Bioinformatics, genetics (Clinical & Health Science Infrastructures and HR platforms)

Andreas Puschmann

1. **Short description of the infrastructure.**
   MultiPark supports up to 50% FTE of a clinical bioinformatician at the Dept. of neurology for projects within the aims of MultiPark.

   The exact name of the support in the 2018-2020 MultiPark budgets was "***Bioinformatics (50% FTE or user fees at NBIS for example) clinic***", placed under "Infrastructure and salaries" – "Clinical" –"Neurology/Neurosurgery".
   Budget for 2018-2020: 425 000 kr (per annum)

   In the MultiPark Budget for 2021, the post was renamed "***Bioinformatics, genetics (local person and/or fees to NBIS)***" moved under a separate heading "Bioinformatics" under "Clinical & Health Science Infrastructures and HR platforms".
   Budget for 2021: 300 000 kr

   <u>Description of the infrastructure, according to its steering document, as submitted to the MP Coordinators:</u>
   "The bioinformatician works with bioinformatic analyses of genetic data from high throughput genetic sequencing from patients and families with Parkinson disease, dementia, other movement disorders and cerebrovascular disorders. Tasks/expertise include the identification of disease-associated genetic variants in families and in case series and comparison with international genetic databases. Sequence alterations and copy number variants in human DNA can be detected. Algorithms to detect short tandem repeats are being developed during 2021.
   A typical research question can be to screen datasets from a number of patients for potentially disease-causing variants, or to compare variants between affected and unaffected members of one family in order to identify a genetic variant that co-segregates with the disease.
   Such work leads to the identification of patients with (rarer) monogenetic forms of neurodegenerative and other neurological disorders. These rare patients may then be included in other research, or biomaterial may be collected for experimental studies (outside of this infrastructure)."

   This infrastructure complements the faculty-wide infrastructure for high-throughput genetic analyses, the *Center for Translational Genomics* (CTG). CTG perform the actual (wet lab) analyses of human DNA from patients (whole exome and now almost exclusively whole genome sequencing) and CTG also do the first steps of bioinformatic processing, including alignment of the reads to reference genomes and variant calling and some annotation. In other words, CTG provides us with lists of all genetic variants identified in an individual.
   However, CTG do <u>not</u> provide further analyses such as comparing data from affected and unaffected family members to identify new genetic loci or variants causing disease, or genetic association studies, or any other explorative research analyses of the raw data provided.
   Neither CTG nor Region Skånes counterpart for clinical NGS analyses, the *Center for Diagnostic Medicine*, can perform more advanced calculations such as those necessary

to identify novel disease-associated repeat expansions in human DNA. These require very large computing space and access to WGS data on extensive numbers of individuals for comparison. For these operations, the infrastructure uses short term support from *National Bioinformatics Infrastructure Sweden (NBIS)* as stipulated in the budget.

2. **Is this infrastructure receiving support also from other Strategic Research Areas (SRAs) or organizations at Lund University (e.g. Medical faculty, LBIC). If yes, please specify the type of support and its amount.**
   No. This infrastructure is co-funded by individual researchers and projects. At present, a bioinformatician (MSc in medical genetics, MSc in bioinformatics) works 80% within the projects named above, amounting to more than 50% co-funding from individual researcher and project grant within these projects.

3. **Number and names of MultiPark senior researchers using the infrastructure in the period 2018-2020[1].**
   Two: Andreas Puschmann, Elisabet Englund (also in 2021)

   In 2021, MultiPark senior researcher Prof. Paul-Visse has approached the infrastructure as she requires its services for bioinformatic analyses of mitochondrial genetic data in a collaboration with Dr. Johannes Ehinger, Mitochondrial Medicine research group, Clinical neurophysiology. The infrastructure now awaits a formal application but probably the infrastructure can be used for this project.

4. **Number and names of senior researchers outside of Multipark and/or non-academic partners using the infrastructure 2018-2020.**
   Two:
   I 2020-2021, Maria Landqvist Waldö MD PhD, Kliniska Vetenskaper, Helsingborg and Kognitiv medicin, Ängelholms sjukhus, Region Skåne used the infrastructure for analyses of whole exome data from patients with rare forms of familial frontotemporal dementia.
   In 2021, The bioinformatician works 20% for prof. Arne Lindgren, Lund Stroke Register, within a multi-center study on stroke genetics, not receiving financing from MultiPark.

5. **Does the infrastructure have a steering document accessible to the users? If yes, when was it last updated?[2]**
   Yes. 22 March 2021, submitted to MP coordinators.

6. **Is the infrastructure charging user fees? If yes, state the amount and what is covered by the user fees.**
   Yes. 50% of the costs need to be financed by individual projects/researchers. The infrastructure coordinator finances more than 50% of time for the bioinformatician because MP support has decreased in 2021.

7. **List publications generated with the help of this infrastructure during the past 3 years (2018-2020). Do not include manuscripts in preparation and please give**

[1] If the infrastructure was first established in 2020, please include this information.

[2] Note that the Multipark leadership may ask to see this document with a very short notice.

the full reference (i.e., complete author list, complete title, journal name with year, volume, pages)[3].

1. Ilinca A, Englund E, Samuelsson S, Truvé K, Kafantari E, Martinez-Majander N, Putaala J, Håkansson C, Lindgren A, Puschmann A. *MAP3K6* mutations in late-onset neurovascular disease causing stroke, cognitive impairment and tremor. *Neurol Genet* **2021**;7:e548.
2. Gorcenco S, Ilinca A, Almasoudi W, Kafantari E, Lindgren AG, Puschmann A. New generation genetic testing entering the clinic. *Parkinsonism Relat Disord*. **2020**;73:72-84.
3. Kafantari E, Andréasson S, Säll T, Puschmann A. Do variants in *IRF2BPL* cause both neurological disorders and keratoconus 8? *Parkinsonism Relat Disord*. **2020** Oct 12:138-140.
4. Ilinca A, Martinez-Majander N, Samuelsson S, Piccinelli P, Truve K, Cole J, Kittner S, Soller M, Kristoffersson U, Tatlisumak T, Puschmann A, Putaala J, Lindgren A. Whole-Exome Sequencing in 22 Young Ischemic Stroke Patients With Familial Clustering of Stroke. *Stroke*. **2020**;51:1056-63.
5. (Ygland Rödström E, Soto-Ortolaza AI, Englund E, Kafantari E, Dickson DW, Wszolek ZB, Puschmann A, Ross OA. Genomic analyses of a large Swedish multi-incident kindred with autosomal dominant Parkinson's disease with dementia. *Manuscript included in a doctoral thesis in print, **2021-10-12***.)

The infrastructure was newly established in 2018 and there were no publications in the first 2 years as a lot of technical details had to be solved initially.

[3] If the infrastructure was first established in 2020, please include this information here too.