Establishing the UK DNA Bank for motor neuron disease (MND)

Smith, Lucy; Cupid, B. C.; Dickie, B. G M; Al-Chalabi, A.; Morrison, K. E.; Shaw, C. E.; Shaw, P. J.

DOI:
10.1186/s12863-015-0236-6

License:
Creative Commons: Attribution (CC BY)

Document Version
Peer reviewed version

Citation for published version (Harvard):
https://doi.org/10.1186/s12863-015-0236-6

Link to publication on Research at Birmingham portal

Publisher Rights Statement:
Eligibility for repository: Checked on 14/12/2015

General rights
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

• Users may freely distribute the URL that is used to identify this publication.
• Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
• Users may use extracts from the document in line with the concept of ‘fair dealing’ under the Copyright, Designs and Patents Act 1988 (?)
• Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.
Establishing the UK DNA bank for motor neuron disease (MND)

Abstract:
In 2003 the Motor Neurone Disease (MND) Association, together with The Wellcome Trust, funded the creation of a national DNA Bank specific for MND. It was anticipated that the DNA Bank would constitute an important resource to researchers worldwide and significantly increase activity in MND genetic research. The DNA Bank houses over 3000 high quality DNA samples, all of which were donated by people living with MND, family members and non-related controls, accompanied by clinical phenotype data about the patients. Today the primary focus of the UK MND DNA Bank still remains to identify causative and disease modifying factors for this devastating disease.
<table>
<thead>
<tr>
<th>Expert in the field</th>
</tr>
</thead>
<tbody>
<tr>
<td>No competing interest</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Robert Swingler, Medicine</td>
</tr>
<tr>
<td>Consultant, Ninewells Hospital, NHS Tayside, Dundee</td>
</tr>
<tr>
<td>Expert in the field</td>
</tr>
<tr>
<td>no conflict of interest</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Jochen Weishaupt, Professor</td>
</tr>
<tr>
<td>Universitat Ulm</td>
</tr>
<tr>
<td><a href="mailto:jochen.weishaupt@uni-ulm.de">jochen.weishaupt@uni-ulm.de</a></td>
</tr>
<tr>
<td>expert in the field</td>
</tr>
<tr>
<td>no conflict of interest</td>
</tr>
</tbody>
</table>
Establishing the UK DNA Bank for motor neuron disease (MND)

*Smith L¹, Cupid BC¹, Dickie BGM¹, #Al-Chalabi A², #Morrison KE³, #Shaw CE², #Shaw PJ⁴

*corresponding author lucy.smith@mndassociation.org

#Principal Investigators for the DNA Bank, listed in alphabetical order.

¹Motor Neurone Disease Association, PO Box 246, Northampton, NN1 2PR, UK;
belinda.cupid@mndassociation.org; brian.dickie@mndassociation.org

²NIHR Biomedical Research Unit in Dementia, Department of Clinical Neuroscience, King’s College London, London, SE5 8AF, UK; ammar.al-chalabi@kcl.ac.uk; chris.shaw@kcl.ac.uk

³Institute of Biomedical Research, College of Medical and Dental Sciences, University of Birmingham, Birmingham B15 2TT, UK; k.morrison@bham.ac.uk

⁴Sheffield Institute for Translational Neuroscience, University of Sheffield, 385A Glossop Road, Sheffield, S10 2HQ, UK; pamela.shaw@sheffield.ac.uk

The authors declare that they have no competing interests.

Abstract

In 2003 the Motor Neurone Disease (MND) Association, together with The Wellcome Trust, funded the creation of a national DNA Bank specific for MND. It was anticipated that the DNA Bank would constitute an important resource to researchers worldwide and significantly increase activity in MND genetic research. The DNA Bank houses over 3000 high quality DNA samples, all of which were donated by people living with MND, family members and non-related controls, accompanied by clinical phenotype data about the patients. Today the primary focus of the UK MND DNA Bank still remains to identify causative and disease modifying factors for this devastating disease.

Keywords: Motor Neurone Disease (MND), Amyotrophic Lateral Sclerosis (ALS), Biobank.
Motor Neuron Disease (MND) is a fatal, rapidly progressive disease that affects the brain and spinal cord and which ultimately leads to respiratory failure around 2-5 years following symptom onset [1, 2]. Approximately 1 in 300 people develop MND but its prevalence is low, at about 6-8 in 100,000 because of short life expectancy [3]. There is no diagnostic test and treatment is largely palliative, with only one agent, riluzole, having a modest effect in extending survival. Genetic factors undoubtedly play a role in most cases of the disease, both in pathogenesis and rate of progression, with about 5-10% of all patients having a clear family history of MND and in some cases, frontotemporal dementia [4, 5, 6]. Over 100 genes have now been implicated in the causation of MND [7]. No consistent environmental risk factor has been identified, although it is possible that such factors may trigger disease in genetically susceptible individuals, and therefore it is plausible that apparent sporadic cases of MND will be genetically determined to some degree [8, 9].

An essential starting point for successful genetic research is access to high quality samples, accompanied by detailed clinical information. Large-scale gene sequencing and association studies need many thousands of samples to be screened such that results are statistically significant. Access to such samples had become a major obstacle in exploring the pathogenesis of MND and the concept of a MND DNA Bank was born. The objectives of the initial study were threefold: 1) To collect cohorts of patient, parent/sibling and control samples from sporadic and familial MND; 2) To collect clinical information in order to examine susceptibility traits in clinical subgroups of MND; 3) To make this resource available to the international research community and to foster collaboration between research teams, in order to identify genetic risk factors for MND.

**ORGANISATIONAL STRUCTURE OF THE UK MND DNA BANK**
The UK MND DNA Bank was a collaborative project adopting a ‘Hub and Spoke’ model, with three regional ‘Hub’ centres linking with a total of 16 ‘Spoke’ centres (Table 1). Samples were obtained from sporadic and familial MND patients attending MND clinics in the UK, their spouses (or other genetically unrelated controls) and blood relatives. Samples within the UK MND DNA Bank are housed at CIGMR Biobank, at the University of Manchester. In addition, as one of the Public Health England collections, the European Collection of Cell Cultures (ECACC) manages the transformation and storage of EBV-transformed lymphocytes derived from blood samples from participants providing an everlasting supply of DNA for the Bank.

The MND Association’s Biomedical Research Advisory Panel (BRAP) oversee the governance and the strategic development of the DNA Bank, ensuring that samples are utilised in an appropriate fashion, and that any clinical information requested is appropriate for the proposed study. The Technical Access Committee (TAC) at CIGMR Biobank, determine sample requirements for the technology platform to be used, the quantity of sample required and ensure any leftover samples are returned or destroyed. All applications for access to the samples are judged on merit. In order to receive material and clinical information from the DNA Bank, all applicants must agree to the terms and conditions of sample use (see supplementary info). This specifies the user and specific purpose for which the samples and data are to be licensed, including standard terms as to the ownership, exploitation and dissemination of results, and requirements that the user conforms to the terms of the participants’ consent.

SAMPLE COLLECTION, STORAGE AND QUALITY CONTROL

Sample collection began in 2003. All participants were over 18 years of age. In order to ensure that the patient cohort was representative of disease prognosis, patients must have experienced symptom onset (significant muscle weakness) on or after January 2002. All patients fulfilled El Escorial
criteria for probable or definite Amyotrophic Lateral Sclerosis (ALS) [10]. Patients presenting with Progressive Muscular Atrophy (PMA), Primary Lateral Sclerosis (PLS) or Progressive Bulbar Palsy (PBP) were also included in the study. Patients were recruited by consultant neurologists with a specialist interest in MND in participating centres. Patients participating in other clinical research projects were not excluded from the study. Blood samples were also collected from consenting partners/carers, providing some degree of matching in terms of age, education, environmental exposure and often ethnicity. Where patients presented with familial MND, blood samples were collected from family members for linkage analysis. Where patients presented with sporadic MND, where possible, blood samples were also collected from parents or from a parent and sibling, to give so-called ‘Trio samples’ increasing the amount of genetic information available for researchers.

Informed consent to participate was sought from all patients, family members and controls. Ethical approval for the collection of samples and the creation of the UK MND DNA Bank was given by the Trent Research Ethics Committee in February 2003 ref MREC/02/4/107 and in July 2009, ref 09/HO405/32. Participants were provided with detailed information and contact details and could withdraw from the study at any time. The samples were pseudo-anonymised and an online clinical database was developed to facilitate data entry and collection by the research nurses and enable tracking of trends in clinical parameters such as symptom onset and presentation for data analysis. Storage and access to this data set is in accordance with the UK Data Protection Act 1998 [11].

Prior to 2010, DNA extraction from donated blood samples was carried out at individual Hub centres using the Nucleon BACC3 protocol (Amersham, UK). Extracted DNA was sent to CIGMR Biobank for long-term storage. On receipt, all DNA samples were run on 1% agarose gels alongside molecular weight markers of appropriate size to check integrity. From August 2010, DNA extraction was carried out at CIGMR Biobank using automated robotic processing under ISO900:2000
operating standards. In all cases, both when imported from Hub centres, or extracted by CIGMR Biobank themselves, DNA concentration was measured using a nanodrop spectrophotometer. Samples with OD ratios outside the normal range were removed from the cohort and contaminants washed using ethanol precipitation. Final DNA concentration was measured using Quant-iTTM Picogreen® dsDNA Assay Kit (Invitrogen™ Life Technologies, UK). Samples were measured on 96 well plates, in triplicate against standards of known concentration for quality control.

DNA aliquots are stored in 2D bar coded tubes for sample tracking purposes. A relational database recorded the 2D barcodes associated with each patient/donor ID. All samples within the collection were screened for gender using PCR on presumed duplicate samples according to standard protocols. Samples with a mismatch between the expected gender as recorded in the patient information, and actual gender as confirmed by PCR, were rescreened using an alternative PCR method of gender identification based on the absence/presence of Alu sequence [12]. Any samples with a confirmed discrepancy were ring fenced from the collection and suspended from the in-house laboratory management system.

Peripheral blood lymphocytes (PBLs) were isolated from whole blood samples at ECACC using density gradient centrifugation. An aliquot of untransformed PBLs was stored in liquid nitrogen for safekeeping, whilst the remaining PBLs were transformed using the Epstein Barr virus according to standard protocols [13]. The resulting lymphoblastoid cell lines were cryopreserved and are used to restock the DNA Bank when stock levels become low.

4. THE UK MND DNA BANK

In October 2012, at the end of the collection period, the UK MND DNA Bank comprised 3159 high quality DNA samples. Of these 1344 samples were taken from individuals diagnosed with sporadic MND (see figure 1A and 1B). There were 133 familial MND samples within the collection and a
further 500 samples taken from family members, including samples that form 28 parent trio sets and 27 sibling trio sets. The remaining 1085 samples were taken from controls. In line with population-based demographic for the disease [14] the breakdown of gender in the collection is around 60% male (Figure 1A). The average age of onset was approximately 62 years of age (Figure 1C). Each sample is accompanied by a minimum dataset of: age at which the samples were taken; gender; disease status; and where appropriate diagnostic certainty (El Escorial Status) and age of onset (calculated from date of birth and date of symptom onset). An extended dataset has been collected for as many participants as possible but it is not a complete dataset for the entire collection (see Table 2). In total 2653 frozen lymphoblastoid cell lines are held in storage at ECACC following a PBL transformation success rate of 97%. Of these, 1267 samples were generated from whole blood taken from patients with sporadic MND. 115 cell lines were generated from familial samples and the remaining 1058 cells lines have been established using blood samples obtained from control or family members (see Figure 1D).

Table 3 shows the success rates for PCRs performed on the DNA samples within the collection. The failure rate of the quality control assay was less than 1.5% suggesting that the quality of DNA within the collection is very high. The gender results from these assays were directly compared to the gender recorded for individuals on the clinical database. Where there was a discrepancy between the expected gender and that determined in the assay, patient clinical notes were rechecked. In the absence of a clerical error, samples were rescreened using both the original AMEL marker and an alternative gender marker, the Human ALU expansion [12]. Sixty eight samples that continued to show a discrepancy between the expected gender and the assay gender were ring fenced from the collection and suspended from the laboratory management system.
The UK MND DNA Bank was designed to be available to the international research community. DNA samples from the bank represent an incident not prevalent population and are unlikely to be biased. The fundamental guarantee that any DNA Bank must be able to give is that it can provide high quality DNA samples with good integrity and accompanying high quality clinical data.

Understandably, a constraint of the DNA Bank is that the genomic DNA supply itself is limited and although cell lines have been established, the DNA from such cell lines may have sequence changes compared with the original genomic samples. This fact must be considered when choosing to use cell line derived DNA even if the DNA itself is of a high standard as demonstrated by the rigorous quality control assays in place. As part of the governance of the DNA Bank, the MND Association must ensure compliance with legal and regulatory requirements. The Association must also guarantee that the resource adheres to rigorous research standards and is used in the further understanding of motor neuron disease, this includes prioritising access to those parts of the DNA Bank that are limited in availability, clarifying intellectual property rights and disseminating the results that flow from it.

To date more than twenty projects have withdrawn samples from the DNA Bank. DNA samples have been used in complex, technical protocols such as genotyping, gene sequencing and genome-wide association studies and numerous papers have been published or are in press [15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27 and 28]. Importantly, projects using samples from the DNA Bank have directly led to the detection of several MND causing genes including *C9orf72* and more recently *Tub4A* [15, 17, 22, 23 and 28]. With researchers now encouraged to publish in an open access format as part of the DNA Bank governance, and to deposit data from sequencing projects within accessible databases such as ALSOD: the Amyotrophic Lateral Sclerosis Online Database [7] and European Genome-Phenome archive [29], the dissemination and discussion of results by the research community is ensured. In 2014 a proposal to perform whole genome sequencing on DNA samples from the UK MND DNA Bank as part of the international collaboration called Project MinE [30]
was approved. This exciting project will allow Next Generation Sequencing data to be collected from DNA Bank samples and shared across research groups. The data will also confirm the accuracy of existing studies through imputation. It is hoped that sequencing DNA Bank samples will allow the identification of rare variants responsible for sporadic disease, continuously widening our knowledge about how genetic changes can contribute to MND.

Ethical approval to extend the use of the cell lines beyond their original scope of providing an everlasting supply of DNA was granted in 2014 by the Derby- East Midlands Research Ethics Committee ref no. 14/EM/1088. This change in permission will potentially allow researchers to generate primary neuronal cultures and highly desirable induced pluripotent stem (iPS) cell lines from the cell lines stored at ECACC. The iPS cell lines could act as new disease models for drug screening and other potential treatments, as well as acting as tools for analysing downstream mechanisms involved in disease pathogenesis. Clearly the role of the DNA Bank in the governance of such samples will be paramount; it is simply not enough to provide high quality samples, but following how those samples have been used and ensuring the results are disseminated and discussed is the only way to ensure research continues to move forward. The original scope of the DNA Bank was to make a quantal difference in our understanding of MND and it is well on the way to fulfilling this promise.

Acknowledgments: This project was supported by a programme grant from the Motor Neurone Disease Association (grant ref 6700). Information regarding the DNA Bank, including the terms and conditions for sample use can be accessed on the MND Association website [31]. The creation of the lymphoblastoid cell lines was supported by a grant from the Wellcome Trust (grant ref 070122/A/02/Z). Ethical approval for the creation of the UK MND DNA Bank was given by the Trent Research Ethics Committee in February 2003 ref MREC/02/4/107 and in July 2009, ref 09/HO405/32. Subsequent approval for the extended use of the cell lines within the DNA Bank was approved by the Derby- East Midlands Research Ethics Committee in Sept
2014, ref no. 14/EM/1088. We are grateful for the support of the Dementias and Neurodegenerative Disease Research Network (DeNDRoN). We would like to acknowledge our partners at the CIGMR Biobank (formerly BioBanking Solutions (BBS), UK DNA Bank Network (UDBN)) based at the University of Manchester and the European Cell Culture Collection at Public Health England. This project would now have been possible without the support of collaborators in the participating centres across the UK [31]. Thank you to all the people with MND and their families who participated in this project.

References:


29. The European Genome-Phenome archive website: https://www.ebi.ac.uk/ega/home

30. The Project MinE website: https://www.projectmine.com/

31. The MND Association Website: http://www.mndassociation.org/dnabank

FIGURE LEGENDS

Table 1: Hub and Spoke Model for Sample Collection

The UK MND DNA Bank was a collaborative project adopting a ‘hub and spoke’ model with three regional ‘hub centres’ linking with a total of 16 ‘spoke centres’. These included hospitals that are part of the MND Association’s Care Centre Network and centres which form part of the Department of Health / NIHR Dementias and Neurodegenerative Diseases Research Network (DeNDRoN). The three Hub Centres were established at London (King’s College Hospital), Sheffield (Royal Hallamshire Hospital), and Birmingham (Queen Elizabeth Hospital). Recruitment to the study and sample collection was coordinated at the Hub centres by a DNA Bank Co-ordinator based in London. Samples were obtained from sporadic and familial MND patients attending MND clinics in the UK, their spouses (or other genetically unrelated controls) and blood relatives. A DNA Bank research nurse was affiliated to each Hub centre to act as patient liaison, collect clinical information from patients, controls and family members, and take blood samples. Laboratory Technicians were also employed at each Hub Centre to assist in the preparation of sample collection packs for participating satellite centres, DNA extraction and final sample storage.

Figure 1: The UK MND DNA Bank
The UK MND DNA Bank comprises 3159 high quality DNA samples. Of these 1344 samples were taken from individuals diagnosed with sporadic MND (see figure 1A and 1B). There were 133 familial MND samples within the collection and a further 500 samples taken from family members, including samples that form 28 parent trio sets and 27 sibling trio sets. The remaining 1085 samples were taken from controls. In line with previous findings, where MND has been diagnosed, the breakdown of gender in the collection is around 60% male (Figure 1A). The average age of onset was approximately 62 years of age (Figure 1C). In total 2653 frozen lymphoblastoid cell lines are held in storage at ECACC. Of these 1267 samples were generated from whole blood taken from patients with sporadic MND. 115 cell lines were generated from familial samples and the remaining 1058 cells lines have been established using blood samples obtained from control or family members (see Figure 1D).

Table 2: Clinical information available from the UK MND DNA Bank

Each sample withdrawn from the UK MND DNA Bank is accompanied by a minimum dataset of: age at which the samples were taken; gender; disease status; and where appropriate diagnostic certainty (El Escorial Status) and age of onset (calculated from date of birth and date of symptom onset). An extended dataset has been collected for as many participants as possible but it is not a complete dataset for the entire collection. The clinical information was collected by the Research Nurse using a brief clinical questionnaire. Identifying data was kept at each Hub centre in secure locations in accordance with the Data Protection Act.

Table 3: Quality Control PCR Assay fail rate

Quality control assays were carried out across the collection. In an initial analysis of the collection 768 samples were screened using the ABI Identifiler kit. Following this, a further 2170 samples were
screened in a gender-based assay using the AMEL marker. Only 62 samples showed a continued
discrepancy between the gender of the actual DNA sample and that stated in the clinical notes.
Whilst this is still a low level of error for a collection of this size, the samples were ring fenced from
the collection.
Table 1: Hub and Spoke Model of Sample Collection

### King's College Hospital, London
- Royal Free Hospital London
- Bart's and The London NHS trust
- Poole NHS trust, Poole
- Cambridge University Hospital
- Derriford Hospital, Plymouth
- Southampton University Hospital
- Queen's Hospital, Romford

### Queen Elizabeth Hospital, Birmingham
- John Radcliffe, Oxford
- Belfast City Hospital, Belfast
- Walton Neurological Centre, Liverpool
- Southmead Hospital, Bristol

### Royal Hallamshire Hospital, Sheffield
- Royal Preston Hospital
- Greater Manchester Medical Centre, Manchester
- Ninewells Hospital, Dundee
- Queen's Medical Centre, Nottingham
- Royal Victoria Infirmary, Newcastle
Figure 1: The UK MND DNA Bank Collection

1A: MND status and gender of samples held in the DNA bank

<table>
<thead>
<tr>
<th>Status</th>
<th>Female</th>
<th>Male</th>
<th>Total numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sporadic</td>
<td>530</td>
<td>834</td>
<td>1,364</td>
</tr>
<tr>
<td>Familial</td>
<td>53</td>
<td>80</td>
<td>133</td>
</tr>
<tr>
<td>Family*</td>
<td>290</td>
<td>210</td>
<td>500</td>
</tr>
<tr>
<td>Unknown</td>
<td>26</td>
<td>46</td>
<td>72</td>
</tr>
<tr>
<td>Control</td>
<td>677</td>
<td>408</td>
<td>1,085</td>
</tr>
<tr>
<td>Total number</td>
<td>1,576</td>
<td>1,578</td>
<td>3,154</td>
</tr>
</tbody>
</table>

1B: Diagnostic certainty of samples within the DNA bank

- ALS definite
- ALS probable
- Control
- PBP
- PLS
- PMA
- Unconfirmed

1C: Age range of symptom onset within the DNA bank

1D: MND status and gender of cell line samples held at ECACC

<table>
<thead>
<tr>
<th>Status</th>
<th>Female</th>
<th>Male</th>
<th>Total numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sporadic</td>
<td>489</td>
<td>798</td>
<td>1,287</td>
</tr>
<tr>
<td>Familial</td>
<td>44</td>
<td>71</td>
<td>115</td>
</tr>
<tr>
<td>Family*</td>
<td>96</td>
<td>79</td>
<td>175</td>
</tr>
<tr>
<td>Unknown</td>
<td>25</td>
<td>40</td>
<td>65</td>
</tr>
<tr>
<td>Control</td>
<td>627</td>
<td>366</td>
<td>993</td>
</tr>
<tr>
<td>Total number</td>
<td>1,281</td>
<td>1,354</td>
<td>2,635</td>
</tr>
</tbody>
</table>
Table 2: Clinical information available from the UK MND DNA bank

### Minimum Dataset

All approved users of the samples will receive a **minimum dataset** of anonymised information on participants.

The fields within the minimum dataset include:

- Age
- Gender
- Affectation status (control, familial ALS, sporadic ALS etc)
- Diagnostic certainty (El Escorial status)
- Age of onset

### Extended Clinical Dataset

An **extended dataset** has been collected from as many participants as possible. The extended dataset is only available to researchers who collaborate with the Principal Investigators of the DNA Bank.

The fields within the extended dataset include:

- Clinical history (including site of presentation, dominant hand, family history)
- Family tree pedigree
- Investigation and results (including EMG, MRI, Nerve conduction studies)
- Medications (including riluzole)
- ALSFRS score
- Physical examination history (including % FVC, % VC)
- MRC scores
Table 3: Quality Control PCR fail rates

<table>
<thead>
<tr>
<th>Type of Assay</th>
<th>No. samples screened</th>
<th>% Assay fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abi Identifilier Kit - AMEL Marker</td>
<td>768</td>
<td>1.30</td>
</tr>
<tr>
<td>Gender based PCR - AMEL marker</td>
<td>2750</td>
<td>0.62</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total number of individuals screened to confirm gender</th>
<th>% Gender error across collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>3415</td>
<td>1.82</td>
</tr>
</tbody>
</table>