomics (microbial biotechnology), proteomics, veterinary sciences, and ecology (marine genomics). The **largest ongoing database** in development comprises the national cancer initiative targeting to establish a comprehensive cancer genomic database. **Flagship tools** with a total citation over 1000 include the IUPRED (protein structure prediction), PDBTM (transmembrane protein databank), the KM-plotter (cancer survival analysis), and HMMTOP (topology prediction of transmembrane proteins). Members of the Hungarian ELIXIR node expressed interest to joint all together nine different ongoing and planned **use cases**.

A3: Snapshot of services

Services to be provided by the Node (delete the ✓ in each category not appropriate)

Data: ✓	Compute:	Training: ✓	Tools: ✓	Standards:
---------	----------	-------------	----------	------------



Part B: Administrative details

B1: Organisation(s) involved in the Node

Organisations involved in the Node	Coordinating organisation: Research Centre for Natural Sciences, Hungarian Academy of Sciences (RCNS HAS) Organisations involved: <ul style="list-style-type: none">• Eötvös Loránd University (ELTE)• University of Debrecen (UD)• Pázmány Péter Catholic University (PPCU)• Semmelweis University (SE)• University of Veterinary Medicine (UVM)• National Institute of Oncology (NIO)• National Agricultural Research and Innovation Centre (NARIC)• Centre for Ecological Research, Hungarian Academy of Sciences (CER HAS)• Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences (RENYI HAS)• Institute for Veterinary Medical Research, Centre for Agricultural Research, HAS (CAR HAS)• University of Pécs, Szentágotthai Research Centre (PTE SZKK)
---	--

B2: Contact details of Head of Node

Name	<input type="text" value="Balázs Gyórfy"/>	Title	<input type="text" value="Dr"/>
Contact address	<input type="text" value="Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences
H-1117 Budapest, Magyar tudósok körútja 2., Hungary"/>		
Tel	<input type="text" value="+36 1 382 6744"/>	Fax	<input type="text"/>
Email	<input type="text" value="gyorffy.balazs@ttk.mta.hu"/>		
Job title	<input type="text" value="group leader"/>		
Organisation	<input type="text" value="Research Centre for Natural Sciences, Hungarian Academy of Sciences (RCNS HAS)"/>		



Lead Scientific Representatives

Name	<input type="text" value="Gábor Tusnádý"/>	Title	<input type="text" value="Dr"/>
Contact address	<input type="text" value="Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences H-1117 Budapest, Magyar tudósok körútja 2., Hungary"/>		
Tel	<input type="text" value="+36 1 382 6709"/>	Fax	<input type="text"/>
Email	<input type="text" value="tusnady.gabor@ttk.mta.hu"/>		
Job title	<input type="text" value="group leader"/>		
Organisation	<input type="text" value="Research Centre for Natural Sciences, Hungarian Academy of Sciences (RCNS HAS)"/>		
Name	<input type="text" value="István Csabai"/>	Title	<input type="text" value="Prof"/>
Contact address	<input type="text" value="Dept. of Physics of Complex Systems, Eötvös Loránd University, H-1117 Budapest, Pázmány Péter sétány 1/A, Hungary"/>		
Tel	<input type="text" value="+36 1 372 2826"/>	Fax	<input type="text" value="+36 1 372 2866"/>
Email	<input type="text" value="csabai@complex.elte.hu"/>		
Job title	<input type="text" value="professor"/>		
Organisation	<input type="text" value="Eötvös Loránd University (ELTE)"/>		
Name	<input type="text" value="István Kenessey"/>	Title	<input type="text" value="Dr"/>
Contact address	<input type="text" value="National Institute of Oncology H-1122 Budapest, Ráth György utca 7-9., Hungary"/>		
Tel	<input type="text" value="+36 20 666 3448"/>	Fax	<input type="text"/>
Email	<input type="text" value="kenessey.istvan@oncol.hu"/>		
Job title	<input type="text" value="head of center, National Cancer Registry"/>		
Organisation	<input type="text" value="National Institute of Oncology (NIO)"/>		



Name Title

Contact address

Tel Fax

Email

Job title

Organisation

Name Title

Contact address

Tel Fax

Email

Job title

Organisation

Name Title

Contact address

Tel Fax

Email

Job title

Organisation

Name Title

Contact address

Tel Fax

Email

Job title

Organisation



Name Title

Contact address

Tel Fax

Email

Job title

Organisation

Name Title

Contact address

Tel Fax

Email

Job title

Organisation

Name Title

Contact address

Tel Fax

Email

Job title

Organisation



Name	<input type="text" value="Norbert Solymosi"/>	Title	<input type="text" value="Dr"/>
Contact address	<input type="text" value="Department of Animal Hygiene, Herdhealth and Veterinary Ethology, University of Veterinary Medicine
H-1078 Budapest, István utca 2., Hungary"/>		
Tel	<input type="text" value="+36 1 478 4242"/>	Fax	<input type="text"/>
Email	<input type="text" value="solymosi.norbert@gmail.com"/>		
Job title	<input type="text" value="associate professor"/>		
Organisation	<input type="text" value="University of Veterinary Medicine (UVM)"/>		

Name	<input type="text" value="Attila Gyenesei"/>	Title	<input type="text" value="Dr"/>
Contact address	<input type="text" value="University of Pécs, Szentágothai Research Centre,
H-7624 Pécs, Ifjúság útja 20, Hungary"/>		
Tel	<input type="text" value="+36 72 501 668"/>	Fax	<input type="text" value="+36 72 501 654"/>
Email	<input type="text" value="gyenesei.attila@pte.hu"/>		
Job title	<input type="text" value="senior research fellow"/>		
Organisation	<input type="text" value="University of Pécs, Szentágothai Research Centre (PTE SZKK)"/>		

B3: Curriculum vitae of Head of Node

CV of Balázs Gyórfy, the head of the node is the first annex to this Application.

B4: Short curriculum vitae of each Lead Scientific Representative

CVs of 12 representatives of RCNS HAS, ELTE, UD, PPCU, SE, NIO, NARIC, CER HAS, RENYI HAS, CAR HAS, UVM and PTE SZKK are the second annex to this Application.

Part C: Operational description of Node

C1: Description of services to be provided

The Hungarian ELIXIR node will provide following services, tools, and training platforms:

Data services

- **UniTmp:** Unified source for transmembrane protein topologies and structures, including five servers:
 - a. Human Transmembrane Proteome (HTP, <http://htp.enzim.hu/>),
 - b. Protein Data Bank of Transmembrane Proteins (PDBTM, <http://pdbtm.enzim.hu/>),
 - c. Topology Data Bank of Transmembrane Proteins (TOPDB, <http://topdb.enzim.hu/>),
 - d. Conservatively Located Domains and Motifs in Proteins (TOPDOM, <http://topdom.enzim.hu/>) and
 - e. Target Selection for human TransMembrane Proteins (TSTMP, <http://tstmp.enzim.ttk.mta.hu/>).
- **National Cancer Registry:** a population-based database that collects all cases with malignancies in Hungary. In addition to ICD-10 diagnostic codes, oncology care hospitals reported follow-up data, characteristic features of the tumor as well as basic treatment information can be found.
- **Hungarian Oncogenome Portal:** interactive large scale data depository initiated in 2017 for whole genome and/or exome sequencing data generated from ~400 tumor biopsies from Hungarian patients with various hematological malignancies, colon cancer, breast cancer, melanoma, and lung cancer with detailed clinical annotation. The database is under construction and will include a functional module for performing all relevant bioinformatic analyses in order to carry out reliable variant calling in tumor samples. (<http://genomics.hu/projects/oncogenome/>)

Tools services

- **IUPred** and **ANCHOR:** Tools to characterize disordered protein regions. IUPred is a widely used tool for predicting disordered regions in proteins, ANCHOR predicted disordered binding regions. These tools are targeted for both the life science researchers, as well as ELIXIR Nodes.
- **IsoMut:** is a tool designed for experimental scenarios whenever unique mutations (single nucleotide variations or insertions/deletions) are sought in multiple isogenic samples. The software uses information gathered from all the available samples to rule out false positive mutations with great efficiency. (<http://www.genomics.hu/tools/isomut/isomut.html>)
- **Kooplex:** is an infrastructure for flexible collaboration of multicenter research units. Kooplex combines collaborative data sharing and analysis in a virtual research environment. (<https://kooplex.vo.elte.hu/>)
- **UniTmp:** Unified source for transmembrane protein topologies and structures, including four tools:
 - a. Constrained Consensus TOPology prediction server (CCTOP, <http://cctop.enzim.ttk.mta.hu/>),

- b. Transmembrane helices and protein topology prediction (HMMTOP, <http://www.enzim.hu/hmmtop/>),
 - c. Detection of transmembrane regions (TMDET, <http://tmdet.enzim.hu/>) and
 - d. Transmembrane proteins' membrane segments folding recognizing (TmFoldWeb, <http://tmfoldweb.enzim.ttk.mta.hu/web/>).
- **Kaplan-Meier Plotter:** The Kaplan-Meier Plotter is capable to assess the effect genes on survival using breast, ovarian, lung and gastric cancer samples. Primary purpose of the tool is a meta-analysis based biomarker assessment. (<http://kmplot.com>)
 - **G-2-O:** linking mutation (or CNV) of a gene to clinical outcome (survival) by utilizing next generation sequencing and gene chip data. (<http://www.g-2-o.com/>)
 - **RecurrenceOnline:** a tool capable to predict response to hormonal treatment, to targeted therapy and survival (recurrence score) for breast cancer patients using gene expression data obtained by Affymetrix gene chips. (<http://www.recurrenceonline.com/>)
 - **MÉTA Programme:** website encompassing vegetation heritage data of Hungary, including distribution maps of habitat types, plant invasion of habitat types. (<http://novenyzetiterkep.hu/english>)
 - **CoNSEnsX+:** is a tool to analyze structural ensembles of proteins by assessing their compliance to structural parameters derived from NMR measurements. Can be used both for ordered and disordered proteins and includes an option to select sub-ensembles with better correspondence to the parameters chosen.

Training

- Involved organizations may provide expert guidance in using the abovementioned tools.
- **UD** will follow the model of training courses established at the Pasteur Institute in Paris where MSc and PhD students participate in focused practical courses of 1-2 weeks together with experienced researchers. The courses will be open to wider audience for a fee. Topics of the courses will follow a wide spectrum from database mining to basic programming and using open source bioinformatics tools.
- **PPCU** will offer the analysis of metagenomic data as well as biomedical images with their Field Programmable Gate Arrays (FPGA) platforms. Analysis of complex data sets using systems biology tools like SiComPre and deep learning approaches will also be offered. It will provide expert guidance in using the tools for the generation and analysis of protein structural ensembles, and will also provide the necessary infrastructure for calculation and analysis tasks that exceed the capacity of the open services
- **PTE SZKK** will provide short (2-3 half-day) targeted courses in bioinformatics/data analysis, biostatistics and scripting/programming for academic research groups and industrial participants. These training activities will be coordinated with other members of the Hungarian Node. We will contribute to TESS (ELIXIR's Training Portal – eSupport System) and participate in the Train the Trainer, Researcher and Developer training programme. We also plan to be active member of the Software and Data Carpentry Working Group.

C2: Description of provenance and technical sustainability

Tools

- **IUPred, ANCHOR, ISOMUT, Kooplex, Hungarian Oncogenome Portal:** ELTE central infrastructure serves 30000 users on more than 1000 devices. The management is highly trained and regular presenter on conferences; many of them are working with much larger external systems as consultants. In collaboration with the Hungarian Academy of Sciences (MTA) cloud services and other research centers, the web services affiliated with the ELTE university provide services that are built on public data are freely available for academic research.
- **UniTmp, Kaplan-Meier Plotter, RecurrenceOnline, G-2-O:** All data that will be provided by RCNS HAS come from academic sources and are now publicly available. Where possible, tools are distributed with an open source license. All resources, pipelines, online services, and databases can be used for free by scientists working at universities and (non-profit) research institutes but are licenced for companies. Data sources and analysis tools will, for the moment, be maintained by the research groups and institutes where they are currently developed and deployed. Since, especially in the beginning, most of the data sources and pipelines offered to ELIXIR will be the ones that have a long and proven track record already, we do not foresee major problems with their technical sustainability in the short and middle-long term.
- **MÉTA:** During the MÉTA Programme the whole territory of the country was surveyed. Satellite images were superimposed with GIS layers, topographic maps and forest management data to support the field surveys. Data sources and analysis tools will, for the moment, be maintained by Functional Ecology Department, Institute of Ecology and Botany, CER HAS where they are currently developed and deployed.

Database

- **National Cancer Registry:** a population-based database that collects all the cases with malignancies in Hungary by reports of oncology care hospitals. Registration is mandatory. The Registry is regulated by the order of the Hungarian Government and maintained by the National Institute of Oncology.

Training

- **UD** is already running bioinformatics courses for MSc students under the guidance of Dr Endre Barta, Dr Beata Scholtz. Several departments have pieces of bioinformatics trainings based on their specific interests. UD has a large supercomputing unit and several mid and large sized training facilities with computer terminals that can be used for training purposes. The team behind the application has significant experience in organizing international practical courses (HHMI, FEBS, Research Training Networks etc).
- The majority of the code developed at **PPCU** is open access and the rest is provided freely. Protocols including our own and third-party programs are available openly. The necessary hardware for the applications and trainings will be provided using the machine park available at PPCU. If a given collaboration exceeds the available capacities, services provided by the Hungarian Supercomputing Centre will be used on a case-by-case basis.
- **PTE SZKK** has a new, state-of-the-art research building (called "Science Building") with all the necessary equipment and tools for the training activities. It also hosts a supercomputer and has standardized data analysis pipelines and workflows.



C3: Service provision tables

Application area of data services to be provided by the Node

Medical: 2	Agri-food: 4	Environmental: 3	Pharma: N/A	Biology: 1	Other:
------------	--------------	------------------	-------------	------------	--------

Application area of tools and/or services to be provided by the Node

Medical: 1	Agri-food: 4	Environmental: 3	Pharma: 5	Biology: 2	Other:
------------	--------------	------------------	-----------	------------	--------

C4: Expected Impact of the Node

Scientific and technical impact:

Providing the developed resources and services to the ELIXIR community has impact not only for the ELIXIR but for Hungarian scientists as well by developing more effective and quick applications and by standardizing these resources and procedures. The ELIXIR Node will be important by promoting collaboration and increasing the visibility of the tools. In addition, the ELIXIR Node would be also important for helping the increasing number of users of various bioinformatics tools for next generation sequencing and single cell transcriptomic in various biological projects, including oncology, immunology, neurology and microbiology. UD will have a strong interaction with local clinicians to involve them in the collection and genomic as well as proteomic analysis of samples obtained from patients. PPCU with deep learning, and the use of highly parallel platforms such as FPGAs is expected to help closing the gap between theoretically possible and technically feasible data analysis pipelines. The International Committee for Taxonomy of Viruses (ICTV) plans to introduce higher level viral taxons, but viruses do not have common genes thus it is very hard to infer their full evolutionary tree. To overcome this obstacle bioinformatics expertize and infrastructure needed to predict pathogenicity and function of non-isolated viruses to facilitate food safety and the advancement of science.

Impact on national bioinformatics community:

The ELIXIR Node is also expected to strengthen the education of bioinformatics at both undergraduate and graduate levels at universities. It can help students to access current bioinformatics research problems through practical hands on projects. This type of approach will also be central to a novel Biotechnology MSc course that will be launched by the joint effort of ELTE and Budapest University of Technology and Economics (BME) in 2018. As a result of training activities the number of researchers who have access to the research services can increase, which brings efficiency to the research programs and as a network effect, the number of future collaborations can thrive.

Economic and societal impact:

The new HPC cluster that will be installed in NARIC in 2017 will host not only the suitable programs and databases for Hungarian agricultural research but also the genomics related data of several Hungarian breeds such as the mangalica pig and the red deer, which was *de novo* sequenced and annotated in NARIC. Research in CER HAS with objectives like sustainable fisheries and agro-ecology can also contribute to the agricultural sector.

Lastly, it is crucial for the diverse and scattered Hungarian bioinformatics community to strengthen collaboration. Providing the aforementioned services by ELIXIR will require the organization and coordination of resources at the national level. ELIXIR will drive the integration of the national bioinformatics community and will further increase its visibility.



C5a: Description of organizations involved and what services they currently run

RCNS HAS

At the Institute of Enzymology, RCNS HAS four research groups have working experience on computational chemistry, bioinformatics of proteins, genomics and transcriptomics in the field of oncology, prenatal development among others. The provided tools for ELIXIR are highly accessed, based on the usage stats visible on their websites. Major research groups participating in ELIXIR include the Oncology Biomarker Research Group, the Systems Biology of Reproduction Research Group and the Bioinformatics Research Group.

ELTE

Bioinformatics groups at the Eötvös Loránd University develop various bioinformatics tools. These include the IUPred and ANCHOR servers. IUPred is one of the most commonly used sequence based disorder prediction methods. IUPred was also built into several other resources such as Pfam (<http://pfam.xfam.org/>), ELM (<http://elm.eu.org/>), D2P2 (<http://d2p2.pro/>). IUPred is part of MobiDB-Lite which provides disorder annotation to Interpro (<https://www.ebi.ac.uk/interpro/>). The ANCHOR server is another widely used method to identify disordered binding regions in proteins. Both web servers have >99% site uptime serving and several hundred thousands requests per year. The programs are also available as downloadable program packages. Additional tools developed at the ELTE include programs for gene enrichment and phylogenetic analyses, and tools to study brain connectivity. Groups also apply and develop machine learning and data mining techniques for specific problems. They are also involved in providing bioinformatics services for collaborative research (Kooplex) as well as for next generation sequencing (IsoMUT). The analyses of next generation sequencing and single cell transcriptomics is applied in various research projects related to the fields of immunology, neurology and microbiology.

UD

UD was a pioneer among the Hungarian universities in using the most advanced functional genomics techniques in biomedical research. Together with an Illumina HiScan sequencer a high performance computing cluster was installed as well (84 cores, 35TB disk capacity, high memory node etc.). This cluster is dedicated completely to the analysis and storing of next generation sequencing data and will be significantly expanded by the end of 2017. The Department also hosts two proteomics databases. TRANSDAB (http://genomics.dote.hu/wiki/index.php/Main_Page) transglutaminase substrates and interaction partners collection. FuzDB (<http://protdyn-database.org/>) compiles experimentally observed fuzzy protein complexes, where intrinsic disorder is maintained upon interacting with a partner. During these years UD implemented the education of Bioinformatics both at MSc and PhD level. Moreover UD organized several bioinformatic and genomic training courses with the involvement of experts from the EMBL, DKFZ, EMBO, HHMI and relevant commercial players of the field (e.g. GeneSpring). These training activities will be the basis of further to be organized bioinformatic training activities.

PPCU

Special emphasis is put on massively parallel architectures and machine learning methodologies. This has been increasingly complemented by gaining expertise in biomedical applications such as image processing, metagenomics and protein structural analysis. Online services currently run: The CoNSEnsX+ server for the evaluation and selection of dynamic structural ensembles (consensx.itk.ppke.hu)

NIO

Aggregated epidemiological data is freely available from the website according to basic categorizations. Based on individual permission deeper data is also available from the department directly.

PTE SZKK

PTE SZKK is a new research institute established on the basis of modern international science organizational and management normative. It covers all aspects of education, research and innovation at the fields of biomedical, natural and environmental sciences. The infrastructure, instrumentation and expertise of the 22 research groups operating in the building provide an excellent basis to become a well-known, leading research facility in Hungary, as well as in Central-Europe with an extensive and fruitful collaboration network. The newly establishing bioinformatics group already provides data analysis services and statistical support for pharmaco-, functional- and molecular genomics projects.

C5b: Table of involvement in relevant pan-European or international consortia

<i>Name of consortium</i>	Collaborative management platform for detection and analyses of (re-) emerging and foodborne outbreaks in Europe (COMPARE)
<i>URL</i>	http://www.compare-europe.eu/
<i>1-sentence summary of what the consortia does</i>	COMPARE is a multidisciplinary research network that has the common vision to become the enabling analytical framework and globally linked data and information sharing platform for the rapid identification, containment and mitigation of emerging infectious diseases and foodborne outbreaks.
<i>1-sentence summary of the role of ELTE in the consortium</i>	Developing a Data Portal for a collaborative data sharing and analysis in a virtual research environment.

<i>Name of consortium</i>	COST Action BM1405 - Disprot- Central Consortium
<i>URL</i>	https://ngp-net.bio.unipd.it/
<i>1-sentence summary of what the consortia does</i>	The aim of this COST Action is to understanding of overlapping phenomena concerning non-globular proteins (NGPs) and their joint consensus classification.
<i>1-sentence summary of the role of ELTE in the consortium</i>	We actively participate in this consortium by providing annotation, maintaining and developing tools (IUPred, ANCHOR, Mobidb-Lite) and participating the in the development of functional ontologies for disordered proteins.

<i>Name of consortium</i>	eLTER RI
<i>URL</i>	http://www.lter-europe.net/document-archive/elter-esfri-files
<i>1-sentence summary of what the consortia does</i>	Long-Term Ecosystem Research in Europe
<i>1-sentence summary of the role of CER HAS in the consortium</i>	Still in the planning phase



<i>Name of consortium</i>	European Clinical Research Infrastructure Network (ECRIN)
<i>URL</i>	http://www.ecri.org
<i>1-sentence summary of what the consortia does</i>	ECRIN, as one of the ESFRI initiative, is public, non-profit organization that links scientific partners and networks across Europe to facilitate multinational clinical research.
<i>1-sentence summary of the role of PTE SZKK in the consortium</i>	HECRIN, the Hungarian Center of ECRIN, is located at PTE SZKK and coordinates all the related clinical research activities in Hungary.

<i>Name of consortium</i>	Joint Action on Rare Cancers (JARC, 3rd Health Programme)
<i>URL</i>	http://jointactionrarecancers.eu
<i>1-sentence summary of what the consortia does</i>	JARC is aimed to integrate and maximize efforts of the EU Commission, EU Member States and all stakeholders to advance quality of care and research on rare cancers.
<i>1-sentence summary of the role of PTE SZKK in the consortium</i>	To delineate optimal resources for undergraduate medical education, to promote the improvement of European medical expert training instruments via the European Union of Medical Specialists (UEMS), as well as SIOPE – European Society for Paediatric Oncology, to provide recommendations on education of non-medical experts, patient advocates and patient communities involved in patient care, as a means to improve rare cancer patient empowerment in Europe.

C6: Interaction with other ELIXIR partners

ELTE

Active collaboration with Silvio Tosatto (University of Padova, Italian ELIXIR node). Zsuzsanna Dosztányi is a member of the Interpro consortia through Mobidb-Lite, developed in collaboration with Silvio Tosatto to provide disorder related annotation to Interpro and Uniprot (EBI, ELIXIR Central Node).

In collaboration with the Data Centre at Wigner Research Institute, Budapest we are part of the COMPARE H2020 collaboration. In that project, together with EBI ENA (ELIXIR Hub), UK and DTU, Denmark we are developing a Data Portal (<http://compare-europe.hu/>), a collaborative data sharing and analysis a virtual research environment. The planned platform enables a better surveillance system for infectious diseases, speed up the detection of, and response to disease outbreaks among humans and animals worldwide using genomics approach.

One of the near future established Biomarker Lab (<http://genomics.hu/projects/biomarkerlab/>) objectives is to sequence the functional antibody repertoire from patients with systemic autoimmune diseases using high-throughput DNA sequencing technologies. One of their main focuses is to examine the effects of novel therapeutic approaches, such as monoclonal antibodies, on the B cell populations and functional antibody repertoire. Different tools of bioinformatics analyses will be applied for the studies, including IMGT, the international ImMunoGeneTics information system (ELIXIR, French node). The active usage of this tool can possibly turn into partnership in the future.

UD

The EMBL node can interact with the University of Debrecen by providing expertise both in the field of Genomic bioinformatics but also through their extensive training activities. Since the University of Debrecen, and other Hungarian Units are in a shared network with EMBL (Teaming H2020 Molmedex) this shared work can provide synergies and can enforce the ELIXIR activities.

PTE SZKK

Already ongoing collaborations on training activities with VIB, leader of the ELIXIR Belgium Node, and especially with Alexander Botzki who is leading Training Coordinator of the Belgium node. Attila Gyenesei, the Lead Scientific Representative of PTE SZKK, has many ongoing EU projects with ELIXIR CZ (Brno) on training and research infrastructure activities, and on Galaxy with ELIXIR Germany and ELIXIR Belgium.

C7: Use cases

The list below shows shared interest between members of ELIXIR Hungary and the ongoing and planned Uses Cases:

Use case	Contact	Email	Institute
Marine metagenomics	Ferenc Jordán	jordan.ferenc@gmail.com	CER HAS
	Sándor Pongor	pongor@itk.ppke.hu	PPCU
Human Genomics and Translational Data	Bálint Bálint	balintblaszlo@gmail.com	UD
	Balázs Gyórfy	gyorffy.balazs@ttk.mta.hu	RCNS
Rare disease community	István Balogh	balogh@med.unideb.hu	UD
	Attila Gyenesei	gyenesei.attila@pte.hu	PTE SZKK
Proteomics	Gábor Tusnády	tusnady.gabor@ttk.mta.hu	RCNS
	Attila Csikász-Nagy, Zoltán Gáspári	csikasz-nagy.attila@itk.ppke.hu	PPCU
Galaxy	Balázs Gyórfy	gyorffy.balazs@ttk.mta.hu	RCNS
	Attila Gyenesei	gyenesei.attila@pte.hu	PTE SZKK
Microbial biotechnology	Sándor Pongor	pongor@itk.ppke.hu	PPCU
Intrinsically Disordered Proteins	Mónika Fuxreiter	fmoni@med.unideb.hu	UD
	István Simon	simon.istvan@ttk.mta.hu	RCNS
	Zsuzsanna Dosztányi	dosztanyi@gmail.com	ELTE
Human Cell Atlas	Zoltán Gáspári	gaspari.zoltan@itk.ppke.hu	PPCU
Human copy number variation	Katalin Monostory	monostory.katalin@ttk.mta.hu	RCNS

Ferenc Jordán is a scientific adviser at the Danube Research Institute, Centre for Ecological Research HAS, Budapest, Hungary and research associate at Stazione Zoologica, Napoli, Italy. He has expertise in computational systems ecology and marine ecosystem modelling.

The **Pongor** group is interested in agent based modelling of microbial communities with special regard to stability and collapse phenomena. The traditional interest of the group is the identification specific genes in genomic and metagenomic datasets, especially genes involved in intercellular communication.

Bálint L. Bálint MD PhD is the Laboratory Head of Centre for Clinical Genomics and Personalized Medicine and UD-GenoMed Medical Genomic Technologies Ltd., both based at University of Debrecen. He has cooperated with Ivo Gut previously, which can be a favorable connection to the Rare Disease Use case.

István Balogh MD, PhD is the Head of Department of Clinical Genomics at University of Debrecen. The

department can join in the Rare Disease work group as a member of the European Reference Network (MetabERN). His Inherited Diseases Research Group investigates the molecular background of genetic, primarily monogenic diseases.

Gábor Tusnády's Membrane Protein Bioinformatics Research Group in RCNS, HAS focuses to transmembrane proteins. They have launched 7 new or renewed publicly available web servers and databases in the last couple of years - including the Human Transmembrane Proteome Database (<http://htp.enzim.hu>)

The group of **Attila Csikasz-Nagy** at PPCU works on integrative approaches to quantitatively predict protein complex compositions from proteomics data. The group also works on phosphoproteomics and proteomics data analysis in collaboration with researchers from Denmark (University of Copenhagen) and Switzerland (ETH Zurich).

The group of **Zoltán Gáspári** is interested in the structure and function of proteins and their complexes in the postsynaptic density. Their aim is the mathematical and structural modeling of these complexes making extensive use of experimental data related to changes in protein levels.

Balázs Gyórfy is active at the Research Centre for Natural Sciences of the Hungarian Academy of Sciences (RCNS). He has a major focus on human diseases and genome-level data analysis. His particular interest includes the development of online tools and analytical services.

Mónika Fuxreiter has established a connection already with Silvio Tosatto and Norman Davey through the Biomedicine and Molecular Biosciences COST Action BM1405.

The group of **Katalin Monostory** deals with interindividual differences in drug metabolism and elimination that is one of the main sources of therapeutic failure or undesired side effects. The research group is involved in deciphering the biological and clinical significance of copy number variations of drug metabolizing CYPs in therapy resistance of various tumors.

István Simon is a member of the Hungarian Academy of Sciences, professor emeritus researcher. His group has been working with intrinsically disordered proteins (IDP) for more than a decade. Recent work denotes a subcategory of the IDPs, which can form stabile complexes only by themselves.

Attila Gyenesei, lead scientific representative of PTE SZKK, has an ongoing collaboration with Björn Grüning and Frederik Coppens, leaders of the ELIXIR Galaxy team, on Galaxy development. PTE SZKK has a special interest in rare diseases, oncogenetics beside common diseases (member of the Joint Action on Rare Cancers initiative).



Part D: Summary of finances and resources

D1: Staff deployment table

Staff deployment	2018	2019	2020
Head of Node	0.25	0.25	0.5
Technical Coordinator	0.25	0.25	0.5
Training Coordinator	0.3	0.3	0.5
Bioinformatician, Data Curator	1	1	2
Administration	0.5	0.5	0.5
Total	2.3	2.3	4

D2: Operating costs table

Expenditure (k€)	2018	2019	2020	Total
Staff costs	13.8	13.8	24	51.6
Running costs	4	4	4	12
Equipment	0	0	0	0
Other costs	3	3	3	9
Total	21.8	21.8	31	74.6

D3: Description of resources to be committed

The Node's staff resources will be covered by the grants of the partners. Other costs consist of mainly travel expenses by the Node personnel and participants to meetings and also the publication and outreach costs. The running costs include the setting up and maintenance of a website allowing ELIXIRHUN partners and the member states to access resources developed and maintained by ELIXIRHUN. The costs showed in the tables are estimates.

Part E: Financial sustainability of Node E: Funding received or committed

RCNS HAS

- OTKA K119287, Transcriptome analysis of transmembrane proteins, running 2016-2020, 40.7 M HUF budget, PI: Gábor Tusnády
- OTKA, KH125607, Developing high-throughput experimental method for investigation of the structure of transmembrane proteins, running 2017-2019, 20.0 M HUF budget, PI: Gábor Tusnády
- OTKA K115698, Endeavour for a unified description of protein structure formation, running 2016-2020, 42.3 M HUF budget, PI: István Simon
- HAS Momentum 2014-2019, Cancer biomarker discovery, 248M HUF budget, PI: Balázs Gyórfy
- FIEK_16-1-2016-0005 grant, A preeclampsia genetikai vizsgálata, 230M HUF budget, PI: Gábor Than

UD

University of Debrecen could commit staff of the Genomic Core facility as part time members in the training unit. A National Grant (EFOP with allocation of 0,5M EUR) could provide part time allocation for resources (25% from the Grant activities). Moreover being part of the ELIXIR network, would ease the path to become a service and training unit for the HCEMM (Hungarian Center of Excellence in Molecular Medicine) within the already operational H2020 Molmedex EMBL Teaming project.

ELTE

- FIEK_16-1-2016-0005, Developing molecular biomarker research and service center for industrial needs, Running 2017-2021, 2700M HUF, PI: Imre Kacskovics.
<http://genomics.hu/projects/biomarkerlab/>
- Momentum Grant of the HAS LP2014-18/2016. Running 2014-2019. Large-scale exploration of linear-motif mediated protein interactions 277M HUF. PI: Zsuzsanna Dosztanyi

PPCU

- PPCU: Hungarian Human Resources Development Operational Programme (EFOP-3.6.2-16-2017-00013) consortial grant, Thematic collaborations for innovative collaborations in informatics and information technology, running 2017-2020, 200 M HUF budget, PI (from PPCU): Attila Csikasz-Nagy
- NKFIA 120650, Microbiome Bioinformatics: Computational analysis and modelling of bacterial communities, running 2016-2020, 32 M HUF budget, PI: Sándor Pongor

CER HAS

- GINOP 2.3.2-15-2016-00057, Az evolúció fényében: elvek és megoldások, 2017-2020, 749 M HUF, PI: Szathmáry Eörs
- OTKA K 116071, comparative analyses of aquatic food webs, 2015-2019, 48 M HUF, PI: Ferenc Jordán

SE

- Department of Pathology and Experimental Cancer Research, Semmelweis University
- Momentum Grant of the Hungarian Academy of Sciences, Personalised and stratified medicine in cancer. 2015-2019, 100 M HUF budget PI: Dr. Csaba Bödör
- NVKP Grant of the National Research, Development and Innovation Office, Hungarian Oncogenome and Personalised Tumor Diagnostics and Therapy National Program, 1 457 M HUF budget, 2017-2019, PI: Prof. András Matolcsy/Dr. Csaba Bödör

- Department of Pharmacodynamics, Semmelweis University (Gyorgy Bagdy, Gabriella Juhasz):
- Hungarian Brain Research Program 2, „Comparison of brain mechanisms in migraine and tension type headache”, 2018-2021, 100 M HUF budget, PI: Dr. Gabriella Juhasz
- Hungarian Brain Research Program 2, „Novel antidepressant targets: genome-wide gene-environment interaction studies”, 2018-2021, 120 M HUF budget, PI: Prof. Gyorgy Bagdy

PTE SZKK

- GINOP-2.3.3-15-2016-00025 “Kutatási központ kialakítása speciális sejtek funkcionális és genomikai analizésére a Pécsi Tudományegyetemen”, 2017-2020, 480K EUR
- GINOP-2.3.2.-15-2016-00048 "Életet veszélyeztető Akut megbetegedések súlyossági és hálózati mutatóinak javítása transzlációs orvostudományi megközelítésben" - STAY ALIVE, 2017-2021, 624K EUR
- OTKA- K119540 “Ritka-és diagnosztizálatlan betegségek" molekuláris hátterének és eredetének vizsgálata új generációs genomikai módszerekkel”, 2016-2020, 157K EUR
- GINOP-2.3.2-15-2016-00039 Research of the pathogenesis of rare diseases, innovation of underlying new diagnostic and therapeutic procedures, 2017-2021, 1.136K EUR
- GINOP-2.3.2-15-2016-00021 “Chip-technológia alkalmazása a humán in vitro fertilizáció eredményességének javításába”, 2016-2019.07, 6.500K EUR
- GINOP-2.3.2-15-2016-00047 An analytical and diagnostic research excellence center for health and sport performance, 2017-2021, 506K EUR
- GINOP- 2.3.2-15-2016-00021”Chip-technológia alkalmazása a humán in vitro fertilizáció eredményességének javításában”, 2016-2019, 5.000K HUF
- H2020-RUR-06-2016 Crop diversification and low-input farming across Europe: from practitioners’ engagement and ecosystems services to increased revenues and value chain organization, 2017-2021, 506K EUR
- Horizon 2020-732482 Parallel network-based biocomputation: Technological baseline, scale-up and innovation ecosystem. 2017-2020. 22K EUR

Next steps: Submitting your ELIXIR Node Application

Please submit your completed form by e-mailing it to hub@elixir-europe.org.

Each Node application will be presented to the ELIXIR Board by the relevant Board member from that country. Following agreement by Board Members, the SAB will then review the Node Application, before the Application is presented to the ELIXIR Board for approval. The SAB and the ELIXIR Board may request additional information on any aspect of the Node, including reserving the right to interview those individuals involved in the Node if deemed appropriate. Upon its acceptance, the successful Node Applicant shall conclude a Collaboration Agreement with the EMBL, acting on behalf of the ELIXIR Hub. By signing the Collaboration Agreement the Node Applicant shall become an ELIXIR Node.