

Bio21:MMIM: Molecular Medicine Informatics Model

Newsletter 2: April 2005



Type 1 Diabetes evolves over many years



diabetes.

The Bio21:MMIM project is very exciting for diabetes research, as it allows the possibility of 'building resources that would last forever'. The group can 'already ask questions across the database that we hadn't been able to do previously'.

When deciding what data to include in Bio21:MMIM, Colman also included a large amount of clinical research data on patients with diagnosed diabetes. Included in Bio21:MMIM is 10 years of clinical research information about diabetes treatment, cholesterol levels, blood pressure, kidney and eye complications. The possibility for new discoveries are endless and as Colman says, Bio21:MMIM is 'pretty unique actually'.

A/Prof Colman is obviously passionate about his work, as he explains how it was originally thought that pancreatic cells didn't regenerate. Now it is believed they do, which is important for stem cell research. If they regenerate, unfortunately they may be again attacked by the immune system. This makes the research into preventing type 1 diabetes also relevant to future successful transplantation. However, it also suggests that there could be a cure if the auto-immune reactivity can be halted, saving the pancreatic cells.

In summary, he agrees that people often think type 1 diabetes is

a disease in which patients will require insulin injections for the rest of their lives; but he says, 'we hope that our research will alter this'. Hopefully joining the Bio21:MMIM

database and answering the many research questions will accelerate the progress towards improved treatment and diabetes prevention.

Progress Report: February-March 2005

The Bio21:MMIM group has been busy testing systems in the last two months, with numerous aspects of the system being trialled. The colorectal cancer aspect of the database (the most advanced in the system) is now functional for cross-institutional queries, including ACCORD (Austrian Comprehensive Clinical Outcomes and Research Database a cancer patient registry), protein, and public genetic database searches. Researchers can also query across the cancer and diabetes databases. All information for the diabetes database is currently being converted into a new model, which will make querying this data easier and more logical.

Some epilepsy patients have been tested for polymorphism (genetic variation see next issue). Training and testing for the Federated Data Integrator (FDI), and linkage to the public databases has been undertaken.

Other databases have been added to the system, including several public genomic and proteomic databases, the RMH/Austin Tissue Bank, and a microarray database. These are

still being improved and tested.

There are now a number of databases that can be accessed through Bio21:MMIM: Melbourne Health Local Research Repository (LRR); Austin LRR; Peter MacCallum LRR; as well as the public databases: Genbank; PubMed; LocusLink and UniProt/Swiss-Prot. These databases consist of 301 tables which contain over two million records, or segments of information about patients (or more general information in the case of public databases). The databases are quite large, with the biggest (RMH) being over one gigabyte. The FDI, which contains a lot of 'virtual information', contains 281 megabytes (MB) of data. At the last count there were over 35,000 patients in the FDI.

In other news, all hardware is now installed at all sites. The Virtual Private Network (VPN) and Secure Sockets Layer (SSL) internet connectivity and security) have been implemented. The Unique Subject Index (USI) system has been implemented and tested. The USI system had been affected originally by data quality, but each site is now working to 'clean' its data and make it more unified.

Bio21:MMIM was presented at the Bio21 Symposium on the 25th of February and was well-received. A Bio21:MMIM website is also currently in the works. It will include data dictionaries; table descriptions and details of sample selection; information about training sessions; application forms and the process to apply for Bio21:MMIM data; and other useful and interesting information and links.

Next issue: Focus on Epilepsy.

Benefits for researchers and doctors

The Molecular Medicine Informatics Model is a unique system linking genetic and clinical research data, enabling both medical and scientific researchers new opportunities to study diseases, compare treatment outcomes, or analyse detailed genetics.

Bio21:MMIM provides research data within a framework of privacy for patient clinical and genetic data. Five ethics committees have approved the access rules to the system and privacy is protected by de-identifying or codifying the health data. Patients can be re-identified by the ethics committee if something that can help the patient is discovered. Intellectual property issues are also addressed (see What About Intellectual Property?).

A major benefit of Bio21:MMIM is its size. Bio21:MMIM allows analysis of data from thousands of patients across multiple institutions and from various different databases. These include the clinical research data, genetic data, tissue bank and microarray data, and information from public databases such as GenBank and SwissProt. Cross-disease group queries are enabled, so that users can search for patients with cancer and diabetes, for example. All this information can be combined and studied in one place.

The broad scope of data in Bio21:MMIM will allow researchers to infer possible

outcomes and treatments for patients. For example, patients may be more efficiently selected for clinical trials based on their genetic or clinical profile. Treatment with available medication can be turned into a more individual process, as more becomes known about genetics and the effects of medication benefits as well as side effects. Eventually, it could even lead to testing for diseases (such as cancer and diabetes) before symptoms appear, allowing early intervention.

Bio21:MMIM is a system

that unifies information across multiple institutions. Each patient is given a Unique Subject Identifier (USI) to protect their privacy. It also allows researchers to track patients across institutions covered by Bio21:MMIM. Even if a patient changes hospitals for any reason, or undertakes clinical trials or research in a separate institution, a query in Bio21:MMIM will give the whole medical and genetic picture.

Bio21:MMIM is in a growth phase and is planning to keep adding more databases of other

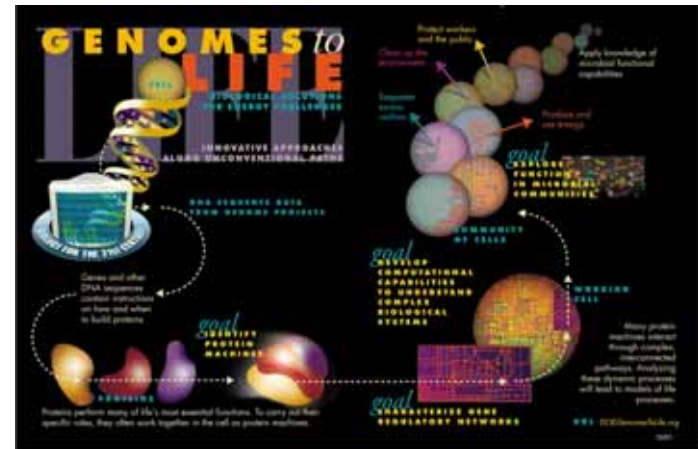


Image credit: U.S. Department of Energy Genomics:GTL Program,

All organisms face three information challenges, and all life on earth uses the same fundamental biochemical strategies to meet these challenges. First, the organism must encode and store, within each cell, all the instructions needed to build, operate, maintain, and reproduce itself and to respond to varied environmental conditions. DNA, the biochemical solution to this coding and storage problem, is made up of four chemical building blocks (nucleotide bases): adenine (A), thymine (T), cytosine (C), and guanine (G). These building blocks are organized in long chains like chemically linked beads, whose precise order spells out the organism's full set of genetic instructions -- its genome.



BIO21:MMIM Contacts: Robert Merriel
Email: robert.merriel@mh.org.au
Marianne Hibbert
Email: Marianne.Hibbert@mh.org.au



diseases and from more institutions. So far the list of involved institutions include Bio21; Melbourne Health; the Royal Melbourne Hospital, including the RMH Tissue Bank; the University of Melbourne; Bayside Health; the Alfred Hospital; Austin Health; the Peter MacCallum Clinic; the Walter and Eliza Hall Institute; Western Health; the Ludwig Institute for Cancer Research (Melbourne), and Cancer Trials Australia.

How will it work?

When a new tool is available to researchers and practitioners, the first question is always 'how does it work?' Bio21:MMIM is a powerful system that can be used to search for specific details within a disease database, as well as for cross-disease queries.

For example, a researcher wishes to find all instances in the diabetes database of people with elevated antibody levels, and to see how many of them have diabetes. Using the Query Management Facility (QMF Bio21:MMIM's Internet-based interface), the researcher selects which database should be searched. The diabetes data set includes three studies following families over time. Data from any of these can be selected, or all can be searched. The antibody level threshold for several different antibodies and diabetes outcome can then be selected. The results of this query will generate separate de-identified lists of patients who are Antibody positive, Antibody negative, and with and without diabetes. Results can be sorted into antibody levels.

Cross-disease queries can be performed, for example all participants of a certain age group or specific gender, who have diabetes as well as colorectal cancer, can be selected.

Bio21:MMIM now

includes several public databases which can be searched. For example, once all patients with diabetes and colorectal cancer have been found, their genetic profiles can be studied further using tools such as GenBank and LocusLink (DNA information), UniProt (protein information), and PubMed (academic literature).

After their queries are complete all results of searches must be analysed statistically by researchers. Bio21:MMIM combines a large amount of data



Image credit: U.S. Department of Energy Human Genome Program, <http://www.ornl.gov/hgmis>

Each DNA molecule contains many genes--the basic physical and functional units of heredity. A gene is a specific sequence of nucleotide bases, whose sequences carry the information required for constructing proteins, which provide the structural components of cells and tissues as well as enzymes for essential biochemical reactions. The human genome is estimated to comprise more than 30,000 genes.

that can allow statistically significant research to be completed, efficiently using existing data to test research questions with the ultimate aim of improving health care.

What about Intellectual Property?

One major concern that researchers may have with a system like Bio21:MMIM is how their intellectual property (IP) rights are protected. A legal team was hired to work with participating institutions to

develop a solution to this problem. They developed detailed guidelines to ensure that each researcher's rights to discoveries are protected. Before gaining access to Bio21:MMIM, researchers and their institutions must sign their acceptance to this protocol.

The IP of Bio21:MMIM projects is defined as either Background IP, which is IP relating to the databases that each institution or research group has contributed to Bio21:MMIM; or Project IP, which is the IP arising from research done using the Bio21:MMIM. Background IP is retained by the party that supplied the relevant data. Each party grants a non-exclusive license for other Bio21:MMIM users to use its Background IP appropriately within their research project. Bio21:MMIM users must comply with any conditions set by the institution supplying the data.

Before beginning a project, each research group must appoint a Principal Investigator, identify the Background IP that they are bringing to the project, the data they wish to access and identify which party will be the Commercialisation Lead. The Commercialisation Lead is responsible for any efforts made to commercialise the Project IP. Each project participant agrees to ensure that they provide all possible assistance in protecting any Project IP.

Identification of Project IP depends on the participants' awareness of IP principles such as 'first to invent' and its requirements, including documentation of progress and workbook management.

Diabetes

Diabetes is a condition in which the body tissues cannot make use of glucose, the body's preferred energy source. As a result, the glucose level in the blood rises to dangerous levels. There are three main types of diabetes: Type 1, Type 2, and Gestational diabetes.

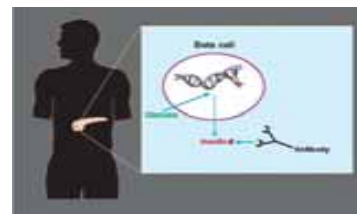
Type 1 diabetes was previously known as juvenile diabetes or insulin-dependent diabetes. In Type 1 diabetes, insulin, the hormone responsible for moving glucose into the body's cells, is not produced by the beta cells in the pancreas. This is thought to be due to an autoimmune response to the pancreas, in other words, the immune system mistakenly attacks pancreatic cells and gradually destroys them, eventually leaving the body with an insulin deficiency.

It is known that this autoimmune process occurs in people with a certain genetic profile. It is presumed that an environmental insult, possibly a viral infection, then triggers the immune process. Type 1 diabetes accounts for around 15% of all people with diabetes. Although type 1 diabetes is usually diagnosed in people aged 30 or younger it can happen at any age. Type 1 diabetes is diagnosed when the body's insulin production falls to a critical level at which glucose begins to build up in the blood stream. At this point, approximately 90% of the pancreatic cells have already been destroyed.

Type 2 diabetes, also known as adult onset or mature onset diabetes, accounts for the majority of patients with diabetes. In Type 2 diabetes, insulin is produced, but the body's cells are resistant (insulin resistance). Type 2 diabetes can be diagnosed at any age, but is more common in adults over 40 years. The mechanisms causing insulin resistance are still not totally understood. Type 2 diabetes can develop without any warning symptoms as the glucose gradually builds up in the blood stream.

Gestational diabetes occurs in pregnancy and is also caused by insulin resistance. It is thought to occur due to production of various hormones from the

Type 1 diabetes is an autoimmune disorder



The autoimmune nature of Type 1 Diabetes

Beta Cells in the pancreas make and release insulin, a hormone that controls the level of glucose in the blood. GAD (Glutamic Acid Decarboxylase) is a normal enzyme found in all cells that initiates the metabolism of glutamic acid. The presence of antibodies to GAD in the blood is an early indication of the start of the autoimmune process in Type 1 Diabetes.

placenta, that interfere with insulin function and result in high blood glucose levels. It usually develops in the second half of the pregnancy, and is now routinely tested for. If untreated, it can cause macrosomia (high birth weight) or hypoglycaemia (low blood sugar) in the baby. Approximately one in 20 pregnant women will develop gestational diabetes, with an elevated risk if there is family history of diabetes. Gestational diabetes is temporary and usually resolves after pregnancy, but it does indicate a higher risk for developing Type 2 diabetes later in life.

Spotlight on Diabetes

Associate Professor Peter Colman is a consultant endocrinologist and is the head of the Royal Melbourne Hospital's Diabetes and Endocrinology Department. After completing his medical degree, he undertook research training in Melbourne and then Boston, before returning to Melbourne. His work now includes research in early diagnosis and prevention of Type 1 diabetes and clinical care of patients with diabetes.

His research focuses mainly on children and young adults with type 1 diabetes. When asked why he chose to work on Type 1 diabetes, Colman noted the difficulties people with diabetes often experience in managing the condition. He hopes that the research will lead to greater knowledge of the genetics, triggering and pancreas immune attack which occurs in Type 1 diabetes. Increased understanding will hopefully ultimately lead to better treatments or even prevention.

Colman is researching the possibility of detecting type 1 diabetes long before symptoms develop. Clearly, the strong genetic background in type 1 diabetes puts relatives of patients at increased risk. However, environmental factors probably determine whether diabetes will actually develop.

In order to study this, A/Professor Colman and Professor Len Harrison at Walter and Eliza Hall Institute are running several studies of people who don't (yet) have type 1 diabetes but are at increased risk because of the presence of antibodies to the pancreas cells. It is hoped that an increased understanding of this 'preclinical' period will lead to future prevention of type 1 diabetes.

Colman, Harrison and his colleagues decided to work with Bio21:MMIM because it allows a spectrum of many people's research data to be put together, giving 'the big picture'. For the first time, relatives of people with type 1 diabetes, children with a parent with type 1 diabetes and randomly selected school children can be compared directly. They wanted to compare data on a large scale, and as Colman states, 'Bio21:MMIM has given us this opportunity'. Hopefully, putting all existing research data together will help determine risk factors and thereby help prevent type 1