International Agency for Research on Cancer

World Health Organization

Governing Council **GC/59/11**
Fifty-ninth Session 28/03/2017

Lyon, 18–19 May 2017
Auditorium

REQUEST FOR USE OF FUNDS FROM THE GOVERNING COUNCIL SPECIAL FUNDS: PURCHASE OF SCIENTIFIC EQUIPMENT

1. Laboratory sciences, biostatistics, bioinformatics and epidemiology are closely integrated in the IARC Medium-Term Strategy. This requires high quality laboratories and the availability of state-of-the-art scientific equipment. Regular upgrade and acquisition of scientific instruments and information technology equipment are thus essential to support this strategy.

2. Centralized platforms (e.g. next-generation sequencing (NGS) or mass spectrometry) are available to scientists across the Agency. A high success rate in obtaining competitive research grants has resulted in a significant increase in the number of samples tested on these platforms, and in the volume of data acquired and analysed. A new plan has been made to support developments in bioinformatics (see document SC/53/8), an area essential to interpret the complex datasets generated.

3. The Director would like, therefore, to request the Governing Council to provide an allocation from the Governing Council Special Fund to purchase equipment in the following areas:

a) Upgrade of the IARC scientific computing capacity

4. The field of bioinformatics is making an increasingly important contribution to research into the causes and prevention of cancer. Similarly, data sharing across the scientific community is creating a vast array of *in-silico* resources. Both have enormous potential in IARC's multi-disciplinary studies but also rely heavily on bioinformatics to deal with these complex datasets.

5. The current computing capacity at IARC mostly comes from the High Performance Computing (HPC) cluster purchased in 2012 and further updated in 2015. IARC scientists also have access to a private cloud environment hosting shared hardware, software and storage resources, dedicated to general purpose computation and development.

6. The Agency progressed from four users of the HPC cluster in 2014 to 19 active users in September 2016, from seven different scientific Groups (EGE, ENV, GCS, GEP, ICB, MMB, and NMB)[1]. The current level of usage is approaching the maximum capacity, recognizing the need to provide for peaks in demand and to avoid long waiting times for analyses.

[1] EGE = Epigenetics Group
GCS = Genetic Cancer Susceptibility Group;
ICB = Infections and Cancer Biology Group;
ENV = Section of Environment and Radiation;
GEP = Genetic Epidemiology Group;
MMB = Molecular Mechanisms and Biomarkers Group;

7. In order to address these pressing needs a modular scalable cluster is required with a minimum of double the current processing and storage capacities. In addition, a dedicated backup and archive environment will be established to protect important raw data and analysis results.

8. Maintenance of in-house capacity and direct access to advanced technologies and bioinformatics is important to maintain the Agency's scientific program and attractiveness to high-quality staff. At the same time, IARC meets part of its requirements through a policy of collaborative partnerships with local centres of expertise, for example with the "*Plateforme de Bioinformatique Gilles Thomas"* of the Synergie Lyon Cancer foundation at the Centre Léon Bérard, Lyon, in order to provide capacity, avoid redundancy and overspecialization in these areas.

9. The proposed equipment would be operated as a shared resource under the responsibility of the IT Working Group of the Bioinformatics and Biostatistics Steering Committee. The Committee will ensure access and support for research groups across IARC, and will continue to advise the Director on the cost benefits of internal computing capacity and adoption of a hybrid model including cloud-based solutions to cover utilization peaks.

b) Upgrade of the IARC next-generation sequencing (NGS) platform

Benchtop sequencer of 100–120 Gb capacity

10. An important line of research conducted by several groups at IARC involves the successful identification and functional evaluation of genetic and epigenetic alterations in different cells and tissues. For this research, IARC scientists currently use small-throughput sequencers which are well-suited for low-complexity sequencing strategies but whose capacity does not allow for more complex applications where cost-effectiveness can be achieved by higher-level multiplexing.

11. Consequently, a desktop sequencer with higher capacity would allow the cost-effective implementation of novel sequencing applications of critical interest for multiple research groups, which are currently only available through fragmented third-party arrangements. A cost comparison demonstrated the better value of the proposed investment compared to an outsourcing option. In contrast, high-throughput applications will continue to be performed through strategic (external) partnerships.

12. The proposed equipment would be operated as a shared resource overseen by existing laboratory technicians under the responsibility of EGE, MMB and GCS, who would provide access and support for other research groups at IARC. Dedicated Information Technology Services (ITS) support is available to this equipment. Data analysis and storage will be accommodated by the above request for an expansion of ITS infrastructure and computational capacity.

c) Automated system to study cancer chromatin at genome-wide level

13. Additional support is needed to maintain and upgrade IARC's capacity to perform analyses of chromatin at the genome-wide level, as the current number of robotics-based chromatin studies is expected to increase four- to five-fold over the coming years. Notably the scale of studies is increasing as the approaches find more routine applications within large-scale projects, thus

NMB = Nutritional Methodology and Biostatistics

requiring specialized robotics to automatically and efficiently perform sample preparation, and to provide high-quality data at a reduced labour cost and greater reproducibility and efficiency.

14. The proposed system will increase the existing capacity and simultaneously allow for efficient and cost-effective library preparation for NGS platforms. This is an important consideration given the proposed upgrade of NGS and would be of interest to multiple research groups (EGE, MMB, GCS, ICB). At the same time it meets the increased need for the cost-effective in-house preparation of libraries, irrespective of a sequencing upgrade.

15. The Scientific Council reviewed the proposal as detailed below:

Requested budget

	Approximate price (€)
a) Upgrade of the IARC scientific computing capacity	300 000
b) Upgrade of the IARC next-generation sequencing (NGS) platform	310 000
c) Automated system to study cancer chromatin at genome-wide level	90 000
Total equipment	700 000

16. The Scientific Council noted that the annual maintenance costs of the requested equipment will be covered by the regular budget as well as by collaborative programmes through grant applications.

17. The Scientific Council, noting the importance of maintaining key research platforms and infrastructures, recommended that the Governing Council approves the above-mentioned purchase of scientific equipment and provide a total allocation of €700 000 from the Governing Council Special Fund.

18. The Governing Council is requested to approve at its 59th Session in May 2017 the use of €700 000 from the Governing Council Special Fund (GCSF). As at 01 January 2017, the GCSF account has an opening fund balance of €12.1 million, of which over 60% have already been committed. More details of the projection of GCSF account from 2017 to 2019 are provided in the document GC/59/Inf.Doc. No.2.