Submission
2016 National Research Infrastructure Roadmap
Capability Issues Paper

Name	**Michael Dobbie** **on behalf of the APN Management Group, Governance Board, and associates**
Title/role	**CEO**
Organisation	**Australian Phenomics Network**

This submission canvases perspectives and insights from the Australian Phenomics Network (APN) partners, board, affiliates, collaborators, and other existing NCRIS capabilities. The APN also contributed to a separate joint submission made by a number of NCRIS capabilities. In addition to the general responses to the full range of questions expressed in the joint submission, we offer the following focussed, capability-specific responses to selected questions.

To meet Australia's existing and emerging needs and drive biomedical research innovation, the APN describes in this submission several new strategic investments. These would build upon and extend the enormous value being derived from our existing suite of infrastructure services that create and supply animal and cellular models of disease. These investments would provide complementary solutions, but through different technologies and services, to those through the existing and proposed research infrastructure provided by other NCRIS capabilities, principally but not exclusively Bioplatforms Australia, Therapeutic Innovations Australia, National Imaging Facility, and the Australian Microscopy and Microanalysis Research Facility. Together through tightly-integrated approaches, we are ideally positioned to accelerate discovery and boost translation across the health and medical research sector.

Over the past 10 years and with NCRIS support, the APN has created tools and developed services that have almost exclusively been deployed to advance biomedical research. Therefore, our stakeholders and experience lie firmly within the Health and Medical Sciences capability area as set out in the Issues Paper, with some additional responses relevant to National Security, Underpinning Infrastructure and Data capability areas.

The proposed transformative measures will overcome existing bottlenecks in Australia's disease model creation and characterisation pipeline, which generates tools for gene function discovery, understanding disease pathogenesis, and identifying new therapeutic targets for drug discovery.

Section 5. Health and Medical Sciences

The post-genome era of biomedical research is fast moving with disruptive technologies emerging at an unprecedented pace. It has become impractical for individual research groups to become skilled in each relevant new technique and **the need for centralised research infrastructure is greater than ever before**. Several factors have led to rapid change:

- An explosion in animal models: large scale international and national projects means that we are nearing the completion of the Aim to generate a mutation in every gene within the mouse genome. Of the 24,433 annotated protein coding mouse genes, 21,826 (90%) have been mutated in cells alone or in experimental mice.
- A revolution in our ability to edit the genome of organisms and cells. This will, in turn, further increase the number of mouse models, with the focus shifting from null alleles (full ablation of gene function) to the production of refined alleles that more closely mimic human disease variation. For the first time, genome modifications can be made in somatic cell lines.
- New high-resolution imaging methods for characterising organisms and cells.
- New technologies for high-throughput drug screening in cell lines.

The bottleneck has shifted from creation to the characterisation of disease models and validation of therapeutic targets and molecules. One factor that has altered the space in which the APN operates is the explosion in mouse models. To date this has been underpinned by chemical mutagenesis and the ES cell to mouse routes, where mutations of 21,826 annotated protein coding genes in the mouse genome have been created. The expansion of this number of targeted genes continues apace and has been largely achieved through concerted efforts by international collaborations, principally through the International Mouse Phenotyping Consortium (IMPC), our key international partner. Moreover, this already staggering number of mouse mutants is set to increase at an unparalleled rate owing to the advent of highly efficient gene editing capability now available through the CRISPR technology. Our ability to modify the genome of any organism has been rapidly and profoundly transformed by this technology – an approach that facilitates not only gene inactivation but also the introduction of subtle variants like those that underpin many human genetic diseases. It is evident that the next phase of human genetic research will involve mimicking putative disease causing DNA alterations (identified in patients) in the mouse in order to assess the pathological relevance of the alteration. CRISPR mutagenesis has effectively removed the bottleneck of generating mouse models of simple Mendelian traits, while at the same time bringing into focus the need to streamline, automate and standardise methodologies for characterising disease models, identifying therapeutic target molecules and pathways, and develop new drugs to ameliorate disease.

Question 15: Are the identified emerging directions and research infrastructure capabilities for Health and Medical Sciences right? Are there any missing or additional needed?

We describe three important platforms and expertise requiring new investments that were missed in the Issues Paper and which would greatly extend the national benefits of biomedical research.

<u>Investment Area 1</u>

Australia needs renewed investment in effective and efficient pathways to therapeutic compound discovery and translation into clinical opportunities for treating patients. This initiative woud be enabled by large-scale, high-quality animal breeding and genome engineering facilities coupled with high-throughput therapeutic target and molecule assay design and screening facilities, and supported by medicinal chemistry capability.

Outcome

To increase the number of therapeutic lead compounds produced in Australia by enhancing our ability to develop models of human disease, screen for compounds that ameliorate the disease phenotype in animal models, and, when promising compounds are found, to use medicinal chemistry to mould these into drug-like compounds.

Background

Australia is very good at basic biomolecular research. We rival most developed countries in terms of total per capita investment in medical research and numbers of researchers. We have found the underlying molecular mechanism for many diseases and published this work in high profile internationally peer-reviewed journals. We attract international funding for our basic research and are well represented on podia at international conferences.

Yet, our prowess at discovery is not matched by our ability to translate our findings into new therapeutics. Much of our work finishes with the discovery of potential new targets for disease therapy and either we leave our work at this point and wait for others to perform the translation, or we sell out cheaply to multinational companies. With a little more effort and some investment in crucial infrastructure, we could convert and leverage our discoveries beyond target identification to the production of valuable lead therapeutic compounds.

Over the past two decades a change has occurred in the value chain of biomedical research. The discovery of genes, proteins and mechanisms underscoring disease has lead to an increased effort to discover therapeutic drugs aimed at suppressing or modifying the expression of disease-specific molecules. Several models of drug discovery (from bench to bedside) have been entertained; in some cases validated drug targets backed up by strong IP have been licenced to pharmaceutical companies where no capacity for drug discovery exits within the research institute itself. In this model, most of the the risk in embarking on a drug discovery programme is taken on by the pharmaceutical company and, accordingly, the financial reward (up-front payment/royalty sharing) is typically disappointing.

A variation of this scenario, is that a research institute with solid IP surrounding a validated disease-specific target would establish a partnership with a pharmaceutical company where an on-going collaboration is established that might be based on, for example, the evaluation of the safety and efficacy of a drug candidate in one or other animal model and even extend to collaborative clinical trials (see case study below).

Serious consideration must be given to increasing Australia's capacity for undertaking drug discovery as well as the ability to pursue lead drug candidates through drug development to ideally emerge as legitimate clinical candidates. Nothwithstanding the need at some stage to partner with a pharmaceutical company, by extending the pipeline from discovery to candidate therapeutic molecules would add considerable IP value as a prelude to any business proposition.

Pathway to therapeutic compounds

Fundamental biomedical research into the cause of disease can lead to the identification of potential drug targets. These are mainly proteins whose altered action as a consequence of interaction with a therapeutic drug, leads to an amelioration of the disease phenotype, either through alteration to normalise function or through an enhancement or repression of expression of the target gene.

The ability to test hundreds of thousands of compounds for biological activity extremely rapidly and in a cost-effective manner ("High-Throughput Screening", or "HTS") is one of the most important approaches to the discovery of new first-in-class drugs[1]. Australia is continuing to establish world-best infrastructure at at several sites, including the Walter & Eliza Hall Institute of Medical Research, Griffith University (Compounds Australia), Peter MacCallum Cancer Centre, and through the CRC for Cancer Therapeutics, to name a few. However, unlike the central support for HTS in Europe[2], or the extensive support for HTS in the USA made possible through NIH support, or facile access to HTS by Chinese scientists through the Shanghai Institute of Materia Medica, Australia will benefit from new investments to build a well-integrated and sustainable HTS ecosystem. Poorly integrated or insufficiently resourced infrastructure can result in important disease targets in Australian Universities and Medical Research Institutes lying fallow or being offered to overseas pharmaceutical companies for further research and development.

As set out in the 2011 STRATEGIC ROADMAP FOR AUSTRALIAN RESEARCH INFRASTRUCTURE (page 49), translational medicinal chemistry is a key capability that dovetails with HTS in the development of drugs to treat human disease. With advances in molecular biology furnishing an ever-increasing list of disease-relevant biological targets, the bottleneck in medicinal chemistry today is even greater than it was in 2011, the two key limitations being synthesis and purification. No technological solution can currently replace skilled organic chemists to undertake synthesis, but in recent years great advances have emerged in the design and implementation of equipment that allows multiple reactions to be undertaken simultaneously followed by automated multiple parallel purification equipment. Such capability is essential in order to dovetail with, and capture maximum value from, advances in upstream high throughput screening technologies so that Australia can develop Australian drugs. It is to be emphasized, however, that support for soft infrastructure (highly skilled staff) of such a multiple parallel reaction and purification facility, is important.

Model systems to test "target" hypotheses. In order to determine whether potential targets are going to be useful, most researchers transfer the disease phenotype into a model organism and to use this as a preclinical model to test potential therapeutic compounds. These models systems are usually rodents, particularly mice or cell lines but zebrafish, worms, flies and others are used. Usually the human disease is transferred into the model system using human polymorphisms that contribute to disease and that have been discovered through basic biomolecular research into the disease. These polymorphisms are engineered into the model organism or cell line either by making similar changes into the same gene in the model, or by moving the entire human gene into the model organism or cell line. The emergence of the CRISPR/Cas9 gene editing system has revolutionised the speed with which human genetic polymorphisms can be engineered into mice, other organisms and

[1] Eder et al., 2014, Nat Rev Drug Discov. 13, 577-587.

[2] Mullard, 2013, *Nature Rev. Drug Discov.*, 12, 734-735.

cell lines. This technology is relatively cheap and highly efficient, therefore affording the ability to be extremely thorough in analysing not only the primary target of interest but also members in parallel and intersecting biochemical pathways. Inactivating or activating mutations can be engineered into these genes and the effect on the disease assessed in disease models. Furthermore, generating large databanks of disease models and concurrent detailed phenotypic analysis provides an invaluable resource for rapidly understanding of drug action when a target is identified in high throughput screen.

Infrastructural needs: Therefore, large animal houses and highly efficient CRISPR/Cas9 genome editing (to generate animal models) and animal production service facilities are required. The animal facilities must have appropriate input from researchers and knowledgeable vivarium design professionals to ensure they meet contemporary and future animal and research needs; this is not always the case in Australia.

Identification of small molecular compounds that reverse disease phenotype. The development of drugs to treat human disease starts with the identification of lead compounds that engage with a validated target or another target that is not necessarily associated with the biochemical pathway containing a disease gene. These compounds are identified through the screening of tens of thousands of compounds representing chemical libraries of very varied composition. In parallel, functional genomics approaches using gene knockdown or gene knockout on a genome scale can help define key molecular pathways and reduce the need for broad scale compound screening. The screening strategy comprises either a biochemical or cellular assay in a cell line model best reflecting the disease being studied, with each compound or gene target assessed in a single well format. Such screening is performed in fully automated liquid handling and imaging facilities by experts in developing laboratory-scale assays to a high throughput scale with robust and sensitive readouts. In this way, both loss or gain of function biochemical assays or cellular phenotypes can be measured. Variations of this theme use whole organisms as the screening assay; for example this has been done successfully with zebrafish, drosophila and C. elegans embryos.

Australia has a number of extremely skilled and highly productive screening facilities that are currently heavily utilised. An expansion of the existing infrastructure will enable us to more rapidly generate a large number of lead compounds to take to the next stage. There is no limit in the number of targets that could be screened however we need to invest in developing more biologically representative assays that can be screened rapidly.

Infrastructure needs: To generate more lead compounds we need to develop a critical mass of centralised high throughput screening facility nodes by increasing robotic automation liquid handling and imaging infrastructure.

Conversion of HTS hits into drug candidates. Ideally, a chemical screen identifies compounds that engage with the 'target' but at this early stage are generally not ideal drugs. These compounds may lack affinity for its target, may not be sufficiently soluble in water, or may not be able to adequately cross biological membranes. A lead compound therefore typically needs significant modification to change its properties to a drug-like compound. This challenge is taken on by medicinal chemists who explore related structures for increased activity, solubility, and permeability into cells. The same

biochemical screening assay is then iteratively analysed and subsequently its effect assessed in a model organism. After manipulation of the structure of the lead compound, an entity emerges that is able to bind to the target with a high affinity, alter its function and change the disease phenotype in an animal model. This is then classified as a drug-like compound that has high-value IP and should form the basis of negotiations for the clinical testing of this compound by international pharmaceutical companies. It is only at this point that significant value can be returned to this country through deals done based on outputs from our drug discovery and development research programs.

In contrast to small chemical compounds, the development of biological agents such as antibodies may be appropriate, especially if the target is on the surface of the cell. The identification of therapeutic antibodies follows a different path but still uses mainly mice to assess the efficacy of the antibody on the disease phenotype.

Functional genomics is just as important as compound screens in directing the discovery of gene targets, which ultimately can circle back to compounds, but can allow a more focused approach, or it can take a number of compounds and assess their specificity if targets are known before pursuing more detailed medicinal chemistry analysis. This enables the researcher to delve into the mechanism of action, which is critical for understanding on and off-target effects of drugs and expertly triaging candidates before they enter expensive trials.

Infrastructure needs: High capacity medicinal chemists with dedicated time and the ability to manipulate the structures of dozens of lead compounds annually and who are closely associated with the HTS facilities to rapidly characterise and assess the altered molecules. While this may look like research, the medicinal chemists represent an essential ingredient in the discovery pipeline to secure clinical grade drugs with specificity for their disease targets.

Case Study: Australian discovery leads to new anti-cancer drugs

Strategic institutional investments have translated fundamental discoveries of promising new anti-cancer therapeutic pathways. Despite a plethora of putative targets, examples where new molecules progress through clinical trials to market are few.

A prototypic case began with the discovery of role of BCL-2 in promoting cancer cell survival. This discovery in 1988 by Prof David Vaux at WEHI sparked a series of biological discoveries highlighting the role of the BCL-2 family of proteins in the regulation of intrinsic cell death and how the deregulation between pro- and anti-survival members of this family plays a critical role in cancer.

Recognising the therapeutic potential, WEHI Director Prof Suzanne Cory made strategic institutional investments to build new structural biology capability, integrated with enhanced screeing and medicinal chemistry capability to create a new state-of-the-art biotechnology centre in 2003.

The first high throughput screen (HTS) against a BCL-2 family member was run in 2004 leading to the successful development of a potent and selective series of inhibitors of the pro-survival protein BCL-XL. A collaboration with the US biotechnology company Genentech boosted the screening-biology-medicinal chemistry pipeline and developed new inhibitors of BCL-2 family proteins. With the addition of the AbbVie pharmaceutical company (then Abbott Laboratories) to the academic-industry team multiple series of drug-like molecules targeting BCL-XL, BCL-2 and MCL1 were discovered. The small molecule ABT-199 was selected as a clinical candidate to target BCL-2.

The WEHI-Genentech-AbbVie team pursued clinical development of ABT-199/venetoclax with the first Phase I patients receiving the investigational treatment at the Royal Melbourne Hospital in 2011. Subsequent successful, multi-centrered, clinical phases led to milestones decisions from the FDA: first granted “breakthrough therapy status” in 2015, “priority review” in January 2016 and approval for use by 17p-deleted CLL patients in April. Approvals for the treatment of hematological and solid cancers are expected within the next few years.

Therefore, new investments in national research infrastructure can mature existing pockets of capability and enable the scale and quality for a potent discovery and development pipeline, overcoming the medicinal chemistry bottleneck, and facilitating early engagement and new partnerships with industry for national benefit and enhanced international reputation.

Governance requirements

Broadly, strategic investments in the governance and management of the multi-faceted drug discovery ecosystem should be carefully considered and designed to ensure there is a collective coherence in the way that the various partner nodes will collaborate to deliver infrastructure with open access and for national benefit that goes beyond serving only host institutional interests.

Investment Area 2 – Imaging

A 3-D Imaging Platform is required to develop Australia's first efficient and sensitive protocols for biological Computed Tomography, enabling cell-level resolution in whole organ and organism images and automated comparative pathology of mouse tissues. The APN currently offers tissue pathology services, to Australian researchers, biotechnology companies and pharmaceutical industry. These services rely on manually cutting sections through tissues and examining them one-by-one to compare between sections and between genotypes. Sifting through such data by trained experts requires a tremendous amount of time. Current state-of-the-art international projects use 3-D imaging of specimens in combination with computer-automated image registration algorithms to circumvent this bottleneck in pathology. The Australia has the world-leading 3-D imaging expertise required to build this new infrastructure to develop algorithms for automated comparative pathology of mouse tissues. This could leverage in-kind support from Australian National Data Service, Research Data Storage Infrastructure and the National Computational Infrastructure and would synergise with infrastructure offerings of other NCRIS capabilities, including the National Imaging Facility and the Australian Microscopy and Microanalysis Research Facility.

Automated processes for interrogating pathology data are needed for moderate throughput, high-resolution, standardised analysis of biological specimens. The use of histopathology techniques to document changes in normal pathophysiology caused by expression of an introduced genome change has long been a cornerstone of disease model characterisation. This is, however, a labour intensive activity. Generally tissues and organisms for investigation are embedded in a suitable medium and sliced into sections. Next, each section is stained to differentiate particular features and then examined using a microscope and the data interpreted by a highly trained expert. It is unfeasible to do this on the scale now required to exploit the information contained in the vast banks of mouse models worldwide. In addition, this method leads to considerable data loss because even trained experts are unable to reconstruct in their mind the complex 3-D morphology of the tissues under comparison. A clear requirement here is to rethink histopathology and to incorporate lessons learned from other areas of the APN services, each of which are shifting towards automation with the concomitant advantages of increased throughput and standardisation. The new infrastructure could be built via collaboration with existing expertise in 3-D imaging to develop algorithms to identify statistically significant anatomical differences between specimens via computer-automated image registration. This would be an Australian-first service that will help biomedical researchers take advantage of the decreased time for animal model generation brought about by CRISPR technology.

Developing the 3-D Imaging Platform will have several important impacts. Providing cellular resolution for whole organ-scale images has the potential to replace histopathology methods for some applications. It offers 3-D imaging of structures instead of the 2-D of standard histopathology. The images will be made available through TissueStack, recently developed by the National Imaging

Facility to enable efficient remote access to large image files. Hitherto, remotely managing large image files has been a bottleneck that has now been solved. The proposed project will increase standardisation of analysis and remove human error and minimise variability; factors that ultimately reduce the number of research animals required – a key goal of global animal experimentation principles.

The 3-D Imaging Platform infrastructure would be highly valuable by our key international partner, the International Mouse Phenotyping Consortium (IMPC). During the this year's Annual Meeting of the IMPC one of the technological capabilities they are seeking to source or boost is high-resolution imaging of mouse embryos and neonates, such as would be provided by an Australian 3-D imaging platform. The first landmark paper demonstrating the enormous value of this approach is set to be published on-line by the journal Nature on 14 September. The provision of quantitative high-resolution imaging service for fixed embryos and tissues offers a potential way to **overcome the geographical and biosecurity barriers we face in Australia in seeking to more fully participate in projects being run by international organisations** such as the IMPC. Animals produced overseas can be stabilised in fixing solution and cost-effectively shipped to us without the need for expensive packaging or subject to international quarantine restrictions. The advantages of easy sample handling together with access to the highest resolution imaging facility of its kind in the world will no doubt attract fee-for-service business from anywhere in the world and significantly boost Australia's reputation in the fields of animal modelling and imaging.

Bioinformatics investments cannot be neglected. Both suggested Investment Areas 1 and 2 (above) will generate increasingly vast volumes of complex data, particularly the structural analysis of lead compounds. The value of those investments will be significantly diminished without substantial investment in bioinformatics resources as eleborated in general terms elsewhere in this submission and in the aforementioned joint submission from other NCRIS capabilities.

Investment Area 3 – Veterinary Pathology

New investmenst are required to overcome the deficit of veterinary pathologists and comparative medicine specialists in Australia. This country suffers from a dearth of boarded veterinary pathologists (e.g., American College of Veterianry Pathologists, ACVP[3]) and boarded laboratory animal veterinarians (e.g., American College of Laboratory Animal Medicine, ACLAM[4]). Each specialty fosters excellence towards protecting and improving human, animal, and environmental health. Ideally, the veterinary pathologist would benefit from mouse pathology specialization training, the US NIH funds such opportunities. Many countries and regions have benefitted from these credentialed, specialist veterinarians; Singapore and China made extenstive hires of these trained individuals to not only plan and design their impressive contemporary research facilties and vivaria, but to assist driving their research programs towards meeting or exceeding international accreditation standards. Australia has the advantage of credentialed veterinary schools the would facilitate veterinary training in these strategic, internationally recognized specialty veterinary areas.

[3] acvp.org
[4] aclam.org

Question 16: Are there any international research infrastructure collaborations or emerging projects that Australia should engage in over the next ten years and beyond?

In the broad area of personalised medicine, both the National Human Genome Research Institute and the Broad Institute, in the USA, are world leaders. We should endeavour to align our investments with emerging trends and advances made by international organisations such as these.

In the area of drug discovery pipelines, Europe has established the European Lead Factory (ELF)[5]. The ELF is underpinned by the Joint European Compound Library (JECL), comprising 320,000 compounds from proprietary collections of seven pharmaceutical company partners and 200,000 compounds from the Public Compound Collection. Researchers can access the JECL through the screening centre. A project proposal is submitted on-line with brief description and suggested screening cascade and ideally evidence of a working minituarised assay. If accepted, the project enters the HTS pipeline. All costs are covered by through the Innovation Medicines Initiative, which is jointly funded by the EU and the European pharmaceutical industry. Upon receipt of a qualified hit list of up to 50 hits, the project team pursues early stage hit-to-lead activities through the ELF. Three years of exclusivity are offered. In short, there is widespread recognition in Europe for the need to widen this bottleneck. Engagement with this and similar international organisations could lead to exchange of expertise, insights into which business models are proving successful, opportunities to build complementary infrastructure, and potential international collaborations.

In the area of mouse phenomics, Australia has benefitted from liberal access to a number of research infrastructures that are wholly-funded and managed by international agencies. Often this access and benefit has been at little or no cost to Australian researchers. For example, the APN has had free access to international libraries of mouse strains (such as those created and managed through the International Mouse Phenotyping Consortium, IMPC) that have been imported for experimental research in Australia. We have also been fortunate to maintain our founding membership of the Steering Committee of the IMPC. This representation has enabled us to contribute and be privy to strategic discussions that identify emerging trends and determine international investments in experimental genetics relevant to modelling human disease. Due to lack of appropriate infrastructure (such as a national mouse clinic providing a comprehensive characterisation service and funding to conduct systematic phenotyping assays), Australia has not been a substaintial co-investor in the IMPC or similar international programs (such as equivalent Asian and North American consortia), nor have we contributed to the vast amount of well-curated, open data that is being generated by these organisations.

As they mature through the establishment to expansion phases of operations, international Genomic and Phenomic projects are starting to discuss the requirement for reciprocal contributions as criteria for consortium membership and access privileges. We cannot rely on long-term access to these resources on a consumer-only basis. Therefore, it is imperative for Australia to plan for strategic investments to maintain beneficial access. For example, Australia could become a valued member of the IMPC if we made focussed investments in Phenomics

[5] See www.europeanleadfactory.eu

infrastructure to generate high quality, well-curated and annotated datasets in areas of national strength such as in Immunology and Developmental Biology. The data would not only directly benefit research projects but could also be aggregated with the IMPC's datasets, now recognised to be unsurpassed in scale and quality and providing new insights in genotype-phenotype linkages. We could also consider hosting major international meetings of the consortia. Substantial co-investments in designing and building data-intensive research resources in this way will lift Australia's international reputation as a contributor not just as a benefactor.

Not only is the IMPC leading the way in the generation of large-scale, high-quality Phenomic dataset, it has conducted a thorough appraisal and ongoing review of the application and impact of the fast-evolving genome editing technology (CRISPR), and our seat at that table has helped our service delivery nodes keep abreast of advances in technology and deployment. The IMPC has also been recognised as a prototype for internationally-resourced and -managed research infrastructure. The IMPC is one of only two research infrastructures (and the only life science project) selected for study by the Group of Senior Officials of the G8 Global Research Infrastructures intiative during their meeting hosted by Australia earlier this year. Our long-term engagement the IMPC should continue to be fostered and strengthened as it enters a new expansion phase of operation.

Question 17: Is there anything else that needs to be included or considered in the 2016 Roadmap for the Health and Medical Sciences capability area?

Through NCRIS and other investments during the past 10 years a number of large-scale, very high quality research infrastructures have been built that have integrated and serviced an extensive research network and continue to operate. A large-scale, resilient and responsive ecosystem of infrastructure capabilities has been created. For example, the APN has co-invested with Bioplatforms Australia, the National Computational Infrastructure, and the Australian National Data Service to build interdisciplinary research infrastructure in the areas of genomics, bioinformatics, and data stewardship and management for discoverability.

To expand expertise, resources and service models, it is sensible to build upon these existing capabilities and established modes of interoperability. The strength of the NCRIS ecosystem serving Health and Medical Sciences is the provision of a suite of technology plaftorms, technical and scientific expertise, project management, and program governance with the capacity to be agile and responsive in the face of technological advances and directional changes in emerging national needs. New investments will ideally be founded in, and built upon, appropriate, existing strengths to effectively serve future needs, whether they be envisioned and unexpected.

National Security

Question 27: Are the identified emerging directions and research infrastructure capabilities for National Security right? Are there any missing or additional needed?

While almost all of the APN's infrastructure has been designed and built to serve the field of biomedical research, we note that the APN's listing as existing infrastructure in the Issues Paper was limited to the National Security capability area. While we do not draw our current clientele from this domain, our services and expertise can certainly be applied to address biosecurity challenges

(including understanding the pathogenesis and control of human and animal disease). For example, the APN partners support and manage Australian Quarantine Importation Service-approved laboratories, that are integrated within our animal facilities. These are mainly used for the import and export of mouse strains for biomedical experiments involving both national and international collaborations. There is additional potential for new interoperability with other containment and animal breeding facilities such as those provided by the Australian Animal Health Laboratory, particulary as would be required for experiments involving virulent pathogens. Credentialed veterinarians play a critical role in developing and analysing these animal models by working closely and collaboratively with research teams.

Underpinning Research Infrastructure

Question 30: Are the identified emerging directions and research infrastructure capabilities for Underpinning Research Infrastructure right? Are there any missing or additional needed?

The APN enables the creation of increasingly large genomic and phenomics datasets (e.g. DNA sequence and image data) that should be aggregated, and rendered discoverable and accessible for optimal value and longevity. There is an increasing reliance on Tier 1/Tier 2 computational infrastructure for the characterisation of new models of disease, both during their creation by the APN and when used for research by our clients. Storage and portability requirements are growing exponentially, and there is an increasing need for sophisticated access and authentication resources.

Data for Research and Discoverability

Question 33 Are the identified emerging directions and research infrastructure capabilities for Data for Research and Discoverability right? Are there any missing or additional needed?

The scale and complexity of new capability requirements are very well described in the Issues Paper. Data-intensive research creates increasingly unwieldy datasets that are more difficult to move, aggregage, discover, share, and interrogate. This problem can be solved by enabling ways to bring powerful data visualisation tools to the data, and simultaneously interrogate, disparate datasets. For example, biomedical researchers using the APN infrastructure to investigate increasingly complex research questions will benefit from access to computational tools and capacity that far exceed current suitability and capability. Additionally, we have already reached the point where the required computational and data management solutions (whether existing or developing) are generally poorly understood by biomedical researchers so that not only are the solutions not available but researchers are poorly equipped to describe the technical requirements and how to resource the solution. Therefore, investments in cross-disciplinary training and the provision of data technology “translators” would effectively bridge this biology-data technology divide enabling the design and provision of innovative packages of computational tools for burgening demands data-intensive research.

Question 35: Is there anything else that needs to be included or considered in the 2016 Roadmap for the Data for Research and Discoverability capability area?

It's one thing to instil the value of data sharing , it's another thing to provide a mechanism to enable, encourage, and reasonably enforce it. For example, the NIH strongly recommends that data (and model organisms) produced from NIH grants be shared publicly[6]. But this cannot be enforced because public databases do not exist for all data types and projects are not always designed to capture set of core parameters (where they exist) and informative metadata. The same deficiencies are found in Australia. Some international granting bodies absolutely require sharing of data (even before publication), for example, some US Center grants, and this is explicit in the relevant funding agreements. Investments in minimizing the logistic and financial barriers to data sharing will return great dividends through decreasing duplication of experimental effort while increasing reproducibility of results through stringent curation. When projects are not fully funded, often the data management components are first to be shaved. Therefore, if data sharing is mandated sufficient funding should be preserved for that purpose. Alignment of funding obligations and provisions will optimally build data-based infrastructure.

Other comments

If you believe that there are issues not addressed in this Issues Paper or the associated questions, please provide your comments under this heading.

There exists a general issue regarding the ability of researchers to secure sufficient funding to enable access to some research infrastructures on a partial cost-recovery basis. The cost-recovery models need to be set at a level that reflects available research project funding. For example, large-scale, high-throughput screens have been historically difficult to fund through national research councils. Therefore, it is important to acknowledge that cost recovery structures for each each technology platform must reflect project funding available to researchers for those kinds of experiments. Optimal infrastructure utilization can be achieved through finely-balanced funding from multiple sources. Therefore, to establish a sustainable business model and optimal utilization of the even the most cost-effective services, national investments need to be tighly coordinated through both research infrastructure programmes and through project granting schemes.

[6] The NIH Resource Sharing Plan covers data and model organisms (https://grants.nih.gov/grants/peer/guidelines_general/Resource_sharing_plans.pdf)