<table>
<thead>
<tr>
<th>CONTENTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflections From the First Chair</td>
<td>4</td>
</tr>
<tr>
<td>Insights From the New Chair</td>
<td>6</td>
</tr>
<tr>
<td>Bioplatforms Australia Profile</td>
<td>8</td>
</tr>
<tr>
<td>Year in Review</td>
<td>10</td>
</tr>
<tr>
<td>National Framework Datasets</td>
<td>11</td>
</tr>
<tr>
<td>Dataset Progress Report</td>
<td>13</td>
</tr>
<tr>
<td>Bioinformatics Initiatives</td>
<td>18</td>
</tr>
<tr>
<td>Access</td>
<td>20</td>
</tr>
<tr>
<td>Collaboration, Partnerships and Governance</td>
<td>24</td>
</tr>
<tr>
<td>Platform Reports</td>
<td>26</td>
</tr>
<tr>
<td>Genomics</td>
<td>26</td>
</tr>
<tr>
<td>Proteomics</td>
<td>27</td>
</tr>
<tr>
<td>Metabolomics</td>
<td>27</td>
</tr>
<tr>
<td>Bioinformatics</td>
<td>29</td>
</tr>
</tbody>
</table>
Innovation has long been recognised as a key to increased productivity and tangible economic benefits. While there is little to debate about the need to enable high quality science, it is much more challenging to establish suitable funding models, investment strategies and cultural frameworks that foster the scientific capabilities that are most valued and beneficial to the Australian community.

As I step down from my six year term as the first Chair of Bioplatforms Australia, I am able to reflect on our many achievements in advancing Australian expertise in biomolecular research. Since our foundation in 2007, Bioplatforms Australia has consolidated, expanded and synchronised the platform technologies that can deliver a systems biology approach. Existing genomics, proteomics and bioinformatics expertise has been significantly boosted by the state-of-the-art infrastructure investments we have delivered. Likewise, brand new capabilities in metabolomics have been established and are increasingly being integrated into many ‘omics projects.

A creditable record in targeted infrastructure purchases and new research expertise has been complemented by initiatives that are enabling a truly holistic systems biology capability. To promote greater cross-discipline collaboration and technology integration we have been funding projects centred on shared infrastructure, collective resources and multi-disciplinary expertise.

Since 2010, we have instigated six separate projects to generate the large-scale quantitative data that are vital for modern day biological research but beyond the capacity of any one lab to create. By funding ‘omic datasets and offering them as a public resource we can support research endeavours of particular importance to Australians. The attraction of building and accessing valuable data assets has harnessed the collective support of leading scientists, research organisations, industry specialists and government resources which leverages existing expertise and ensures the datasets will be meaningful.

New Paradigms of Collaboration
The success of the dataset program is particularly pleasing. In only a short time our dataset collaborations have started to overcome the institutional, disciplinary and geographic boundaries that prevent a systems perspective. Upon reflection, I can comfortably claim that this program has delivered a powerful catalyst in bringing research and industry experts together around specific scientific challenges and, as hoped, they are achieving world class research results. We believe this collaborative model is breaking new ground. It has certainly attracted international interest, not just for the technical challenges associated with generating large-scale, cross-discipline data, or for our approach to developing ‘soft infrastructure’, but as much for the success in stimulating the cultural and structural environment needed for cross-collaboration.

Dataset Success
Wine was our first focus and used to demonstrate the benefits of systems biology. It has resulted in numerous scientific papers and an international collaboration to map the chardonnay genome. A project on wheat followed and was the first to employ our model of broad national collaboration that is unprecedented in Australian science. Our Wheat and Pathogenomics consortiums have delivered DNA sequencing and other ‘omics investigations for several wheat varieties and pathogens which are now publically accessible. The pathogenomics work has been especially applauded for its success in consolidating and building on existing research.

Our other dataset projects are supporting
efforts to map the microbial diversity of the Australian continent; identify genetic mutations related to melanoma cancer; and investigate the resilience of coral to aid climate change protection for the Great Barrier Reef. All these projects have recorded exciting progress over the last year.

During 2013, we were pleased to announce our latest project to support research on Australia’s iconic koala with diverse genomics based datasets. Already, the Koala Consortium have become the first scientists in the world to complete the koala genome and we are delighted to have enabled this achievement.

Growing Momentum in Bioinformatics
The dataset program also encompasses goals to build Australian bioinformatics capabilities. Each dataset throws up technical challenges around data integration, harmonisation, management and accessibility. The new tools and solutions prompted by the datasets are the same as those raised by modern biology in general. Our melanoma dataset project is critically dependent on advanced bioinformatics to analyse and handle the 200+ terabytes of data generated so far. While it is one of the largest genome-based studies in Australia, it is typical in its need for tailored bioinformatics solutions and is building valuable experience and expertise.

To improve data analysis capabilities, we have channelled considerable effort into initiatives to make bioinformatics skills, tools and workflows more accessible. Through tailored training schemes and support for portals, web interfaces and eResearch capabilities, we are ensuring that Australia’s growing data resources are translated to innovative outcomes. A tangible achievement has been the roll out of a next-generation sequencing (NGS) workshop pitched specifically at providing hands-on bioinformatics training to biologists. This workshop was purposefully designed for the Australian research community and drew on the internationally recognised training expertise of the EMBL-European Bioinformatics Institute. The courses conducted in the last year have received outstanding feedback on their usefulness and relevance.

I am also pleased with the new industry connections we are building. The Australian Bioinformatics Network is the fruit of our partnership with EMBL Australia and CSIRO. It has reported great progress in supporting Australia’s emerging bioinformatics sector and pleasing membership growth over its first year.

Acknowledgements
Bioplatforms Australia’s growing track record in new scientific infrastructure and initiatives to foster collaboration and technology integration is a direct reflection of the extraordinary effort and commitment of our people and partners over the past six years.

Our platform convenors and facility leaders have been the backbone of our network ensuring Australian scientists have access to first class research expertise and state-of-the-art infrastructure. I have appreciated their on-going commitment and thank them for all they have done to translate our objectives to material results. I wish to note the substantial contributions of our convenors Dr Sue Forrest (Genomics), Dr Jeff Gorman (Proteomics), Dr Tony Bacic (Metabolomics) and Dr Matthew Bellgard (Bioinformatics), as members of the Bioplatforms Australia Executive Committee, which is chaired by General Manager and Secretary, Andrew Gilbert.

I wish to make particular note of the contribution Andrew Gilbert has made in leading the development of Bioplatforms Australia from its infancy, as the then only full time executive, to the organisation it is today and the great progress on which I have been most honoured to report. In many achievements, he has been very ably supported by Anna Fitzgerald and Catherine Shang, who make a small but effective team that consistently works hard, both collaboratively and creatively with our platforms and partners, to implement our strategic programs. Their energy and exceptional dedication are tremendous assets to the organisation and to the Bioplatforms Australia Board.

Lastly but not least, I thank my fellow founding directors, Dr Les Trudzik and Dr John de Jersey, and also Dr Sue Meek, who joined the Board last year, for their personal support and strategic contribution in directing Bioplatforms Australia’s considerable advancement over my term as Chairman.

I am pleased that Les has agreed to take the reins from August 2013 upon my retirement. I have complete confidence in his ability to steward Bioplatforms Australia to continued success in fostering scientific collaboration both within Australia and internationally.

Mr John Grant AM
Since its inception in 2007, Bioplatforms Australia has continued to evolve its role in the research landscape and is now well positioned with much to offer for the future. I am delighted to be appointed Chair and will continue to deepen our support for bioscience research by delivering the infrastructure, tools and strategies that foster innovation and maintain Australia’s first rate reputation in science.

With much owing to the vision and leadership of the previous Chair, John Grant, Bioplatforms Australia now offers a world class infrastructure network that is being leveraged for pioneering multi-disciplinary research. Our challenge now is to capitalise on this strong foundation to ensure that the opportunities presented by the genomics generation and systems biology can be readily applied to our nation’s most important research challenges.

Genomics research is a vital centre piece of Australian life science capabilities. Interwoven advances in genetics, high throughput technologies and bioinformatics are allowing unprecedented insight into the functioning of an organism with profound potential for every field of life science. Next-generation sequencing in particular is dramatically changing biological research. We are now able to generate genomic maps, sequences and other data faster and more cheaply than ever before which is, in turn, fuelling considerable demand for high-throughput services. These ‘big’ and rapidly growing data resources are invaluable for genomics-based research. A key limitation, however, is ready and affordable access to the computational power and expertise required to analyse these data. Likewise, demands for multi-technology integration and data sharing necessitate a different culture of collaboration along with practical solutions to multiplex workflows and achieve data harmonisation. The consequences of such rapid technology advances are exciting but demanding. They require astute action and a deep awareness of international trends to address bottlenecks and take full advantage of new capabilities.

Bioplatforms Australia’s most fundamental role is to ensure Australian scientists have access to the full range of biomolecular technologies and tools that enable cutting edge research, especially where it can promote systems biology approaches and analysis.

At the same time we need to remain proactive in tackling the numerous opportunities presented by the ‘genomics age’ and do all we can to support and sustain Australia’s momentum in life science research. Important in doing this is to engage with scientists, policy makers and funding agencies to ensure our strategies are relevant and effective.
Future Priorities

Our priorities over the next few years will be in a number of key areas:

• Strengthening partner and stakeholder networks that continue to catalyse new research collaborations.

• Supporting further development of the cultural and practical pre-requisites for greater levels of collaboration and integration across multiple fronts, including the different ‘omics platform technologies; scientists from different disciplines; and academia with industry.

• Fostering deeper engagement with Australia’s peak research organisations to better serve the scientific community and maintain an acute awareness of Australian infrastructure requirements, technology trends and emerging issues.

• Forging new links with international organisations to broaden our insight on international trends and identify beneficial collaborative opportunities.

This Annual Report shows the breadth of our endeavour and achievements to date, and highlights the many strengths of our staff and partners. I very much look forward to Bioplatforms Australia building on this strong base.

Dr Les Trudzik

Bioplatforms Australia’s most fundamental role is to ensure Australian scientists have access to the full range of biomolecular technologies and tools that enable cutting edge research, especially where it can promote systems biology approaches and analysis.

Access for cutting edge research
Bioplatforms Australia builds Australian research capability and capacity in the fields of genomics, proteomics, metabolomics and bioinformatics.

Our national network of leading universities, research institutes and bioscience companies provides access to world class ‘omics expertise and state-of-the-art infrastructure through contracted services and research collaborations.

Bioplatforms Australia leverages government funding to develop and implement investment strategies that boost Australian bioscience capabilities relevant to agriculture, biomedicine and the environment.

We endeavour to accelerate systems biology in Australia through projects and initiatives that facilitate multi-discipline integration. We also strive to foster greater scientific collaboration among researchers and between researchers and industry to tackle research challenges of national importance to Australia.

**Mission**

Bioplatforms Australia seeks to catalyse world class bioscience through access to integrated biomolecular research capability.

**Goals**

- Ensure broad access to state-of-the-art biomolecular research capability.
- Facilitate greater multi-discipline integration in Life Science research.
- Facilitate a collaborative approach to addressing national challenges of economic and social significance.
- Advance world class research and innovation in Australia.

**Capabilities**

Bioplatforms Australia offers a broad range of complementary ‘omics technologies relevant to food, agriculture, biomedicine and environmental research. Users can access services from just one node or multiple network partners for more complex projects.

**Network**

The Bioplatforms Australia network is organised into four technology platforms – genomics, proteomics, metabolomics and bioinformatics. Services are offered through 17 separate facilities located around the nation with expert capabilities that can be readily accessed by the research community.

The Bioplatforms Australia network is organised into four technology platforms – genomics, proteomics, metabolomics and bioinformatics. Services are offered through 17 separate facilities located around the nation with expert capabilities.
BIOPLATFORMS AUSTRALIA PROFILE

Mission
Bioplatforms Australia seeks to catalyse world class bioscience research

Biomolecular research capability
• infrastructure
• expertise
• capacity

Access to ‘omics technology
• contract services
• collaborations
• awards

Enable integration
• multi-platform technologies
• national collaborative projects
• bioinformatics

Increase scientific collaboration
• among researchers
• between researchers & industry

Bioinformatics
Solutions for data acquisition, interpretation and analysis

Proteomics
Protein structure and function

• protein separation
• protein chemistry
• mass spectrometry
• monoclonal antibody development

Metabolomics
Small molecule analysis

• sample preparation
• metabolite profiling
• mass spectrometry
• lipidsomics

Genomics
Gene discovery and genome analysis

• high throughput sequencing
• transcript analysis
• epigenetics
• bioinformatics

Bioinformatics Resource Australia - EMBL
Centre for Comparative Genomics, Murdoch University

Systems Biology Initiative, University of NSW

Queensland Institute of Medical Research
Australian Proteome Analysis Facility

Monash Biomedical Proteomics Facility
Monash Antibody Technology Facility

TGR BioSciences

Metabolomics Laboratory, Murdoch University
School of Botany, University of Melbourne
Bio21 Institute, University of Melbourne

Australian Wine Research Institute
ARC Centre of Excellence in Plant Energy Biology, University of WA

Australian Institute of Bioengineering and Nanotechnology, University of Queensland

Adelaide Proteomics Centre

Western Australian Institute for Medical Research, Proteomics International

The Ramaciotti Centre for Gene Function Analysis, University of NSW

The John Curtin School of Medical Research, ANU

Australian Genome Research Facility

The John Curtin School of Medical Research, ANU
Bioplatforms Australia endeavours revolve around our constant commitment to building Australian biomolecular research capabilities and ensuring they are fully exploited to enable world class innovation.

All our programs incorporate four key themes:

• Building research CAPABILITY
• Fostering scientific COLLABORATION and a systems based approach
• Increasing ACCESS to world class scientific infrastructure
• Facilitating technology INTEGRATION to enable new biological insights

These themes have guided our activities and achievements over the year. Priorities have been:

• Our Framework Dataset Program to provide pivotal data resources in response to national challenges while also facilitating the collaborative culture and cross-discipline approach essential to systems-based analysis.
• Bioinformatics initiatives that seek to address bottlenecks and challenges presented by the data driven nature of modern biology.
• New collaborations that magnify our own efforts to build research capabilities in Australia and deliver highly valued scientific outcomes.

Bioplatforms Australia endeavours revolve around our constant commitment to building Australian biomolecular research capabilities and ensuring they are fully exploited to enable world class innovation.
DATASET PROGRAM OBJECTIVES

- Build ‘soft infrastructure’ (data assets) for Australian life scientists
- Facilitate and advance the emerging field of systems biology
- Catalyse scientific collaboration and partnership around national research challenges
- Maximise the impact of Australia’s investments in ‘omics infrastructure
- Provide a focal point for advancing Australia’s bioinformatics community
- Attract co-investment from philanthropic, commercial, scientific and government partners.

Bioplatforms Australia’s National Framework Dataset Program provides critical biomolecular datasets to help tackle scientific challenges faced by Australian researchers in biomedicine, agricultural or environment science.

Since 2010, we have instigated six separate projects to generate the large-scale ‘omics data demanded by research projects that discover new ideas or solve problems of particular importance to Australians.

By funding the datasets and offering them as a public resource, Bioplatforms Australia can enable a broad scope of research endeavours.

The program has also seeded a new collaborative model that draws on the collective expertise and resources of leading scientists, research organisations, industry specialists and government bodies.

While each of our dataset projects offer significant strategic value to national research efforts, all have resounding potential to benefit from an integrated approach. We are particularly pleased that the program has stimulated deep engagement between Australia’s internationally recognised experts and broader elements of the research community. Furthermore, the nature and extent of the collaborative effort is cementing cultural acceptance of shared infrastructure and greater technology integration – two elements of critical importance to future research collaboration and partnerships.

The dataset program extends to goals that seek to build Australia’s bioinformatics expertise. Each dataset must navigate challenges to manage, integrate and analyse data. New tools and solutions prompted by the datasets both utilise and expand Australian bioinformatics capabilities.

Bioplatforms Australia’s National Framework Dataset Program provides critical biomolecular datasets to help tackle scientific challenges faced by Australian researchers in biomedicine, agricultural or environment science.
BIOME OF AUSTRALIAN SOIL ENVIRONMENTS

Define Australian soil communities to understand ecological processes and land use response

- Map of Australian soil biodiversity
- Metagenomics sequencing of soil microbial communities
- Comparative analysis of soil for environmental, agricultural and discovery applications

KOALA

Protect koalas from debilitating disease and gain genetic insights into their unique traits

- Sequencing the complete koala genome and transcriptome analysis
- Genomic insights into koala infection and immunity
- Enabling koala population health and response

MELANOMA

Identify genetic mutations that lead to melanoma cancer

- Genome sequencing of melanoma tumours
- Multi-disciplinary analysis of melanoma cell function and mutation
- Tailored bioinformatics solutions

WHEAT

Greater food security through increased yields and defence against disease

- Contribute to wheat reference genome - Chromosome 7A
- Genome sequencing of key Australian breeding varieties
- Multi-disciplinary analysis of pathogen-wheat interaction

WINE

Enhance wine characteristics through improved fermentation processes and gain genetic insights to chardonnay varieties

- Systems biology analysis of yeast in wine fermentation processes
- Genome sequencing of key chardonnay varieties
- Comparative analysis of genetic variations and impact on growth patterns and product differences

GREAT BARRIER REEF

Protect and preserve coral reefs from the impact of rising water temperatures

- Genome sequencing of Great Barrier Reef corals and their symbionts
- Metagenomics sequencing of coral symbiotic organisms
- Analysis of coral biology and potential for climate adaptation and protection

DATASET PROJECTS

NATIONAL FRAMEWORK DATASETS
KOALA

During the year, Bioplatforms Australia launched a new dataset project to sequence the koala genome. This project was instigated with the Australian Museum and Queensland University of Technology (QUT) and resulted in a ‘koala consortium’, which is credited with being the first in the world to complete the draft genome of Australia’s iconic koala.

This landmark mapping achievement offers considerable research benefits. For the first time scientists working on koala biology will have access to their complete genetic blueprint. Until now, only a handful of koala genes have been available in public databases but the full genetic map is expected to reach 20,000 genes. Many unique koala genes have been discovered already such as those related to their remarkable ability to survive on a specialised diet.

The datasets will have broad research applications including questions related to the evolution of koalas being pursued by the Australian Museum and others. They will also help to address the alarming susceptibility of koalas to Chlamydia and Koala retrovirus which are decimating koala populations in NSW and Queensland. The datasets are providing critical information on koala immunity which is facilitating development of a new vaccine.

This wealth of data is also relevant to biologists working to protect other species with similar genetic sequences as the koala. And since the koala datasets will be broadly shared, they are expected to serve an even greater array of future research programs.

Availability of the dataset and the associated collaborative effort has seeded new commitment to koala research. The Koala Genome Consortium currently involves more than 12 scientists, veterinarians and bioinformaticians, including experts from our genomics node at the Ramaciotti Centre. It is now seeking greater researcher and funding support to supplement the sequencing data and generate more extensive applications for the data.

WHEAT

This first dataset project was launched in 2010 to support important research efforts into increasing crop yields and improving resistance to disease and challenging climate conditions. The project is now reaching its final stages with a number of the datasets already available for public access.

Our large consortium of wheat collaborators comprise leading researchers and breeders from the CSIRO, universities, industry bodies and Bioplatforms Australia nodes. Two working groups were established; the first focused on wheat variety sequencing, and the second centred around wheat pathogenomics. Data exchange and project integration between the working groups occurred during forums and workshops held throughout the year.
Tremendous progress has been made towards completing the three core components of our wheat project:

**Chromosome 7A**

The International Wheat Genome Sequencing Consortium (IWGSC) is building a wheat reference genome through a shared effort in sequencing the 21 chromosomes of wheat. Chromosome 7A constitutes Australia’s contribution to this project. Access to a reference genome provides an invaluable resource for identifying genetic differences between wheat varieties. The complete wheat genome is five times larger than the human genome and presents a huge assembly challenge for scientists. Chromosome 7A alone represents a complex undertaking given its substantial size and high level of repetitive sequences.

Australia’s 7A contribution is a strategic choice as it contains important genes that influence yield quality and disease resistance and delivers a neat fit with current objectives in wheat research. Chromosome 7A sequencing has been completed by our node at the Australian Genome Research Facility and genome assembly work is underway at our bioinformatics node at the Centre for Comparative Genomics. Preliminary analysis has so far identified specific genes relevant to grain yield and grain size.

**Genetic mapping of key Australian wheat varieties**

The wheat datasets include genome sequences of important Australian cultivars to enable meaningful genetic comparisons of different wheat varieties and help researchers identify the characteristics that impact yield. Full sequencing and assembly for 16 key varieties has now been completed. Cultivar selections were based on their productivity characteristics in different environments as well as their economic importance. This data has been publically available since last year and shared through two scientific journals published in 2012/13.

**Wheat interaction with pathogens and pests**

As disease represents a major constraint to wheat production both globally and in Australia, collaborators formed the wheat pathogenomics working group to develop a genetic understanding of pathogen interactions with wheat.

The rapid rate of wheat pathogen evolution and the introduction of new risks linked to changing environmental conditions has made disease protection and resistance a critical and on-going endeavour. Genomic-based data on 22 different pathogens that cause stripe rust, stem rust, tan spot, glume blotch, septoria leaf blotch and crown rot/head blight have been generated and analysed on both an individual as well as comparative basis.

Proteomic and metabolomics datasets are almost complete and should be finalised by the end of 2013.

These data are already enabling new insights on the plasticity of pathogen genomes and identifying common and specialised mechanisms which can affect disease in wheat. Existing publications on the wheat pathogen work will soon be supplemented with a number of additional manuscripts currently in the review process.

Genomic analysis on the *Fusarium* species is attracting particular interest from the international wheat community.

During the year the wheat pathogenomics work was positively reviewed by Professor Anne Osbourn from the John Innes Centre in the UK. As an eminent researcher and expert in molecular plant pathology her input was both welcome and encouraging. In particular, Professor Osborne complimented the ‘simple but elegant’ approach of using the host (wheat) as the common element to link research on diverse pathogens and infection styles. She suggested that this work offered a proof of concept for other important crops including legumes, barley and canola. She also noted that the multi-platform approach to investigating wheat diseases has
DATASETS PROGRESS REPORT

culminated in an attractive, well-integrated and cohesive body of research that both consolidates and builds on previous work in the area. The positive impact of bringing the wheat disease community together was also commended.

BIOME OF AUSTRALIAN SOIL ENVIRONMENTS (BASE)

The BASE project is providing a unique opportunity to quantify the abundance, diversity and structure of different soil communities in Australia. This data will provide a map of soil microbial diversity and allow wide ranging environmental research such as soil response to land use practices and insights into ecological stability, resilience and remediation.

Although, soil hosts diverse microbial communities that play a critical role in many ecological processes, soil is rarely viewed as a natural resource and its microbial communities are not well quantified in Australia or the rest of the world. To help build predictive models of microbial communities across different soil and environmental conditions, the datasets will combine both biological and physicochemical measurements and include an extensive array of contextual data for each site. This will allow researchers to compare microbial communities from contrasting environments and attempt to relate their structure and function to vegetation composition, climate and environmental impacts. This will also allow deeper analysis of the symbiotic and co-evolutionary relationship between soil communities and plants.

Throughout the year, BASE collaborators from the CSIRO, national parks, government agencies, universities and research and development corporations collected samples from over 600 sights across a number of national parks and sites from Antarctica, the Terrestrial Ecosystems Research Network (TERN) as well as minerology sites. A small pilot study to analyse data from 75 different sites was conducted to establish various parameters before proceeding to more detailed genomics analysis. This early data is also important for finalising methods development and sequencing protocols as well as defining project milestones that address the disparate research partners involved and their wide ranging research interests.

The soil datasets will align and support the Earth Microbiome Project which aims to examine microbial communities around the globe but will immediately serve current revegetation projects in Gondawana and Christmas Island and a new study investigating the impact of human activities on polar microbiomes in Antarctica.

‘Diversity models’ across different soil types and geographic space utilising data assembled so far are underway and will be refined as deeper genomics analysis is completed. A national on-line database is also being established to share sequencing and environmental contextual data.
DATASETS PROGRESS REPORT

Our genomics node at the Australian Genome Research Facility has been performing the DNA recovery and sequencing work for this project utilising 454 sequencing. Future work will be completed using Illumina instrumentation as recent advances in this high-throughput platform can better accommodate the novel techniques being employed. A bioinformatics pipeline has also been established to manage and analyse the data.

MELANOMA

Australia has the highest incidence of melanoma in the world and the number of cases is increasing. Until recently there have been no effective therapies for treating melanoma cancer, but with greater knowledge of the genetic basis of the disease, researchers are confident that new treatments and better diagnosis will result.

The melanoma dataset program aims to identify the genetic mutations that cause melanoma cancer. Over 500 melanoma tumours are being sequenced to compare mutant cells with ‘normal’ ones in order to identify common mutations and pinpoint new treatment targets.

This collaborative venture led by the Melanoma Institute Australia involves a large number of research and funding partners. It is currently one of Australia’s largest genome projects. The data being generated are critical for exploring the genetic and biochemical events that can transform cells into a melanoma and will hopefully reveal patterns of mutation and the consequent molecular changes that can occur. As multiple datasets will link various cellular processes, the data will also provide an invaluable resource for other cancer research and biological investigations.

Our genomic nodes at the ANU, Ramaciotti Centre and Australian Genome Research Facility are sequencing 150 of the melanoma samples that have been carefully curated from tissue collections established over the last 20 years. So far sequencing for 75 samples has been completed and bioinformatics analysis has commenced at the John Curtin School of Medical Research. Proteomic and metabolomics analyses will also be undertaken to profile the proteins and other small molecules expressed by the relevant genes. These analyses will provide further clues on the impact of melanoma gene mutations and help identify new drug targets.

The final dataset is expected to include up to 1,000 whole genome sequences from both tumours and normal tissue which represents an enormous amount of data requiring storage in excess of 200 terabytes. Tailored bioinformatics pipelines have been developed in order to extract information from the raw sequencing data and detailed reports on the mutations detected have already allowed researchers to start identifying the most important melanoma variants.

The melanoma dataset program aims to identify the genetic mutations that cause melanoma cancer.
DATASETS PROGRESS REPORT

GREAT BARRIER REEF

Bioplatforms Australia is supporting important research on the Great Barrier Reef by delivering genomic datasets for coral, their algal symbionts and associated microbes. These large-scale datasets will serve a series of research projects seeking to determine the response and resilience of corals to climate change and underpin new solutions for protecting and preserving coral reefs.

Climate change is considered to be the greatest risk facing the Great Barrier Reef with rising temperatures and pH changes placing on-going stress on corals and threatening the long term sustainability of coral reefs and their complex ecosystem.

To provide a scientific response to this problem, the Great Barrier Reef Foundation (GBRF) has established a consortium of experts from Australia and Saudi Arabia to investigate the genetic make-up of corals and their potential for climate adaptation.

The first phase of the research program is generating genomic data on ten important Australian corals and their symbionts which will be used for multiple research objectives. Bioplatforms Australia facilities have commenced DNA sequencing in collaboration with the coral research community. A parallel program to sequence Red Sea corals is also being conducted in Saudi Arabia to allow comparisons between corals in different environments and water temperatures.

Work to date has determined extraction and sequencing protocols and resulted in samples being collected from five coral species located around Orpheus and Herron Islands in the Great Barrier Reef. A pilot program for two of the coral types collected will establish the approach for the remaining eight species yet to be sequenced. Acropora millepora, is a target species and one of only two corals with a completed genome map. Consequently, this species will be one of the first to receive extensive genome based analysis of its microbiome, including related viruses and bacteria.

The research program is being directed by the coral consortium involving experts in marine science, genomics and bioinformatics. Titled ReFuGe 2020 (short for Reef Future Genomics), the consortium meets every six months to share data and track progress. Bioplatforms Australia is an active member of the consortium and responsible for delivering genetic sequencing data and associated ‘omics investigations. The program is also receiving corporate sponsorship from Life Technologies and Rio Tinto.

Climate change is considered to be the greatest risk facing the Great Barrier Reef with rising temperatures and pH changes placing on-going stress on corals and threatening the long term sustainability of coral reefs and their complex ecosystem.
The need to manage and analyse large datasets remains a predominant driver of the bioinformatics boom. Australian researchers are still struggling to fully exploit growing data resources, however, improved training, greater community support and new eResearch infrastructure is expanding capabilities and fuelling research advances.

During 2012/13 Bioplatforms Australia maintained a strong focus on building bioinformatics skills and networks for Australian life science researchers. We rolled out next-generation sequencing (NGS) workshops to help build the data analysis skills of laboratory scientists. The course has been well received and immediately regarded by participants as ‘essential learning’. At the same time, we are still facilitating greater connections and support for the Australian bioinformatics community through on-going funding and involvement with the Australian Bioinformatics Network.

We continue to fund specialist expertise with ‘embedded’ bioinformaticians at our nodes and provide support for new analytical tools, pipelines, databases and other infrastructure to facilitate data integration, data sharing and data management. These efforts are complemented by our collaborations with NeCTAR and the services and data offered through the EBI mirror and computing infrastructure of the Australian Bioinformatics Resource-EMBL (BRAEMBL).

Next-generation Sequencing Workshops

The bottleneck in managing and analysing the abundance of NGS data has been intensifying with the lack of bioinformatics training courses and scarcity of skilled trainers available in Australia.

In 2012, Bioplatforms Australia collaborated with the European Bioinformatics Institute (EBI) and CSIRO to develop a workshop that could provide hands-on NGS tutorials for participants. Eight highly experienced bioinformaticians were recruited for instructor training at the world renowned EMBL-European Bioinformatics Institute (EMBL-EBI) and worked with training specialists to design and deliver a tailored course for Australian researchers.

This new NGS workshop was run in Sydney, Melbourne, Canberra, Brisbane and Perth during the last year.

Outstanding feedback was received on both the content and delivery of the course, with many participants indicating that they would welcome additional training.

The course offers a significant ‘hands on’ time which is enabled through the new NeCTAR research cloud. This new capability provides remote access to infrastructure capable of the intensive computation requirements associated with each training exercise.

In response to the success and feedback of the NGS workshop, we are now developing additional bioinformatics training modules for Australian researchers. A metagenomics course is earmarked for next year and will again involve a partnership with EMBL-EBI for training and course design. The workshops have been over-subscribed by more than double the current capacity and Bioplatforms Australia is working closely with the EBI to meet this demand in the coming years.

Australian Bioinformatics Network

The Australian Bioinformatics Network (ABN) was launched in mid-2012 as an independent body to build greater links between researcher needs and bioinformatics capabilities.

The ABN is a jointly funded initiative of Bioplatforms Australia, EMBL Australia and the CSIRO. Dr David Lovell, Bioinformatics and Analytics Leader for CSIRO’s Transformational Biology Initiative was appointed as Director in July 2012. Dr Benita Vincent joined as an Executive Officer in February 2013.
During the year, the Network’s first priority was to build a website that could act as a forum and national information hub for the Australian bioinformatics community. The website is attracting interest and traffic has been steadily increasing along with ABN membership which reached 330 registrations to July 2013.

The ABN also supports training and community initiatives including the NGS training workshops previously outlined, as well as bioinformatics symposiums and professional events. As a member of the newly formed Global Organisation for Bioinformatics Learning, Education and Training (GOBLET), the ABN anticipates that it will expand Australia’s international links and opportunities to access world-wide bioinformatics training initiatives. To build stronger Australian bioinformatics networks and professional development, the ABN now offers ‘Connection Grants’. The grants are awarded to fund speakers or sponsor meetings and have so far facilitated bioinformatics workshops on wheat, a talk by an international expert in computational biology and two Software Carpentry boot camps.

The ABN anticipates that it will expand Australia’s international links and opportunities to access world-wide bioinformatics training initiatives.
ACCESS AND USE

Bioplatforms Australia collects data from its major nodes annually in order to monitor access trends and patterns of demand. Our baseline data was collated in 2006, the year prior to Bioplatforms Australia being established as a national infrastructure network and implementation of the NCRIS funding strategy. Statistics for 2012/13 represents the sixth full year of a shared infrastructure approach for biomolecular research and nationally co-ordinated investment.

Client Growth Continues

The upward trend in accessing ‘omics infrastructure and associated revenue has continued in 2012/13. The total number of fee-for-service clients has trebled since Bioplatforms Australia was established in 2007 with a ten percent increase recorded in the last 12 months. Researchers are now spending more than $20 million per annum on ‘omics services offered by our research network.

A ‘client’ is defined as a researcher or research group who access a service or services for a specific research project. As shown in Figure 1, there are now more than 4000 clients accessing biomolecular research services from our network. We can also report that 88 per cent of these clients are external to host institutions.

When analysed by platform, genomics services continue to show the greatest rate of growth and account for the majority of services provided. Figure 2 depicts the dominance of genomics within the mix of available services over the last five years. This record is unsurprising given the broad relevance of genomic investigations to life science research and is likely to continue as growing bioinformatics capabilities and cloud computing infrastructure further improve the ability to analyse data outputs from next-generation sequencing.

Clients and Revenue by Research Sector

The ratio of access by universities, publicly funded research organisations and commercial clients is similar to last year as shown in Figure 3. In 2012/13 universities were still the biggest users of ‘omics infrastructure and services accounting for 58 per cent of clients. 31 per cent of users derive from publicly funded research organisations and 11 per cent are from the commercial sector.
As shown in Figure 4, total recorded revenue for 'omics services increased slightly above the high of last year reaching $23.6 million in 2012/13. This is more than double the amount recorded before Bioplatforms Australia was established six years ago. Over this period, more than $100 million has been spent on research services supported by Bioplatforms Australia.

Commercial clients have a much higher average spend than those from other sectors and constitute about a third of all recorded revenue. Spending by this sector slightly increased over the last year to 28 per cent of the total. The public research sector is responsible for 39 per cent of expenditure in 2012/13 while universities contributed 33 per cent of revenue.

Access by Industry Sector

Infrastructure access by biomedical/diagnostics researchers remains the most significant accounting for 71 per cent of clients and 77 per cent of revenue. Figure 5 depicts the relevance of genomics for investigating disease and medical research and the funding allocated to this sector is an obvious driver for this trend. The increasing interest of the agricultural/wine/food sector is reflected in a growing number of projects. This sector recorded 23 per cent of clients and 18 per cent of revenue.

Services provided to the environment sector were fairly static at seven per cent of the total, however, an overall decline in spending by this sector was recorded.
INFRASTRUCTURE & TECHNOLOGY INVESTMENTS

Next-generation sequencing is still the main genomics investment focus in order to keep up with continuing demand. A Fluidigm BioMark HD system, and a C1 cell auto-prep unit, have been installed at the Ramaciotti Centre for Gene Function Analysis to extend capabilities in genomics analysis. This system is sensitive enough to study gene expression down to the single-cell level.

Additional mass spectrometry instrumentation placed at the Australian Proteome Analysis Facility will deliver new capability for high mass accuracy and high resolution peptide measurements. A Q Exactive benchtop Orbitrap is providing additional flexibility for mass analysis and is particularly useful for multiplexed protein quantitation experiments. An upgraded BioPlex system for bead-based ELISA has also been installed to meet increasing user demand.

Metabolomics technology investments have filled some identified gaps in capability and capacity. New services are now available as a result of investment in a high-end tissue imaging mass spectrometer from Bruker. This instrument can determine the spatial distribution of many small molecules as well as lipids across various biological tissues.

Bioinformatics investments include expert personnel, infrastructure upgrades, training and industry development initiatives. Embedded bioinformatics remains central to our holistic approach to building life science research capability.

During the year, the Centre for Comparative Genomics installed peta-scale data storage to support the framework datasets and other initiatives. This capability has been closely aligned with the national Research Data Storage Initiative (RDSI) which aims to support retention and integration of nationally significant data assets.

PROMOTIONAL ACTIVITIES

Our promotional activities seek to build awareness of Australian ‘omics capabilities and services to ensure they are fully exploited by Australian researchers. During the year we worked with a public relations company to more broadly convey the research outcomes and successful collaborations formed through our framework dataset initiative. We also continued our advocacy role and widely consulted with ‘omics researchers, service users and investment stakeholders.

Our network partners are also very active in showcasing their research capabilities and technical expertise. During the year nodes published in scientific journals, attended scientific conferences, gave presentations and showed posters, exhibited at relevant conferences, and ran and participated in user group workshops.

We complemented the promotional activities of our network partners by sponsoring the following:

- Metabolomics Symposium at the 18th Australasian Proteomics Society conference at Lorne.
- Biodiversity Genomics Conference at the Australian National University.
- Scientific workshops to share data and facilitate collaboration on the wheat framework data initiative (attended by approximately 50 domain scientists).
- Planning workshops to develop and execute the Great Barrier Reef (ReFuGe) framework data program, attended by leading Australian and international scientists.
- A $10,000 prize voucher awarded to an early career scientist at the 2013 ComBio Conference.
### Table 1: Infrastructure Purchases

#### Genomics

<table>
<thead>
<tr>
<th>Institution</th>
<th>Equipment Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australian Genome Research Facility (AGRF)</td>
<td>- Illumina HiSeq sequencing upgrades and ancillary equipment</td>
</tr>
<tr>
<td></td>
<td>- Sanger Sequencing equipment</td>
</tr>
<tr>
<td></td>
<td>- Sample preparation equipment</td>
</tr>
<tr>
<td></td>
<td>- Ancillary laboratory equipment</td>
</tr>
<tr>
<td>Ramaciotti Centre for Gene Function Analysis</td>
<td>- Fluid BioMark HD System</td>
</tr>
</tbody>
</table>

#### Proteomics

<table>
<thead>
<tr>
<th>Institution</th>
<th>Equipment Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australian Proteome Analysis Facility (APAF)</td>
<td>- Thermo Scientific Orbitrap ELITE Mass Spectrometer with Electron Transfer Dissociation</td>
</tr>
<tr>
<td></td>
<td>- QExactive Orbitrap</td>
</tr>
<tr>
<td></td>
<td>- BioPlex ELISA</td>
</tr>
<tr>
<td>Queensland Institute of Medical Research</td>
<td>- Orbitrap Fusion instruments with ETD</td>
</tr>
<tr>
<td></td>
<td>- AB Sciex Triple TOF 5600</td>
</tr>
<tr>
<td></td>
<td>- Two nano UHPLC pumping systems for existing LTQ-Velos Pro – Orbitrap Mass Spectrometers</td>
</tr>
</tbody>
</table>

#### Metabolomics

<table>
<thead>
<tr>
<th>Institution</th>
<th>Equipment Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australian Wine Research Institute</td>
<td>- Bruker GC-APCI</td>
</tr>
<tr>
<td>University of Melbourne</td>
<td>- Gas Chromatography Triple-Quadrupole Mass Spectrometer</td>
</tr>
<tr>
<td></td>
<td>- Imaging Mass Spectrometer</td>
</tr>
<tr>
<td></td>
<td>- Bruker FITR Mass Spectrometer</td>
</tr>
<tr>
<td></td>
<td>- Two Liquid Chromatography Triple-Quadrupole Mass Spectrometers</td>
</tr>
<tr>
<td></td>
<td>- Quadrupole TOF Mass Spectrometer and related accessories</td>
</tr>
<tr>
<td>University of Queensland</td>
<td>- Quadrupole TOF Mass Spectrometer and related accessories</td>
</tr>
</tbody>
</table>

### Table 2: Bioinformatics Investments for 2012/13

<table>
<thead>
<tr>
<th>Institution</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australian Genome Research Facility</td>
<td>- On site (embedded) bioinformaticians for integrated project support</td>
</tr>
<tr>
<td>Ramaciotti Centre for Gene Function Analysis</td>
<td></td>
</tr>
<tr>
<td>Australian National University</td>
<td></td>
</tr>
<tr>
<td>Australian Proteome Analysis Facility (APAF)</td>
<td></td>
</tr>
<tr>
<td>Monash Biomedical Proteomics Facility</td>
<td></td>
</tr>
<tr>
<td>Queensland Institute of Medical Research</td>
<td></td>
</tr>
<tr>
<td>WA Institute for Medical Research</td>
<td></td>
</tr>
<tr>
<td>Australian Wine Research Institute</td>
<td></td>
</tr>
<tr>
<td>Centre for Comparative Genomics, Murdoch University</td>
<td>- eResearch and Data Management Infrastructure</td>
</tr>
<tr>
<td>Australia Bioinformatics Resource-EMBL</td>
<td>- Data Curation and Repository</td>
</tr>
<tr>
<td>Systems Biology Initiative, University of NSW</td>
<td>- Biological Analyses</td>
</tr>
<tr>
<td>Scientific Community</td>
<td>- Next-generation Sequencing Training Workshops attended by over 200 participants in Perth, Adelaide, Melbourne, Brisbane and Canberra.</td>
</tr>
<tr>
<td></td>
<td>- Sponsorship of 2012 BioInfo Summer Bioinformatics Symposium and the 2012 Winter School in Mathematical and Computational Biology.</td>
</tr>
<tr>
<td>Australian Bioinformatics Network</td>
<td>- Joint funding of Director’s salary with CSIRO and EMBL Australia.</td>
</tr>
</tbody>
</table>
Bioplatforms Australia works hard to foster scientific collaborations and partnerships that aim to address scientific challenges, build Australian research capabilities and deliver world class results. The Framework Dataset program is a key tactic in this mission but is only one element of our effort to more deeply engage with the scientific community.

Over the last year we have been establishing new collaborations that can accelerate high calibre research projects by leveraging innovative biomolecular research capabilities. ARC Centres of Excellence, Joint Centres of Excellence, NHMRC Programs, CRCs and CSIRO flag ship programs undertake pioneering and potentially transformational research that advance Australia's international standing and frequently lead to new capabilities and knowledge. Bioplatforms Australia is in a position to support leading research programs with the objective of linking existing Australian research strengths with aligned interdisciplinary and collaborative research programs.

**New collaborations established during 2012/13:**

**Baker IDI – Cardiovascular Disease**

Cardiovascular disease is responsible for more deaths in Australia than any other disease and costs the Australian economy more than $6 billion a year. Lipid profiling has been shown to have considerable value in detecting patients at risk of cardiovascular disease. The Baker Heart & Diabetes Institute is one of Australia's leading medical research institutes. It possesses particular expertise in lipid profiling while Bioplatforms Australia’s metabolomics nodes have established state-of-the-art analytical and bioinformatics platforms for lipodomics/metabolomics. This collaboration brings together unique expertise in cardiovascular disease, access to clinical samples and lipid analysis to identify new biomarkers for the diagnosis and prognosis of cardiovascular disease. The project also offers potential to leverage public funding with philanthropic contributions and will advance technology that is widely applicable to other human diseases.

**Centre of Biodiversity Analysis**

The Centre for Biodiversity Analysis is a collaboration between the Australian National University and CSIRO which employs new capabilities in evolutionary genomics, bioinformatics and ecological modelling for the discovery, understanding and conservation of Australia’s unique biodiversity. The research collaboration will provide next-generation sequencing on Australian terrestrial vertebrate utilising preserved specimens held by Australian museums as well as modern samples. This project hopes to progress investigations into the genetic basis of climate adaptation and deepen knowledge of Australian biodiversity. The research collaboration will provide next-generation sequencing on Australian terrestrial vertebrate utilising preserved specimens held by Australian museums as well as modern samples. This project hopes to progress investigations into the genetic basis of climate adaptation and deepen knowledge of Australian biodiversity. The results will be applied immediately in several ARC-funded projects and projects across multiple universities and museums. In addition, parallel proposals within CSIRO seek to apply these same technologies to insect and plant collections. The long-term goal is to foster development of genetic resources for all major Australian vertebrates as well as selected plant and invertebrate taxa. The project will require substantial genomic resources and investment in bioinformatics pipelines.

**Centre of Excellence in Human Proteomics**

Bioplatforms Australia is providing funding support in a strategic proposal to establish an Australian Centre of Excellence in Human Proteomics. If successful, the proposal will bring together 16 research and industry organisations to capitalise on and grow Australia’s research excellence in proteomics. The main research objective of the collaboration will be to map the human proteome as it exists in a healthy state. The project would leverage and accelerate advances in proteomics technology, e-Research and bioinformatics capabilities and generate ‘big data’ with on-going research relevance to the bioscience, pharmaceutical, and health care sectors. The project interfaces with the international Human Proteome Project (HPP) and involves significant international collaboration. This initiative would have considerable impact in maintaining Australia’s leadership and international reputation in this field.

**Centre of Excellence in Nanoscale BioPhotonics**

A proposal to establish a Centre of Excellence in Nanoscale BioPhotonics has been established by the University of Adelaide with support from Bioplatforms Australia. The research focuses on the nanoscale control of light-matter interactions for
understanding chemical, biological and physical processes within complex and dynamic nanoenvironments in living organisms. Research outcomes are expected to understand how single cells react to and communicate with their microenvironment in real-time and will underpin a new generation of devices capable of probing the response of cells within individuals to environmental conditions or treatment.

This interdisciplinary program will incorporate ‘omics technology and capitalises on the strong standing of Australian optics research, biosensing, optical and reproductive sciences and high-end photonic technologies. The Centre will also strengthen international research connections established in this field in the UK, US, and China and secure high international visibility.

Centre of Excellence in Microbial Diversity, Evolution and Control

The proposal to create an ARC Centre of Excellence in Microbial Diversity, Evolution and Control (CMDEC) seeks to explore the causes and consequences of microbial diversity relevant to humans, animals and crops and play a leading role in the response to novel infectious diseases including improved methods for infectious disease control. The research plan takes a multidisciplinary approach to address three core challenges:

- **DIVERSITY**: To reveal the extent and patterns of microbial genetic and phenotypic diversity in broad-ranging host populations in Australia, and the factors that shape this diversity
- **EVOLUTION**: To determine the mechanisms, both ecological and genetic, that allow microbes to emerge, adapt and spread in novel hosts, and to evade control programs.
- **CONTROL**: To develop multidisciplinary and evolutionarily-informed methods that will assist in disease prevention and control.

Outcomes of an enhanced understanding of the epidemiology and evolution of key infectious diseases will lead to new avenues of protection and control. It will assist in the conservation of indigenous wildlife species that are under threat from invasive infections and increase awareness of emerging threats and enhance pandemic preparedness.

GoVernance

Bioplatforms Australia is governed by a Board of Directors with support from a Scientific Advisory Committee and Executive Management Committee. Founding chairman, John Grant announced his retirement in July 2013 and will be replaced by Dr Les Trudzik from September 2013.

The Directors met on six occasions during the 2012/13 financial year:

- 19 July 2012
- 30 August 2012
- 29 November 2012 (AGM)
- 14 February 2013
- 12 March 2012
- 6 June 2013

Board members perform the following roles in addition to their fiduciary responsibilities:

- **John Grant** – Financial oversight, Commonwealth Government liaison, NSW State Government liaison, NSW subcontractor liaison.
- **Les Trudzik** – Performance and impact framework development, Victorian State Government liaison, Victorian subcontractor liaison, Western Australian State Government liaison, Western Australian subcontractor.
- **John de Jersey** – Chair of Scientific Advisory Committee, Queensland State Government liaison, Queensland subcontractor liaison, South Australian State Government liaison, South Australian subcontractor liaison.

- **Sue Meek** – Communications, DIISR liaison, BPA representative to the Great Barrier Reef Foundation (ReFuGe) governance committee.

Executive Management Committee

The Executive Management Committee is responsible for executing all funding agreements and implementing the strategic initiatives endorsed by the Board. Committee members are:

- Andrew Gilbert, Chair
- Dr Sue Forrest, Convenor Genomics
- Prof Jeff Gorman, Convenor Proteomics
- Prof Tony Bacic, Convenor Metabolomics
- Prof Matthew Bellgard, Convenor Bioinformatics
PLATFOR...
Australia’s Bearded Dragon

Australia is home to over a million native species of plants and animals, with 80 percent uniquely present in Australia. This unique diversity is believed to have evolved since Australia split from Antarctica about 50 million years ago and presents an exciting opportunity to study evolutionary mechanisms. In collaboration with the University of Canberra, the GDU have used high throughput sequencing technology to understand the molecular basis of sex chromosome dosage compensation in snakes and lizards; identify sex chromosome and possibly sex determining genes in lizards; understand evolutionary relationship between lizards; and immunity development in marsupials.

Geraniums Offer Insight to Genetic Evolution

Researchers at the Australian National University are incorporating the latest genomics technology to investigate the evolutionary lineages of the geranium. There are dozens of Australian Pelargonium which display significant physical diversity and a remarkable ability to flourish and adapt to a wide range of habitats. The research has combined morphological and reproductive data with DNA sequencing in order to define and compare different species. This genetic investigation has revealed that the Australian ancestry of Pelargonium is actually quite recent and may encompass only eight species. It is also offers unique insights into the relationship between environmental variation and genetic factors that give rise to a new species.

PROTEOMICS

Capabilities in protein identification, analysis and characterisation at our proteomics nodes continue to expand along with services offered to both academic and industry based researchers. Multiplexed quantitative proteomic analysis using various forms of mass spectrometry is showing high demand. Requests for high quality amino acid analysis and biomarker detection also remain high.

In addition to the support provided to the national dataset program, the following projects highlight other collaborative research:

Biomarkers in Colorectal Cancer

The Australian Proteome Analysis Facility (APAF) is collaborating with the Royal North Shore Hospital and a multinational pharmaceutical company to undertake a clinical trial involving analysis of biomarkers of response to VEGF (vascular endothelial growth factor) inhibition in colorectal cancer.

APAF has developed a series of highly specific, quantitative mass spectrometry assays to measure putative protein biomarkers in patient plasma over treatment time courses. These data will be linked to patient survival to assess their predictive utility. Project outcomes are expected to lead to improved choices for patients and clinicians in managing treatment plans.

Impact of Malaria in Pregnancy

APAF has been collaborating with Professor Stephen Rogerson from the University of Melbourne who is investigating how pregnant women become susceptible to malaria, and how this impairs the growth of the unborn child. As part of this work, APAF has studied amino acid levels in maternal blood and cords to confirm the nature of amino acid transport in mothers/babies infected with the parasite that lead to reduced birth weight. These studies have been reported in the journal *PLOS Pathogens* and offer the first explanation for foetal growth restriction in malaria pregnancies.

Personalised Cancer Vaccines

Researchers at the Queensland Institute of Medical Research (QIMR) are combining next generation sequencing, state of the art mass spectrometry and classical immunological techniques to identify peptides that signal an immune response in the effort to develop personalised cancer vaccines. The project is utilising patient samples that demonstrated a positive immune response in clinical trials of a novel melanoma vaccine. The peptides present on these patients’ cells will be analysed using a state-of-the-art Orbitrap Elite mass spectrometer and compared with next-generation sequencing data to identify which cells were mutated by the cancer and which of these also demonstrated an immune response. The outcomes will facilitate development of algorithms for predicting immunogenic peptides and will hopefully lead to outcomes of clinical benefit.

METABOLOMICS

Metabolomics continues to grow in recognition and importance as a critical ‘omics’ technology.
of specialties around cancer, nutrition, lipidomics and other fields demonstrate growing demand and research interest. Our metabolomics nodes continue to build an international reputation as both service providers and technology innovators having introduced new and emerging technologies into the Australian research landscape. Project highlights include:

Grape and Wine Research

The Australian Wine Research Institute is collaborating with a leading Italian research group to comprehensively map non-volatile grape and wine metabolites. The research aims to discover and identify grape metabolites that may relate to positive effects on human health.

A new service to analyse volatile phenolic glycosides in grapes has also been developed. These compounds represent biomarkers for smoke exposure of grapes and potential smoke taint in juice. The ability to screen for them allows the wine industry to assess the impact of smoke exposure during bushfire events or controlled burns.

Lipidomics Clues in Plant Stress Tolerance

The University of Melbourne is using a new technique utilising MALDI-FT-MS based imaging, to monitor spatial distributions of lipids in the roots of different barley genotypes with contrasting salinity tolerance levels.

The aim is to investigate the link between plant metabolite composition and functionality and gain a better understanding of the role of lipids, lipid composition and signalling in plant stress tolerance. The study will combine a tissue type-specific transcriptomics analysis and integration of ‘omics data with physiological measurements. The project also has generated a number of international collaborations.

Pollution Biomarkers

The University of Melbourne is collaborating with the Centre for Aquatic Pollution Identification and Management to identify biomarkers for chemical pollutant exposure in aquatic organisms as indicators of pollution in waterways. Using an untargeted metabolomics approach, the University has shown that different metal contaminants were differentially affecting the metabolic pathways in an Australian midge (chironomid). This is a key pathway that produces metabolic precursors that protect against oxidative stress. These studies have the potential to provide more sensitive tools for early detection and accurate identification of pollution and to assign the appropriate corrective actions.

Biomarkers for Malaria

Researchers at the University of Melbourne and QIMR are working to identify metabolite biomarkers useful for early diagnosis of malaria. The research aims to measure changes in blood and urine metabolite levels in human volunteers who have been infected with the malarial parasite, Plasmodium falciparum. State-of-the art mass spectrometry is used to simultaneously monitor changes in a large number of metabolites during the course of infection and subsequent drug treatments to provide insights on the effectiveness of drug treatments and the potential role of patient metabolism in determining the outcome of malaria infections.

Biofuels for Aviation

The Australian Institute for Bioengineering and Nanotechnology is seeking to develop and improve the process of converting sugarcane to aviation fuel by engineering microbes to improve fuel production from sucrose. An important element of the research is ensuring that the performance of biofuels match those presently used in aircraft and making the cost of production more economically viable. The research is part of the Queensland Sustainable Aviation Fuel Initiative which involves a number of research and commercial partners.

Vitamin D Standardisation

Development of a new 2D LC-MS based approach to quantitate vitamin D metabolites in human serum by the University of Western Australia has resulted in an invitation to take part in the first International vitamin D Standardisation Program initiated by the US Centre for Disease Control. The high precision and accuracy of this new approach has also resulted in additional invitations to support standardisation programs for other clinically important endocrine hormones.
The demand for bioinformatics expertise and capacity is ever increasing as it pervades all life science research. Our bioinformatics partners all work to expand Australia’s bioinformatics capabilities through the essential mix of building information infrastructure, providing new analytical tools and delivering sophisticated research services.

Centre for Comparative Genomics

The Centre for Comparative Genomics (CCG) at Murdoch University delivers critical bioinformatics support to Bioplatforms Australia projects, nodes and clients. Throughout the year it has advanced challenging gene assembly efforts, developed new analytical tools and provided data management solutions. Highlights include:

**Data Store** - New data storage infrastructure with 500Tb storage capacity has been developed to automate the process of capturing Bioplatforms Australia national framework datasets. The archive has a sophisticated system to automatically store triplicate copies of the datasets and can be synchronised to other national data stores.

**Wheat Dataset Project** - The CCG delivered seven genome assemblies for fungal pathogens and established a bioinformatics framework to achieve standard data formatting and annotation to enable data sharing. The wheat datasets are being widely shared with the scientific community and have so far been deposited with the International Nucleotide Sequence Database Collaboration.

Work with the AGRF on wheat chromosome 7A has continued. The complex computational challenge of assembling this very large and highly repetitive chromosome has prompted a novel experimental design utilising BAC libraries. The approach essentially split chromosome 7A into approximately 800 segments of 1.2 million bases in average. So far, the segment assemblies have been longer than expected suggesting the technique is highly effective and balances assembly time, costs and accuracy.

A collaboration with the University of Göttingen has been instigated to leverage the open-source gene prediction tool, Augustus for the wheat gene models being developed for the wheat proteogenomics studies.

**Proteogenomics** - The proteogenomics pipeline developed by the University of California - San Diego (UCSD) and previously incorporated into the YABI workflow environment, has been utilised with wheat and grape MS/MS datasets to improve genome annotations. The CCG helped resolve issues in the pipeline code to ensure proper peptide mapping and prediction of divergent proteins (splicing isoforms) for complex genomes which has improved accuracy and outcomes in the wheat dataset analysis.

**MASTR-MS & MAMBO MS** - The MASTR-MS LIMS application, developed in collaboration with Bio21 Institute for the metabolomics nodes, provides comprehensive project management including experimental design and data management. User feedback prompted refinements in system usability including improvements in a synchronisation tool for heterogeneous metabolomics instruments. The MAMBO-MS metabolite library has also been further developed to improve data analysis, data mining and metadata capture. These systems have now been installed at the University of Western Australia.

**Systems Biology Initiative**

The Systems Biology Initiative (SBI) at the University of NSW undertakes basic and applied research in the development and application of bioinformatics for genomics and proteomics. It has a close relationship with our genomics node at the Ramaciotti Centre for Gene Function Analysis and proteomics node at the Australian Proteome Analysis Facility but also serves the broader researcher community.
During 2012, the SBI worked with 15 research groups to produce advances in the fields of microbiology, marine science, plant physiology, neurobiology and cancer research. It continues to develop novel approaches for data integration and analysis and has accrued particular expertise in network-based analyses. Networks have been constructed, modelled and visualised using public domain data and several bioinformatics tools in order to analyse and understand changes in intracellular networks in mammalian cells, as a consequence of perturbation or disease. This expertise has been applied to research on melanoma cancer with the Millennium Institute and Melanoma Australia, and a project to understand the systems-level impact of the deletion of PI3K in cardiac tissue with the Baker Heart Institute. The former project resulted in the development of a new software program to measure the correlative expression patterns of disease-related genes and proteins.

De novo transcriptomics has been central to a project investigating the basis of behavioural changes in brain of the Australian black cricket. The same workflows can be accessed through shared high performance computing services and have proven useful to other research projects such as the analysis of the koala transcriptome.

Bioinformatics Resource Australia - EMBL

In June 2011 an Australian mirror of the services provided by the EMBL European Bioinformatics Institute was launched at the University of Queensland. The mirror, combined with the NCI Specialist Facility for Bioinformatics, represents a major bioinformatics resource for Australian scientists and has been titled the Bioinformatics Resource Australia - EMBL (BRAEMBL).

The services and IT infrastructure delivered by BRAEMBL have made a significant impact to the science community through greater access to data, analytical tools and large computational infrastructure. During the year, the use of all BRAEMBL services recorded very strong growth with an increase of 636 per cent in the total number of site users. Over 250,000 hits to BRAEMBL sites were recorded and the number of researchers downloading data increased by 406 per cent. Nearly one years’ worth of CPU hours were consumed by researchers in analysing bioinformatics data on BRAEMBL’s compute cluster.

The significant increase in the mirror usage reflects growing awareness of BRAEMBL services combined with the successful introduction of new initiatives such as hosting the MEME Suite. MEME and EBI FTP Mirror services were the most popular over the year comprising 63 per cent of all accessed services.

BRAEMBL is seeking to extend its services beyond mirroring. It considers it can leverage the considerable potential represented by Australia’s recent investments in eResearch infrastructure and a deeper collaboration with the EBI to advance Australian bioinformatics services. It wants to become an early adopter of new techniques and help address the broad concern of “lack of expertise” by collaborating on training programs such as those initiated by Bioplatforms Australia.

BRAEMBL also hopes to grow its role in Australian data sharing by developing data curation capabilities. It has recently been appointed to curate the ’Sea-quence’ data generated by Bioplatforms Australia for the Great Barrier Reef dataset project. This is BRAEMBL’s first data sharing project and represents important capability building.
Bioplatforms Australia is responsible for building scientific infrastructure in the specialist fields of genomics, proteomics, metabolomics and bioinformatics. It supports Australian life science research with crucial investments in state-of-the-art technologies and cutting-edge expertise. Investment funding has been provided by the Commonwealth Government’s National Collaborative Research Infrastructure Strategy (NCRIS) and the 2009 Super Science initiative. Co-investments have also been made by State Governments, research institutes and commercial entities. Bioplatforms Australia aims to promote broad access to Australia’s world-class ‘omics’ capabilities to maximise the scientific endeavours of the Australian research community.

For more information, please contact Andrew Gilbert of Bioplatforms Australia at agilbert@bioplatforms.com or call 0410 538 648. www.bioplatforms.com

Bioplatforms Australia Ltd
Level 4, Building F7B Research Park Drive, Macquarie University NSW 2109 Australia
P: +61 410 538 648
F: +61 2 9850 6200
www.bioplatforms.com