

1 Executive summary

1.1 Project rationale and overall objectives of the project

Drug discovery is ideally based upon human models representative of the patient population. However access to relevant tissue is difficult. This bottleneck has been overcome with the advent of stem cell based disease modelling whereby tissue from a patient can be reprogrammed to induced pluripotent stem cells (iPSC) and then differentiated into a cell type closely resembling the tissue bearing the brunt of disease. Modelling disease with human iPSC in a dish offers many opportunities but for it to be included into drug discovery, there are still many challenges. The StemBANCC consortium was an academic-industry partnership tasked with addressing these challenges by advancing science and implementing novel, cutting edge technologies. It provided well characterized patient iPSC accessible for academia and industry and demonstrated proof of concept for the utility of stem cell based models.

1.2 Overall deliverables of the project

StemBANCC was a 5 year programme divided into 12 workpackages (WP) with the following objectives:

- a) Recruitment of 500 subjects with defined diseases, drug responses and healthy volunteers
- b) Re-programming subjects' somatic cells into 3 high quality iPSC per subject
- c) Research conducted with robust research governance, ethical and legal compliance
- d) Developing differentiation protocols for cell lineages relevant to diseases
- e) Cellular phenotyping of iPSC from subjects' iPSC and thereby identify disease relevant cellular pathophysiology or toxicology end-points.
- f) Molecular profiling and generation of integrated molecular networks underpinning disease
- g) Development of assay platforms for high-throughput drug screening using iPSC.
- h) Establishing a single point of access database detailing subject clinical information, stem cell lines, quality control indices and molecular profiles.

1.3 Summary of progress versus plan since last period

The StemBANCC programme has successfully completed many of the key tasks set out in the description of work and achieved its core objectives. The research plan required coordination of 12 substantive workpackages with several inter-dependencies. Furthermore, we sought to conduct consortia wide experiments in order to take full advantage of the multi-disciplinary expertise available in the consortium.

1. Subject recruitment was highly successful and we achieved recruitment of 496 subjects of the planned recruitment of 500.
2. Over the course of StemBANCC, the reprogramming plan was adapted such that three categories of reprogramming would be undertaken. 120 subjects were classed as Priority 1 for which 3 iPSC clones per subject were generated, fully expanded and quality controlled. 140 subjects were classed as Priority 2 for which 1 iPSC clone was generated, fully expanded and quality controlled. The remaining 240 subjects were classed as Priority 3 and iPSC clones

were generated but no further expansion was performed. This prioritisation helped manage the available resources. The strategy has been very successful and reprogramming has completed as planned. Furthermore, we were able to upgrade further Priority 3 lines and completed 177 of the total 195 category III lines assigned to UNEW, UOXF and KCL.

3. Throughout the StemBANCC programme research has been conducted with robust research governance, ethical and legal compliance
4. Many novel differentiation protocols have been developed which are published/pending publication. Importantly, differentiation protocols have been reproduced in multiple labs confirming their robustness. Nevertheless, through a multi-centre experiment conducted in StemBANCC the challenges of working with iPSC differentiated cells for molecular studies was clearly established.
5. The disease and toxicology workpackages have in the last year been able to generate novel data on disease phenotypes including in pain disorders, Alzheimer's Disease, Parkinson's Disease, autism and diabetes. These have again resulted in publications or in manuscripts that will shortly be submitted. Furthermore a very important study on the effect of drugs on a range of iPSC derived cell types has now been completed, which will provide the international research community an important resource for future work on human chemogenomics.
6. The conversion of disease relevant phenotypes to high throughput assays is challenging but several of the partners in academia and industry have been able to achieve this such as the use of the MEA platform for high-throughput screening of drugs affecting nociceptor excitability or a 384-well based high content screening autophagy assay in cortical neurons.
7. StemDB has operated as the single point of access database and has been housing the clinical data, the cell line information and the SNP typing data of the iPSC lines. The transcriptomic, proteomic and metabolomic experiments were conducted through multiple individual experiments and were not housed in StemDB. The relevant, de-identified data in StemDB is being transferred to EBISC and EBI in order that the cell lines will always have an associated dataset when they are requested by researchers.

1.4 Significant achievements since last report

- Long term sustainability of iPSC lines and Data secured: the Biomaterials and Data Access Policy will continue to be applied through EBISC and Coriell (see Deliverable D2.05 and section 1.5 below). These biorepositories will ensure the relevant ethical responsibilities and requirements for the transfer of StemBANCC Biomaterials and Data to recipients of biomaterials and data.
- WPO2 Subject identification and recruitment: the final recruitment of 496 subjects of the 500 planned is also a major achievement since the last report. The publication of a reference paper covering the availability of data (biomaterial and clinical) and a detailed description of the StemBANCC cohort has been prepared and is now in submission.
- UOXF and UNEW delivered Priority 1 and Priority 2 iPSC lines were delivered to the Biorepository in Birmingham. 177 of the total 195 category III lines assigned to UNEW, UOXF and KCL were reprogrammed. Subsequently, Birmingham has successfully transferred to EBISC who have made the StemBANCC iPSC lines available on their catalogue. Newcells Biotech Ltd, a spin-out company of Newcastle University continued conversion of 52 priority III lines to priority II thereby

increasing the availability of these lines to the scientific community. These will be transferred to EBiSC along with the other completed lines when available.

- StemBANCC established efficient iPSC gene editing methodologies and generated of multiple gene-edited iPSC lines by UOXF, Janssen and AZ for use by StemBANCC partners. These will be passed to EBiSC for use by the international research community.

1.5 Scientific and technical results/foregrounds of the project

- An important outcome for StemBANCC was the reprogramming of iPSC lines for use by the scientific community. We are therefore pleased that we have been able to deliver to EBiSC and soon Coriell iPSC lines fully reprogrammed and expanded for rapid distribution on demand. Furthermore, priority 3 lines will also be deposited at EBiSC that could be revived and expanded for distribution. Already within StemBANCC 1337 vials of iPSC lines have been ordered and we expect that demand for StemBANCC iPSC lines will steadily increase.
- At least 40 separate transcriptomics experiments for StemBANCC including the large scale multi-drug experiment involving analysis of more than 1500 samples. Exome sequencing was performed on the samples and proteomics and metabolomics have also been successfully applied to the multi-centre experiments. The molecular profiling and analysis has added significantly to many aspects of StemBANCC research. This includes demonstration with an iPSC dopaminergic Parkinson's Disease (PD) model molecular signature, that a drug (clioquinol) can rescue PD cellular phenotypes. In the coming 6-12 months, the results of these experiments will be published and data will be deposited in international sequencing data repositories such as NIH GEO and EBI for other researchers to examine.
- StemBANCC neuroscience researchers have made significant progress in developing novel differentiation methods that are published/being prepared for publication. This includes peripheral peptidergic nociceptors, microglia, and more synaptically mature glutamatergic neurons. Uniquely, protocols for eliciting different types of synaptic plasticity at glutamatergic and inhibitory synaptic connections between cortical neurons have been developed. Disease models have been established for pain (Nav1.7 and TRESK), neurodegeneration (tau mutations, microglial responses, familial AD, GBA, LRRK2) and neurodysfunctional disorders (Shank3). Finally, the development of mixed co-cultures comprising neurons and glia, associated tools such as reporter constructs and work with 3D human mid-brain like organoids has been progressed.
- The diabetes work-package has comprehensively been addressing the challenges of using iPSC in disease modelling and diabetes including mapping cell specific fetal pancreatic development (eLife 2017; PMID 28731406); generation of reporter lines *NKX6.1* and *MAFA* to further optimize the efficiency and/or maturation of the existing endocrine lineage differentiation protocols; epigenetic and transcriptional profiling of pancreatic differentiation (published in Stem Cell Reports and Diabetologica); molecular profiling and physiological characterisation of diabetic lines (HNF4A and HNF1A mutations). This has for example generated a valuable omics biobank that will be invaluable for future research.
- The focus of the toxicology workpackage has been on improving differentiation protocols in order that the models become more predictive. The achievements have included small molecular based hepatocyte differentiation and 3D hepatocyte models; renal epithelial and podocyte differentiation, cardiomyocyte in 2D and 3D. Additionally, relevant pathways using a system

toxicology approach in liver, heart, CNS and kidney have been identified. The work in StemBANCC has led to over 10 publications since the last reporting period.

- A variety of iPSC based assay methodologies and techniques have been established including microscopy based high content screening of autophagy and synaptic transmission; electrophysiology cell-based assays with multi-electrode arrays combined with live cell imaging; and development of microfluidic methods for liver, neuron and hepatocytes. A large number of SOP protocols for cell based assays with iPSC lines are available on www.stembancc.org and a book has been commissioned by Springer Nature Verlag to make these protocols publicly available.

1.6 Potential impact and main dissemination activities and exploitation of results

Socio-economic benefit for European citizens

The outputs of StemBANCC have established resources, methods and tools to undertake disease modelling and drug discovery using human induced pluripotent stem cells (iPSCs). During the project lifetime there has been considerable knowledge exchange between the consortium partners and as a result, many of them have established in-house expertise for working with iPSC models. We have also been able to demonstrate that iPSC models can recapitulate disease relevant phenotypes and can form the basis of a drug discovery programme. We believe this will encourage academia, biotechnology companies and pharma to expand preclinical science activities either directly or as external partnerships and collaborations. This will result in investment as well as job and career opportunities in this sector for European citizens.

Contribution to the health of European citizens

StemBANCC has been addressing highly challenging, prevalent and difficult to treat disorders of our time. There has been a significant reluctance by industry to invest in these areas since they are considered high risk despite the unmet medical need. In order to find effective treatments and ultimately improve health, it is necessary to demonstrate that a rational lower risk approach to drug discovery is feasible. The work of StemBANCC has significantly contributed to establishing methods and platforms to de-risk drug discovery. Ultimately, the use of human cellular models will enable improvement in selection of compounds that are efficacious and reduced risk of toxicity. These benefits will take many years to materialise, but StemBANCC has set the highly important foundations.

Increased European competitiveness and attractiveness for biopharmaceutical research

StemBANCC has been a major research programme involving 42 institutes in Europe and approximately 200 researchers. The partners in StemBANCC have significantly advanced in the skills and expertise of working with human cellular models and there has been considerable knowledge exchange. Furthermore, StemBANCC has provided new opportunities and projects to work together. A significant impact of StemBANCC is therefore the highly desirable skills and resources available to European research groups. This will ensure that Europe remains an internationally leading centre for Biopharmaceutical research.

Main Dissemination Activities

The main outputs of StemBANCC have and will be disseminated by publications. More than 80 papers have been published already and this number will increase further in the next 6-12 months (see Deliverable 12.05: planned publications). Furthermore, StemBANCC partners have been very actively participating in conferences and research meetings as well as public talks and patient support groups to ensure dissemination to the scientific community and general public.

Exploitation of Results

The work of StemBANCC has led to commercial outputs. Newcells is a spin-out company that has arisen from StemBANCC partner UNEW, at least in part due to the increasing reprogramming activities assigned by StemBANCC to UNEW. Furthermore, StemBANCC has led to the commercial exploitation of R&D results (e.g. method for differentiation of pluripotent stem cells into multi-competent renal precursors) as several patent applications show (see Deliverable 12.05).

1.7 Lessons learned and further opportunities for research

StemBANCC has been a highly successful research programme involving approximately 200 scientists in leading-edge research over the last 5 years. It has been complex and there have been challenges, but by careful planning and cooperation, we have achieved our important objectives.

The most important challenge we faced was the highly demanding work required for iPSC reprogramming. This was undertaken initially at an academic institution (UNEW) that did not have a track record of generating iPSC. However, the scale of iPSC production turned out to be too challenging and we took appropriate steps to apportion some of the reprogramming to two additional groups with a re-channelling of resources. This required a re-allocation of funds which would have been simpler if a reserve fund had been established at the outset. For future programs of similar size and complexity, we would recommend a portion of the funds to be kept as “reserves”.

Reprogramming approaches have now changed since the start of the StemBANCC project and production has become an activity routinely undertaken by biotechnology companies so any future reprogramming should be undertaken by such companies.

The other significant challenge for StemBANCC was the conversion of disease phenotypes into drug screening assays. This will remain a difficult area because firstly, significant resources are required to build confidence that a phenotype is disease relevant using a number of complementary approaches. Secondly, reducing the phenotype into a reproducible assay that is suitable for screening also requires significant investment in time and resources. Nevertheless, StemBANCC has been demonstrably successful in developing disease relevant screening assays.

As with all good research programmes, StemBANCC prompts many new questions and opportunities for future research. This includes:

1. Methods to improve the maturation of differentiated cells produced from iPSC in order to better model late-onset disorders
2. Approaches to tackle the heterogeneity of cell types inherent in iPSC differentiation, in order that biological and disease relevant processes are detectable, particularly for molecular phenotypes
3. Platforms that could reduce the cost of iPSC based phenotypic high throughput assays in order to enable their widespread use for compound screening

4. Development of human physiological systems that more accurately reflect the native physiology of tissues such as the brain, pancreas or liver to allow validation studies and to replace animal use

Further research on the effect of drugs in human cells – StemBANCC conducted a large scale study but even so we only examined the effect of 12 drugs. Extending this study to evaluate how currently marketed drugs effect human cell types would be provide important insights into mechanisms of action and the development of new drugs.