Open-source ImmGen: mononuclear phagocytes

To the Editor:

ImmGen (Immunological Genome Project) is a collaborative group of immunology and computational biology laboratories that aims to determine, on a broad and highly granular scale, the patterns of gene expression and genetic-regulatory circuits of immune-system cells in the mouse[1]. Each of the participating immunology laboratories brings deep knowledge about a given cell lineage and the expertise to frame questions directly pertinent to its function. Data are used for the computational reconstruction of gene-regulatory networks, to define modules of coregulated genes and key regulatory factors predicted to govern their expression and generally explore new algorithmic space to benefit from the unique breadth of the group's data[2,3]. In its current endeavors, ImmGen is exploring the immune system's genomic response to diverse challenges, relating gene expression to chromatin organization, and mapping molecular landscapes beyond mRNA. Perhaps most importantly, and consistent with its support from the National Institute of Allergy and Immunological Diseases, ImmGen is conceived as a community resource, whose data are publicly available through a website (www.immgen.org), custom data browsers, and smartphone apps.

To generate data that can be reliably cross-compared across a very wide range of immunocyte lineages, ImmGen uses a common data-generation pipeline and rigorously follows Standard Operating Protocols, with close attention to contamination issues. ImmGen laboratories were first vetted via the sorting of reference cell types. In practice, cell populations are sorted in ImmGen laboratories, frozen, and shipped to the core team for profiling by microarray or RNA sequencing through the use of a common genomic platform and bioinformatics data pipeline. This configuration has ensured a coherent data set, limiting the platform and batch variations that confound computational analyses. It has also had drawbacks. Unfortunately, it has not been possible to incorporate data from other sources or to benefit from the unique expertise of other colleagues in the preparation of particular immunocytes. Finally, the inter-related

cell populations within a lineage can be defined in different frames of reference through the use of different sets of markers, and different groups can use different such 'cosmologies'. This issue arises in particular for populations with inherent complexity or cell plasticity or that have been defined only recently, such as innate lymphoid cells or the mononuclear phagocyte subsets. ImmGen profiling is, quite naturally, performed according to the framework used by the ImmGen laboratory responsible for a given lineage, but it would be valuable to establish the correspondence with definitions used by others.

Thus, and in the context of revisiting its compendium with new profiling technology, the ImmGen group will attempt to complement its usual mode of operation with an open 'call for samples'. In this format, investigators who do not belong to the ImmGen group are invited to contribute and prepare, using their expertise and marker criteria but following the ImmGen Standard Operating Protocols, cell populations that they are interested in and that they feel would make a valuable addition to the ImmGen compendium. Our current protocol for 'ultra-low-input RNA sequencing' has been optimized to generate high-quality profiles from 1,000 sorted cells, which brings the requisite latitude for the analysis of rare cell populations. Lysates of sorted cells will be shipped to the ImmGen core team for profiling by this technique. Immediately after being generated, raw and pre-processed data will be sent to contributors, who will help verify the integrity and quality and will otherwise be free to use the data as they wish. Investigators will also interact with ImmGen computational biologists, who will lead the integrated exploitation of the data. Results will be included on the ImmGen website for public display after curation of the larger data group, per the 'Fort Lauderdale rules' for pre-publication release of community resource data (https://www.genome.gov/10506537/). ImmGen will assume all costs associated with expression profiling and data processing, but contributing investigators will be asked to cover the costs associated with cell sorting.

In the spirit of the 100th anniversary of the passing of Eli Metchnikoff, who discovered these

cells[4], this 'open-source' project will begin with the area of mononuclear phagocytes (macrophages, dendritic cells and monocytes). These are populations for which recent research has uncovered a rich diversity of origin and phenotype. In line with ImmGen scope, submissions may encompass finely resolved mononuclear phagocyte populations across differentiation, tissue location or specific activation (*ex vivo* only). We anticipate that investigators might contribute somewhere between a few to a dozen cell populations, representing a unique contribution that complements the reference populations already profiled by ImmGen core groups.

We are asking interested investigators to send a detailed description of the cell populations they would propose to contribute (i.e., tissue of origin and exact sorting markers used for identification and/or exclusion) and a short justification of the inclusion of these cells in this project (form available at http://www.immgen.org/OS.MNP.xls). The mouse used as a standard for ImmGen profiling is a 5-week-old C57BL/6 male mouse, but transgenic and knock-in reporter mice with the same genetic background can also be used as needed. An evaluation of sex differences is planned. Some degree of redundancy is expected and is in fact desirable, but to limit duplication and ensure a good fit with the ImmGen scope, the proposals will be reviewed, discussed and refined with the contributing investigators and will be collated with the cell populations profiled by the ImmGen groups into a rich and coherent program.

We believe that experimenting with this mode of discovery has the potential to generate exciting and useful new data and knowledge and to accentuate the community's involvement with the ImmGen effort.

Christophe Benoist for the ImmGen Consortium

Department of Microbiology and Immunobiology, Harvard Medical School Boston, Massachusetts, USA.
e-mail: cbdm@hms.harvard.edu

1. Heng, T.S. *et al. Nat. Immunol.* **9**, 1091–1094 (2008).
2. Jojic, V. *et al. Nat. Immunol.* **14**, 633–643 (2013).
3. Mostafavi, S. *et al. Cell* **164**, 564–578 (2016).
4. Kaufmann, S.H. *Nat. Immunol.* **9**, 705–712 (2008).