Science 2 February 2007:
Vol. 315. no. 5812, pp. 583 - 584
DOI: 10.1126/science.315.5812.583b

NEWS OF THE WEEK

METABOLIC RESEARCH:
Canadian Group Claims 'Unique' Database

Eliot Marshall

No media splash greeted the completion of the Human Metabolome Project last week, although the sponsor boasts it is "the chemical equivalent of the Human Genome Project." The University of Alberta in Edmonton, Canada, announced that a small group of researchers with $7.5 million from the Canadian government has created a "comprehensive" database of human metabolites, calling it "the starting point for a new era in diagnosing and detecting diseases."

Since 2004, the scientists have assembled an inventory of 2500 molecules produced by metabolic reactions in the body's tissues and fluids. The Canadian project has been low profile until now, but the trumpeting of a completed "first draft" metabolome astonished some observers. Noting that the human genome was 90% complete when geneticists announced a draft, Gary Siuzdak of the Scripps Research Institute in San Diego, California, says, "I would be surprised if 2500 [entries] represents even 10%" of the human metabolome.

Still, the metabolome project leader, David Wishart, a biophysicist and computer scientist at the University of Alberta, says he's pleased. The product, described in a recent article in *Nucleic Acids Research*, is a free, public database (www.hmdb.ca) that some observers view as a solid first effort. Wishart and 39 co-authors--including five from the University of Calgary in Canada--describe the database as "unique."

The collection, which opened on 1 January, covers endogenous human

metabolites, mainly from human tissue or gut bacteria. (A separate database covers drug metabolites.) Wishart says the project began with "text mining" of "dusty textbooks and obscure journals" to scoop up and validate previously identified metabolites.

Importantly, says Wishart, the collection provides more than 400 searchable "fingerprints," images of the atomic spectra of metabolites captured with nuclear magnetic resonance or mass spectroscopy. Users studying a specific metabolite can call up a "MetaboCard" that shows this fingerprint, if available, along with data on disease relevance, biofluid concentration, metabolic pathway, and many other topics.

Yet other metabolome researchers point out that there's a lot missing from the collection. Wishart and his colleagues narrowed their

inventory to metabolites found at a concentration of 1 micromole or more. This simplified the task but eliminated tens of thousands of substances. Although Wishart says the present collection is 95% complete at 2500 entries, the group's paper acknowledges that if every small molecule in the body were to be included, "the number of compounds might exceed 100,000."

It's hard to define what's been left out, says Jeremy Nicholson of Imperial College London, head of an industry-backed project on potential toxicities in metabolic interactions--many of them difficult to detect (*Science*, 11 November 2005, p. 965). He views the database as "just a list" of detectable metabolites, although a useful list.

Ian Blair, a metabolomics leader and vice chair of pharmacology at the University of Pennsylvania, says the Canadian project made "quite a good start" on a complete database, although comparing it to the human genome is "over the top." Siuzdak, biochemist Julian Griffin of the University of Oxford, U.K., and several others are each leading independent efforts to build metabolome databases that could rival the Canadian effort.

Wishart acknowledges that the response to the metabolome's first draft has been "mixed." And he agrees that the estimate that the catalog is 95% complete is based on "a bit of a fuzzy number." But he considers the Canadian metabolome database the best and most user-friendly available. His big concern now is how to keep the project afloat after 2007, when government funding is scheduled to end.

Science. ISSN 0036-8075 (print), 1095-9203 (online)