SERVICES & FACILITIES ANNUAL REPORT - FY April 2018 to March 2019

SERVICE	FUNDING	AGREEMENT	ESTABLISHED as S&F	TERM
NERC Biomolecular Analysis Facility, NBAF	PAYG & Block	F14/G6/48 (NBAF–B: R8/H10/61)	1998 (NBAF–S) 2005 (NBAF–E & NBAF–L) 2009 (NBAF–B)	N/A

TYPE OF SERVICE PROVIDED:

NBAF provides fully-supported access to a wide range of advanced molecular genetic, genomic, metabolomic and bioinformatic technologies to the UK environmental science community in order to address ecological or evolutionary questions. Service is provided at four nodes: Birmingham (NBAF–B: metabolomics), Edinburgh (NBAF–E: sequencing and bioinformatics), Liverpool (NBAF–L: sequencing and bioinformatics) and Sheffield (NBAF–S: genotyping and population genetics). Access to the Facility is organised centrally through competitive peer-reviewed proposals that are assessed by the independent members of the Steering Committee (application form at http://www.nbaf.nerc.ac.uk) to ensure that (i) only the best science is supported, (ii) access to more than one node is coordinated, and (iii) projects are followed through to dissemination of the results. Each node is embedded in a well-equipped and vigorous research environment that, together with continuing developments in equipment and training, makes 'state-of-the-art' services available to the NERC community. NBAF provides access to high-level capability, and the associated training, that are rarely available elsewhere. NBAF–B supports metabolomic analyses using both mass spectrometry and NMR methods. NBAF–S is equipped to train and supervise researchers (mostly PhD students) in their own analyses at the bench. At NBAF–S most studies require metabarcoding (for biodiversity, dietary or microbiome analyses) or genotyping (microsatellites, single nucleotide polymorphisms [SNPs], target amplicons or telomeres) methods. Samples are prepared and analysed using Illumina, PacBio or ABI sequencers, SNP typing or qPCR platforms. NBAF-E supports whole-genome, transcriptome and reduced-representation sequencing, including targeted resequencing, amplicon sequencing and genotyping-by-sequencing, at any scale, using short-read (Illumina NovaSeq, HiSeqX and MiSeq) and long-read (PacBio Sequel, ONT Promethion) instruments. NBAF–L offers sequencing, gene expression and bioinformatic services, particularly for environmental diversity through amplicon sequencing on long-read (PacBio) platforms, targeted resequencing of exons and reduced genomic regions, and gene expression analysis on short-read (Illumina) platforms. All NBAF nodes provide integrated experimental design, data generation and bioinformatic analysis; NBAF–L also offers statistical and network-based interpretation of results. NBAF provides a wide diversity of training courses in genomics and metabolomics, as well as project-specific wet lab and bioinformatics training. Access to NBAF–E and NBAF–L is "pay-as-you-go" (PAYG), to support grant-funded projects, but NBAF coordinates an annual, fully-funded Pilot Project scheme to introduce NERC researchers to next-generation sequencing approaches in ecological and environmental science.

ANNUAL TARGETS AND PROGRESS TOWARDS THEM

Capacity is defined by the availability of staff time, and all four nodes make >85% of funded staff time available to users, with any remainder allocated to R&D. Projects have run largely as scheduled, except where users have failed to submit adequate samples on time.

SCORES AT LAST REVIEW (each out of 5)		Date of Last Review:		2011
Need	**Uniqueness**	**Quality of Service**	**Quality of Science & Training**	**Average**
5	4.5	5	5	4.88

CAPACITY of HOST ENTITY FUNDED by S&F %	Staff (Grade, FTE): NBAF–S: DA Dawson (G7 100%), G Horsburgh (G7 100%), H Hipperson / K Maher (G7 100%), C Pagnier (G5 80%), R Tucker (G6 20%); NBAF–B: D Varshavi (G7 100%); NBAF–E: U Trivedi (UOE7 50%), S Wardlow (UOE5 100%), M Arnot (UOE06 15%) R Talbot (UoE7 30%); NBAF–L: K Jackson (Res 9 20%), L Parsons (Cler 5 30%), R Gregory (Res 7 20%), J Kenny (Res 8 10%), C Owen (Res 6 100%), X Lui (Res 7 80%).	Next Review (March)	Contract Ends (31/3)
		TBC	2021

FINANCIAL DETAILS: CURRENT FY						
Total Resource Allocation	**Unit Cost £k**			**Capital Expend**	**Income £k**	**FCC**
	Unit 1	**Unit 2 Variable**	**Unit 3**	4.94		£2,027.8
FINANCIAL COMMITMENT (by year until end of current agreement) £k						
2018-19	**2019-20**	**2020-21**				

STEERING COMMITTEE	Independent Members	Meetings per annum	Other S&F Overseen
NBAFSC	Chair +5	1-2	None (one SC for 4 nodes)

APPLICATIONS: DISTRIBUTION OF GRADES (current FY — 2018/19)													
	10	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**R***	**Pilot**
NERC Grant projects*	1	1	8	1	0	0	0	0	0	0	0	0	N/A
Other academic	0	0	0	1	0	0	0	0	0	0	0	0	16
Students	0	0	5	6	2	1	0	0	0	0	0	2	7
TOTAL	1	1	13	8	2	1	0	0	0	0	0	2	23
APPLICATIONS: DISTRIBUTION OF GRADES (per annum average previous 3 financial years — 2015/2016, 2016/2017 & 2017/18)													
	10	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**R***	**Pilot**
NERC Grant projects*	1	6.7	5	1.7	0	0	0	0	0	0	0	0	N/A
Other academic	0	0.3	2	2	1.3	0	0	0	0	0	0	0.3	16
Students	0	1	5.7	4	3.3	1	0.3	0	0	0	0	1.3	13
TOTAL	1	8	12.7	7.7	4.7	1	0.3	0	0	0	0	1.7	29

PROJECTS COMPLETED (current FY – 2018/19)												
	10	**9**	**8 (α4)**	**7**	**6 (α3)**	**5 (α2)**	**4**	**3 (α1)**	**2**	**1 (β)**	**0**	**Pilot**
NERC Grant projects*	4	17	36	2	0	0	0	0	0	0	0	0
Other Academic	0	0	1	3	0	0	0	0	0	0	0	14
Students	0	2	14	7	5	0	0	0	0	0	0	2

Project Funding Type (current FY – 2018/19) (select one category for each project)											
Grand Total	**Infrastructure**					**PAYG**					
	Supplement to NERC Grant *	**PhD Students**		**NERC Centre**	**Other**	**NERC Grant***	**PhD Students**		**NERC Centre**	**Other**
		NERC	**Other**				**NERC**	**Other**		
107	3	15	13	0	4	56	0	2	0	14
Project Funding Type (per annum average previous 3 financial years — 2015/2016, 2016/2017 & 2017/18)										
Grand Total	**Infrastructure**					**PAYG**				
	Supplement to NERC Grant *	**PhD Students**		**NERC Centre**	**Other**	**NERC Grant***	**PhD Student**		**NERC Centre**	**Other**
		NERC	**Other**				**NERC**	**Other**		
109	4.5	11.7	21	1.3	9	44.8	1.3	6.3	0	9

User type (current FY – 2018/19) (include each person named on application form)				
Academic	**NERC Centre**	**NERC Fellows**	**PhD Students**	**Commercial**
225	2	10	36	3
User type (per annum average previous 3 financial years — 2015/2016, 2016/2017 & 2017/18)				
Academic	**NERC Centre**	**NERC Fellows**	**PhD Students**	**Commercial**
215	18	9	32	1

OUTPUT & PERFORMANCE MEASURES (current year)										
Publications (by science area & type) (calendar year 2018)										
SBA	**ES**	**MS**	**AS**	**TFS**	**EO**	**Polar**	**Grand Total**	**Refereed**	**Non-Ref/ Conf Proc**	**PhD Theses**
0	2.3	12.8	0	76.5	0	3.3	95	63	15	17
Distribution of Projects (by science areas) (FY2018/19)										
Grand Total	**SBA**	**ES**	**MS**	**AS**	**TFS**	**EO**	**Polar**			
107	0	5.5	20	1	79.5	0	1			
OUTPUT & PERFORMANCE MEASURES (per annum average previous 3 years)										
Publications (by science area & type) (Calendar years 2015, 2016 & 2017)										
SBA	**ES**	**MS**	**AS**	**TFS**	**EO**	**Polar**	**Grand Total**	**Refereed**	**Non-Ref/ Conf Proc**	**PhD Theses**
0	0	23	0	96	0	4	123	78	24	21
Distribution of Projects (by science areas) (2015/2016, 2016/2017 & 2017/18)										
Grand Total	**SBA**	**ES**	**MS**	**AS**	**TFS**	**EO**	**Polar**			
109	0	3.3	20.5	0.6	81.9	0	3			

Distribution of Projects by NERC strategic priority (current FY 2018/19)							
Grand Total	**Climate System**	**Biodiversity**	**Earth System Science**	**Sustainable Use of Natural Resources**	**Natural Hazards**	**Environment, Pollution & Human Health**	**Technologies**
107	12	73	0.5	8	8	5	0.5

OVERVIEW & ACTIVITIES IN FINANCIAL YEAR (2018/19):

NBAF-B took the decision to close last year and therefore completed only previously approved projects in the current year. NBAF-B has continued to utilise the analytical and computational capabilities within Phenome Centre Birmingham, a human-health focused metabolomics centre. Progress has continued across NBAF-B's full service portfolio, including experimental design, metabolite extractions, metabolomics data acquisition, processing and statistical analyses, and metabolite identification. As of 31 March 2019, all but three projects had been completed, with the final three due to finish by 30 June 2019. Multiple training courses in metabolomics have continued to be provided to the community through the Birmingham Metabolomics Training Centre (BMTC). NBAF is having discussions with NERC on the possible renewed provision of metabolomics.

NBAF-E is delivered from within Edinburgh Genomics (EdGe), one of the largest genomics facilities in the UK. EdGe brings a decade of experience of next generation sequencing to support clinical, biomedical, agricultural, evolutionary and ecological science. NBAF users have access to a wide range of the most advanced sequencing technologies. The support offered extends from experimental design through to assistance with publication, and includes long experience with "difficult" – especially low quantity – sample sets. EdGe runs the new Illumina NovaSeq platform, a high-volume, low-cost sequencer that can be used for the full range of genomics applications including de novo genome sequencing, resequencing, targeted sequencing, RNASeq and many others. EdGe also has five Illumina HiSeqX and two Illumina MiSeq instruments. EdGe also offer PacBio Sequel long-read sequencing for genomics, amplicons and full-length transcriptomics, and has installed the new Oxford Nanopore Promethion long-read system explicitly for NERC science use. NBAF-E continues to offer Sanger sequencing in bulk. Larger projects are delivered using dedicated robotics, which are particularly used for RNA-Seq, double-digest RAD-Seq and multiplexed targeted selection sequencing. EdGe offers standard and bespoke support in bioinformatics analyses of high-volume data, and has a dedicated genomics and bioinformatics training strand accessible to NERC science, delivering 40 short courses reaching over 600 scientists.

NBAF-L has developed single-cell genomic methods on its 10X Chromium platform. These allow gene expression within each of thousands of cells to be individually identified. Using this we have explored the variety of different immune cell types present within wild rodent populations, including non-model species. The methods are more widely applicable to environmental genomics. We have also developed genome assembly and annotation pipelines for long-read (PacBio) or 10X gDNA data, combined with PacBio IsoSeq transcriptomic data. We are developing more powerful methods for the rapid identification of species diversity within bacterial and fungal communities. We are trialling an APOBEC method to assay genome-wide methylation patterns. We have developed rapid and cost efficient means to produce sequencing libraries from 96- or 384-well plates of DNA, which can be applied to sequencing of large bacterial cohorts or of identifiable individuals in population genomic applications. NBAF-L ran short courses in gene analysis and metagenomic analysis of communities, which were well received and over-subscribed.

NBAF-S continued to see the majority of its new projects requiring metabarcoding techniques, including dietary analysis in nightjars from faecal samples and in cichlid fish from stomach contents, biodiversity analyses of earthworms in agricultural landscapes, assessment of bacterial infection and parasite detection in birds and bats, and microbiome analysis of public water systems. Microsatellite-based genotyping applications included studying genetic relationships in buff-tailed bumblebees, long-tailed tits and monk parakeets, and population genetic analyses of ramshorn snails and pine martens. Other research supported included population structure studies of *Heliconius ethilla* butterflies and of Asiatic wild ass using genotyping by sequencing. These projects exploited our modern qPCR platform, a high-throughput QuantStudio 12k Flex Real-Time PCR system. This system is used to support high-throughput telomere analyses, in particular, and provides a efficient alternative to the node's LGC KASP system for some configurations of SNP typing. The system is also invaluable for precise sample quantification when preparing libraries for multiplexed high-throughput sequencing. The host lab completed several studies using its PacBio Sequel, and this will be available to NBAF-S users for amplicon sequencing (especially MHC typing). NBAF-S ran two courses this year on metabarcoding. NBAF-S's population genomics course was again highly over-subscribed, with over 80 applicants applying for the 30 available places, and was again very positively received by those who could be accommodated.

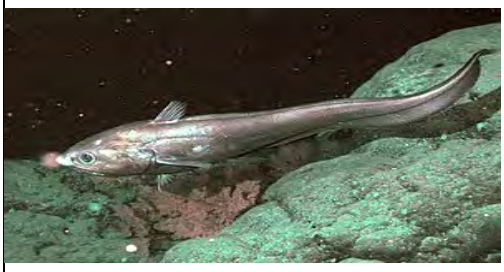
NBAF funded 13 Pilot Grants projects in October 2018, enabling new users (including PhD projects), to gain access to NBAF-E and NBAF-L, which are otherwise only accessible via PAYG.

SCIENCE HIGHLIGHTS

In 2018-19 NBAF serviced 107 projects, including 59 supported by NERC grants; ran 20 courses; supported 225 academics and 36 students; and in 2018 published 63 refereed articles.

Genomics of habitat choice and adaptive evolution in a deep-sea fish (Gaither *et al.* 2018 *Nature Ecology and Evolution* **2**, 680-6870) *NBAF generated the first de novo genome for this species. This involved wet-lab and*

bioinformatic development of assembly methods by NBAF.

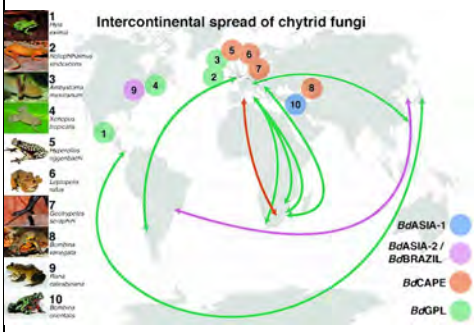


The deep-sea environment is extremely poorly understood. It is a diverse habitat that is host to animal species that are rarely seen in their natural habitat. One large deep-sea fish, the roundnose grenadier (*Coryphaenoides rupestris*) can live at depths ranging from 180–2600 m. The study showed genetic differentiation among populations living at different depths and identified the genes involved. Sixty-nine genes exhibited evidence of differentiation and had functions suggested to be associated with metabolism and morphogenesis – which may point to the adaptations that have been selected for in these different habitats.

The ash dieback invasion of Europe was founded by two genetically divergent individuals (McMullan e*t al.* 2018 *Nature Ecology and Evolution* **2**, 1000-08) *NBAF sequenced and assembled whole genomes of 24 European ash isolates, often from minimal sample input.*

Accelerating international trade and climate change make pathogen spread an increasing concern. *Hymenoscyphus fraxineus*, the causal agent of ash dieback, is a fungal pathogen that has been moving across continents and hosts from Asian to European ash. Most European common ash trees (*Fraxinus excelsior*) are highly susceptible to H. fraxineus, although a minority (~5%) have partial resistance to dieback. Here, we assemble and annotate a H. fraxineus draft genome, which approaches chromosome scale. Pathogen genetic diversity across Europe and in Japan reveals a strong bottleneck in Europe, though a signal of adaptive diversity remains in key host interaction genes. We find that the European population was founded by two divergent haploid individuals. Divergence between these haplotypes represents the ancestral polymorphism within a large source population. Subsequent introduction from this source would greatly increase adaptive potential of the pathogen. Thus, further introgression of *H. fraxineus* into Europe represents a potential threat and Europe-wide biological security measures are needed to manage this disease.

Recent Asian origin of chytrid fungi causing global amphibian declines (O'Hanlon *et al.* 2018 *Science* **360**, 621-627) *NBAF generated the genome resequencing data.*



Globalized infectious diseases are causing species declines worldwide, but their source often remains elusive. We used whole-genome sequencing to solve the spatiotemporal origins of the most devastating panzootic to date, caused by the fungus *Batrachochytrium dendrobatidis*, a proximate driver of global amphibian declines. We traced the source of *B. dendrobatidis* to the Korean peninsula, where one lineage, BdASIA-1, exhibits the genetic hallmarks of an ancestral population that seeded the panzootic. We date the emergence of this pathogen to the early 20th century, coinciding with the global expansion of commercial trade in amphibians, and we show that intercontinental transmission is ongoing. Our findings point to East Asia as a geographic hotspot for *B. dendrobatidis* biodiversity and the original source of these lineages that now parasitize amphibians worldwide.

The decline of the turtle dove: dietary associations with body condition and competition with other columbids analysed using high throughput sequencing (Dunn *et al.* 2018 *Molecular Ecology* **27**, 3386-3407) *NBAF supported the metabarcoding work and data analyses.*

Dietary changes linked to the availability of anthropogenic food resources can have complex implications for species and ecosystems, especially when species are in decline. Here, we use recently developed primers targeting the ITS2 region of plants to characterize diet from faecal samples of four British columbids, with particular focus on the European turtle dove (*Streptopelia turtur*), a rapidly declining obligate granivore. We show considerable change in columbid diets compared to previous studies, probably reflecting opportunistic foraging behaviour by columbids within a highly anthropogenically modified landscape. Nestling turtle doves in better condition had a higher dietary proportion of taxonomic units from natural arable plant species and a lower proportion of taxonomic units from anthropogenic food resources such as garden bird seed mixes and brassicas. This suggests that breeding ground conservation strategies for turtle doves should include provision of anthropogenic seeds for adults early in the breeding season, coupled with habitat

rich in accessible seeds from arable plants once chicks have hatched.

Inclusive fitness consequences of dispersal decisions in cooperatively breeding bird, the long-tailed tit (Green *et al* 2018 *Proceedings of the National Academy of Sciences* **115**, 12011-12016) *NBAF supported the microsatellite genotyping and the genetic analyses.*

Natal dispersal is a demographic trait with profound evolutionary, ecological, and behavioral consequences. However, our understanding of the adaptive value of dispersal patterns is severely hampered by the difficulty of measuring the relative fitness consequences of alternative dispersal strategies in natural populations. This is especially true in social species, in which natal philopatry allows kin selection to operate, so direct and indirect components of inclusive fitness have to be considered when evaluating selection on dispersal. Here, we use lifetime reproductive success data from a long-term study of a cooperative breeder, the long-tailed tit *Aegithalos caudatus*, to quantify the direct and indirect components of inclusive fitness. We show that dispersal has a negative effect on the accrual of indirect fitness, and hence inclusive fitness, by males. In contrast, the inclusive, predominantly direct, fitness of females increases with dispersal distance. We conclude that the conflicting fitness consequences of dispersal in this species result in sexually antagonistic selection on this key demographic parameter.

Life in the intertidal: cellular responses, methylation and epigenetics (Clark *et al.* 2018 *Functional Ecology* **32**, 1982-1994) *NBAF supported the wet lab work and.*

Phenotypic plasticity is essential for the persistence of organisms under changing environmental conditions but the underlying cellular and genetic mechanisms remain unclear. In the Antarctic limpet *Nacella concinna*, two distinct phenotypes are associated with the intertidal and subtidal zones. The in situ gene expression and methylation profiles of intertidal and subtidal cohorts were directly compared before and after reciprocal transplantation as well as after a common garden acclimation to aquarium conditions for 9 months. Expression profiles showed significant modulation of cellular metabolism to habitat zone, with the intertidal profile characterised by transcription modules for antioxidant production, DNA repair and the cytoskeleton reflecting the need to cope with continually fluctuating and stressful conditions including wave action, UV irradiation and desiccation. The subtidal animals transplanted to the intertidal zone modified their gene expression patterns towards that of an intertidal profile. In contrast, many of the antioxidant genes were still differentially expressed in the intertidal animals several weeks after transplantation into the relatively benign subtidal zone. Furthermore, a core of genes involved in antioxidation was still preferentially expressed in intertidal animals at the end of the common garden experiment. Thus, acclimation for 9 months did not completely erase the intertidal gene expression profile. Significant methylation differences were detected between intertidal and subtidal animals from the wild and after transplantation, which were reduced on common garden acclimation. This suggests that epigenetic factors play an important role in physiological flexibility associated with environmental niche.

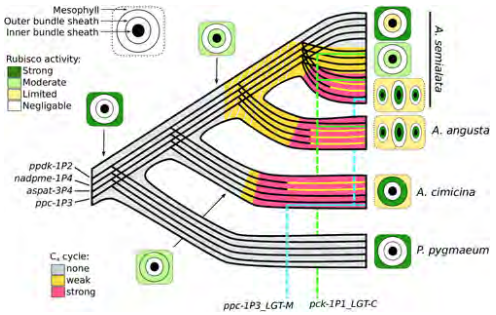
Genomic associations with bill length and disease reveal drift and selection across island bird populations (Armstrong *et al*. 2018 *Evolution Letters* **2**, 22–36) *NBAF supported the ddRAD wet lab work and sequencing.*

Island species provide excellent models for investigating how selection and drift operate in wild populations, and for determining how these processes act to influence local adaptation and speciation. Here, we examine the role of selection and drift in shaping genomic and phenotypic variation across recently separated populations of Berthelot's pipit (*Anthus berthelotii*), a passerine bird endemic to three archipelagos in the Atlantic. Genetic diversity and population structuring supported previous inferences of a history of recent colonizations and bottlenecks. We then tested for regions of the genome associated with the ecologically important traits of bill length and malaria infection, both of which vary substantially across populations in this species. We identified a SNP associated with variation in bill length among individuals, islands, and archipelagos; patterns of variation at this SNP suggest that both phenotypic and genotypic variation in bill length is largely shaped by founder effects. Malaria was associated with SNPs near/within genes involved in the immune response, but this relationship was not consistent among archipelagos, supporting the view that disease resistance is complex and rapidly evolving. Genome scan analyses pointed to several genes related to immunity and metabolism as having important roles in divergence and adaptation. Our findings highlight the utility and challenges of combining association mapping and population genetic analysis to disentangle the effects of drift and

selection on genotypes and phenotypes.

Gene duplication and dosage effects during the early emergence of C4 photosynthesis in the grass genus Alloteropsis. (Bianconi *et al.* 2018 *Journal of Experimental Botany* ***8,*** 1967–80)

The importance of gene duplication for evolutionary diversification has been mainly discussed in terms of genetic redundancy allowing neofunctionalization. In the case of C4 photosynthesis, which evolved via the co-option of multiple enzymes to boost carbon fixation in tropical conditions, the importance of genetic redundancy has not been consistently supported by genomic studies. Here, we test for a different role of gene duplication in the early evolution of C4 photosynthesis, via dosage effects creating rapid step changes in expression levels. Using genome-wide data for accessions of the grass genus Alloteropsis that recently diversified into different photosynthetic types, we estimate gene copy numbers and demonstrate that recurrent duplications in two important families of C4 genes coincided with increases in transcript abundance along the phylogeny, in some cases via a pure dosage effect. While increased gene copy number during the initial emergence of C4 photosynthesis likely offered a rapid route to enhanced expression, we also find losses of duplicates following the acquisition of genes encoding better-suited isoforms. The dosage effect of gene duplication might therefore act as a transient process during the evolution of a C4 biochemistry, rendered obsolete by the fixation of regulatory mutations increasing expression levels.



FUTURE DEVELOPMENTS/STRATEGIC FORWARD LOOK

Assembling and annotating reference quality genomes: In the last year, momentum has gathered for the Earth Biogenome Project, a concerted, global push to generate high-quality reference or near-reference quality genomes for all species on Earth. Largely aspirational at this point, the EBP has a few important funded collaborating projects, including the Vertebrate Genomes Project (aiming to generate reference genomes for all orders, and then all families and ultimately all species of vertebrate), the BGI-led 10K Plants project, and local initiatives on other txa. In the UK, the Wellcome Sanger Institute has signalled its intent to initiate a Darwin Tree of Life (DToL) project, which aims to sequence all British eukaryotes (60,000 species) in the next 10 (or so) years. The EBP and DToL projects will impact directly on NERC science and NBAF activity. NBAF nodes have been working to be included in the DToL, and current node director Blaxter is taking up the position of lead on the Tree of Life programme at Sanger.

It is unclear as yet how NERC (and indeed UKRI) will contribute to the DToL programme. The BBSRC-supported Earlham institute has been funded to start undertaking DToL activities, but no DToL-related funding has yet been awarded via NERC. If NERC does engage with DToL it will represent an opportunity for NBAF, since the availability of reference genomes will allow investigators to perform genomic studies (RNAseq, GWAS, etc) on any organism present in the UK. The current genome sequencing technology for DToL involves high-fold coverage in long reads from PacBio SEQUEL2 or Promethion, with “polishing” through Chromium 10X read cloud data, and annotation via RNAseq. All of these data types are available through NBAF, but the Facility currently does not yet operate the new high-volume SEQUEL2. A capital bid by NBAF-L is planned to address this.

For species targeted by research programmes within NERC science, additional data from genetic linkage mapping can add final chromosomal contiguity. NBAF nodes already offer RAD-seq and other genotyping-by-sequencing technologies for cost-effective linkage mapping of new species. NBAF also offers cost-effective resequencing for population genomics of wild species and ecosystems, and ChipSeq, RNAseq and other analyses for post-genomic interrogation of organismal biology.

Long-read single molecule data: Promethion and PacBio SEQUEL2. The future sequencing technology landscape continues to evolve rapidly. While long-range genomic data can be generated from Illumina platforms using Chromium 10X and chromatin conformation capture (HiC) approaches, a major strand in long-range data is long reads. Pacific Biosciences have released their SEQUEL2 platform, which offers increased throughput (~60 Gb/run). They have also released library construction technologies that reduce input sample requirements and sequencing strategies that improve per-base read quality (approaching that of Illumina at 99.9% accuracy) but reduce total yield. The Oxford Nanopore Technologies (ONT) Promethion sequencer offers similar high throughput (>60 Gbase/run) in long reads (for example 30 kb N50). This technology, while not as mature as that of PacBio, has much promise, as it can deliver very long sequences (the record currently stands at 2 Mb) and is improving rapidly. NBAF-L have run PacBio instrumentation for 8 years, and NBAF-E for 3, and in 2019 NBAF-E is installing the ONT Promethion.

Excitingly, the nanopore systems offers the promise of direct RNA sequencing, and of detection of non-canonical bases in RNA and DNA (i.e. direct methylation detection). The NBAF Promethion service will go live in summer 2019. This

technology space is changing, and maturing, rapidly. NBAF will continue to offer emerging technologies in this area, seeking capital investment to install and prove new systems.

Short-read sequencing: capacity and flexibility. Developments in the landscape of short-read sequencing, dominated by Illumina technology, are driven by the market for cheap clinical genomics. This has a positive knock-on for NERC science, as NBAF-E and NBAF-L are embedded within larger facilities that have the throughput and the ambition to remain at the forefront of the ongoing revolution in clinical genomics. NBAF is therefore able to offer access to the latest in Illumina sequencing platforms, including especially the NovaSeq. The Novaseq is a redesign of the core Illumina platform, offering up to 2.5 billion read pairs and over 700 Gbase of data per 3-day run, and now offers smaller volumes of data similar to a lane of the previous HiSeq machines. Importantly, unlike the previous high-volume Illumina platform the HiSeqX, there are no restrictions on use of the NovaSeq for non-clinical genomic applications. The NovaSeq platform has been made available to NBAF collaborators since late 2017.

Gene expression profiling and epigenetics: Knowledge of the underlying genome sequence, and of the pattern of variation across populations, is only the beginning of understanding the environmental biology of a species. Detailed assessment of gene expression and its plastic and epigenetic modulation in the face of environmental challenge can be achieved by using sequencers (in particular the NovaSeq platform) as counting instruments. IsoSeq analysis on the PacBio Sequel is improving the characterisation of full-length RNA transcripts, including splice variants. The facilities will continue to develop new standard assays for single-cell transcriptome and epigenetic analysis. The 10X Chromium system at NBAF-L permits transcriptome profiling and chromatin accessibility of individual cells using droplet technology. NBAF nodes are collaborating to develop best-practice methods for analysis, and exploring the parameters of replicate numbers and other experimental design parameters that will maximise access to biological insight using these new technologies.

eDNA and metabarcoding: NBAF-S has a long history of using mitochondrial barcoding to detect and identity parasites (such as malaria strains) and analyse diets (e.g. moth predation by bats, used to describe habitat use). The advent of NGS, especially longer reads on the MiSeq, has made these kinds of studies ever more feasible. NBAF-S has greatly increased the implementation of eDNA and metabarcoding techniques over the last 2 years. NBAF-S has always fully supported many ecologists with little knowledge of molecular techniques, and the demand from this community (rather than experienced geneticists) is increasing, as the potential for these methods in describing distribution, biodiversity, food chains and ecological interactions continues to expand.

Population genomics: At NBAF-E, installation of the NovaSeq and its extension to non-human genome sequencing and resequencing brings powerful population genomics even more within the reach of NERC science. The NovaSeq offers radically reduced per-gigabase sequencing costs. We are continuing to develop new applications on this new instrument. NBAF-S has now transferred most of its sequencing needs to Illumina platforms, especially the MiSeq, as it provides read-lengths of up to 600 bp, so matching the read-length on the ABI platforms used previously. All the NBAF-S genotyping development work is currently conducted on the MiSeq platform. The use of multiplexed tags within single MiSeq sequencing libraries for microsatellite analyses is becoming standard and is still being trialled in the lab for wildlife applications. We have found there to be costs as well as benefits: it can be difficult to move established markers to the new platform and the required scale of experiments can be inefficient in the use of consumables; this issue will be revisited now that we have access to nano runs on the MiSeq. So while most NBAF genotyping has moved to Illumina, KASP or qPCR platforms, the ABI platform is still the most appropriate and economical for a subset of projects.

Pilot grants scheme: The scheme will be funded again in 2019–20. It is the only means through which non-grantholders are able to access NBAF-E and NBAF-L, and a key avenue to both educating our research community and trialling novel and potentially disruptive technological advances. The cost is relatively small, so that it continues to represent excellent value for money.