

Volunteers tested 15 trillion signatures for Mapping Cancer Markers project

In this update, coinciding with the 17th anniversary of WCG, the research team summarizes the contribution of all volunteers to each type of tumor and signature size that Mapping Cancer Markers has studied so far.

Project: [Mapping Cancer Markers](#)

Published on: 16 Nov 2021

Background

[Mapping Cancer Markers](#) (MCM) aims to identify the markers (sometimes referred to as signatures) associated with various types of cancer. The project is analyzing millions of data points collected from thousands of healthy and cancerous patient tissue samples. So far, these have included tissues with lung cancer, ovarian cancer, and sarcoma.

Looking at the prodigious volunteer contribution

Each Mapping Cancer Markers work unit tests multiple groups of biomarkers against a cancer dataset for use as diagnostic or prognostic signatures. Currently, that dataset is our sarcoma dataset. MCM has explored three datasets so far: lung cancer, ovarian cancer, and sarcoma. For lung and ovarian cancer, MCM explored different signature length (i.e., different number of genes included in the signature), while for sarcoma the signatures have all the same length but different compositions (i.e., they include markers measured with different techniques, in variable percentages). At the time of this report, the volunteers analyzed about 800 million work units.

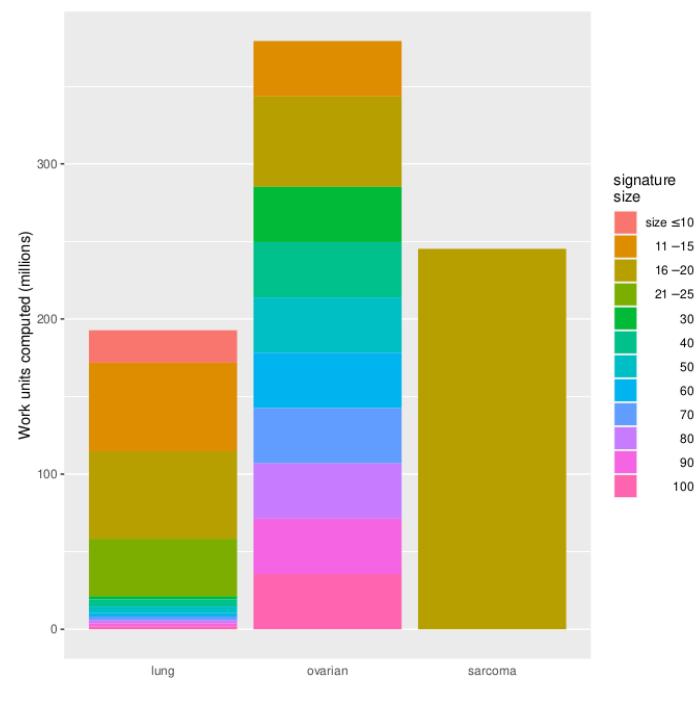


Figure 1. Number of completed work units per cancer type and signature size.

A work unit will test signatures of a specific size (number of biomarkers) against its dataset. Added together, World Community Grid members have tested about 15 trillion signatures, a number that would have been unimaginable to test without your support.

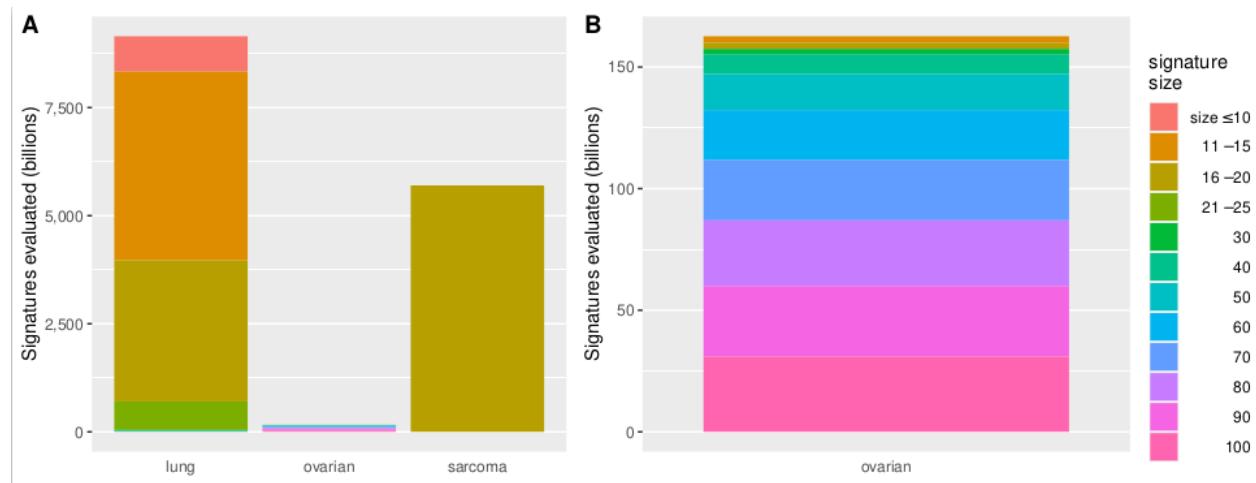


Figure 2. Evaluated signatures by dataset (and size) for all tumours (A) and for ovarian cancer only (B).

The compute time required per signature in any given work unit depends on the signature size and the dataset. Since we try to keep the total amount of computation per

work unit constant, the number of signatures per work unit will also vary with the signature size and dataset.

In most datasets, larger signatures take more compute time than shorter signatures. Our lung dataset generally follows this pattern, while for our ovarian dataset, the opposite is true. For the ovarian cancer dataset, MCM tested prognostic signatures that predicted short or long survival times. This task was inherently more difficult than the diagnostic tasks in lung (cancer/no cancer) or sarcoma (distinguishing sarcoma subtypes). Because of this difficulty, MCM could not compute nearly as many signatures per ovarian work unit as it did for other datasets, regardless of signature size.

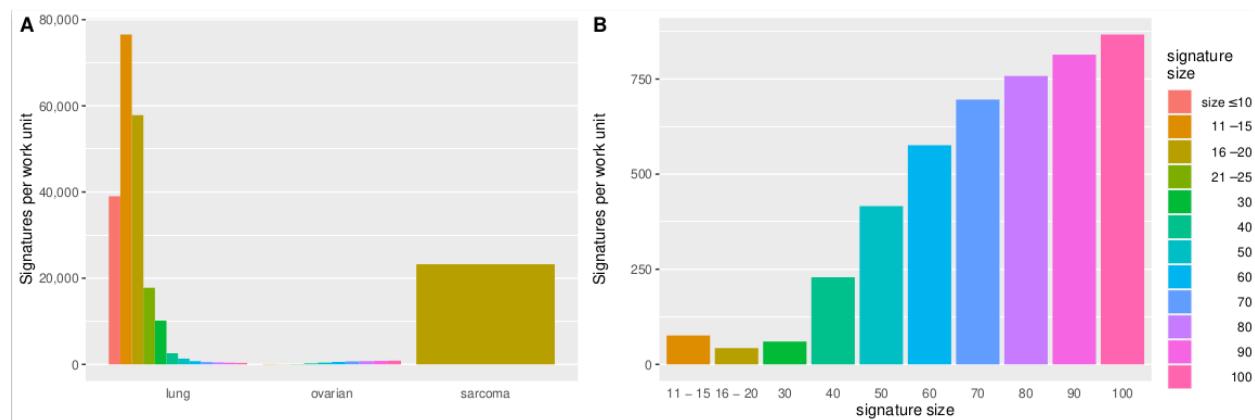


Figure 3. Signatures evaluated per work unit, by dataset (and size), for all tumours (A) and (B) specifically for ovarian cancer.

Our next step will be to evaluate how the different composition of sarcoma signatures affect their predictive ability.

We are grateful to everyone who is supporting Mapping Cancer Markers, and all the important projects on World Community Grid.

If we are celebrating the WCG 17th anniversary it is only because of your dedication.

Thank you for your continued support!