**Daphne Jackson Trust Fellowship application, Oct 2018**

**Applicant:** Dr. Elizabeth Hughes

**Supervisor:** Dr. Edward Wallace

**Host Organization:** Institute for Cell Biology, School of Biological Sciences, Edinburgh University

**Project Title:**  Dynamic mRNA processing in response to environmental stimuli in the fungal pathogen *Cryptococcus neoformans*.

**Reasons for applying for a Fellowship.**

During my undergraduate I was drawn to molecular biology and how it relates to the pathogenesis of disease. My PhD looked at the infecting HIV viral population and its relationship to disease progression, primarily the evolutionary analysis of isolates infecting lymphoid and non-lymphoid tissues 1,2. This led to the discovery of a previously unknown dormant HIV population in the brain and a first author paper with over 100 citation 1.

Subsequently, I examined the replicative processes of HCV by determining whether the NS5B protein (predicted to possess an RdRp activity) was capable of directing HCV replication. I optimized bacterial expression systems and purified the NS5B fusion protein.

Following this I investigated the structure and function of the major outer membrane proteins (MOMP’s) of Chlamydia where I cloned and expressed wild type and VS4 domain mutated proteins and functionally reconstituted them at the single-channel level. Reconstitution in planar lipid bilayers showed that the VS4 domain was not required for pore formation but may help to form the channel vestibule where it may interact with other protein loops 3.

I then joined a leading bio safety testing facility as the research and development scientist and developed a wide range of molecular based assays for clients, and provided technical training and support for colleagues.

I chose to take a career break to raise my family. My husband works in the marine industry as a consultant in risk assessment/safety management and this job takes him away from home on a regular basis. Taking this into account and the prohibitive costs of childcare we decided it would be best if I put my career on hold to raise our family. During this time I have taken over the administrative side of our business.

Both of my children now attend high school and I am confident that they are mature and resilient enough for me to return to my chosen career. I am excited to re-engage with the scientific community. I find the challenges associated with research both rewarding and enlightening and I look forward to returning to the lab environment.

The Daphne Jackson Fellowship with MRS presents an ideal platform for me to resume my biomedical research career. The provision of mentoring, support and retraining are invaluable. This fellowship will give me the opportunity to update my existing skills while learning new invaluable ones to help me develop as a person and move my career forward.

Word count: 398 (400)

**Research Summary**

I will investigate the dynamic processing of mRNA and resulting gene expression profiles of *C. neoformans* in response to environmental stimuli. *C. neoformans* is a fungus that lives in the environment and people become infected after breathing in spores or desiccated yeast cells. Infection is rare in healthy individuals and most cases occur in individuals who have a weakened immune system. Infection of the lung causes a pneumonia-like illness, however, if it spreads to the brain life-threatening meningitis can occur and it is estimated that 223 000 cases occur per annum world-wide resulting in approximately 181 100 deaths.

**Aim 1:** How does the rapid change in environment from soil/vegetation to a mammalian lung affect the gene regulation of *C. neoformans*? I would like to examine what happens when this fungus reactivates in a host environment alien to its normal life cycle.

**Aim 2**: Upon presentation to the lung *C. neoformans* will encounter a number of foreign bodies, some of which will be bacteria that naturally reside within the lung. I would like to investigate the interaction with bacterial proteins to examine what effect this could have on the morphology and pathogenicity of *C. neoformans*. Interactions with other microbes may have a synergistic or antagonistic impact on fungal pathogenesis.

**Aim 3**: In the mammalian lung the first line of defense is the mucosal lining of the airways. This contains soluble effector molecules which help eradicate foreign bodies. One such molecule is SP-D which has anti-microbial and immunomodulatory properties. SP-D will be one of the first effector molecules to interact with this pathogenic fungus and has been shown to help with its survival. I would like to investigate the binding of this molecule to *C. neoformans* to elucidate what changes occur in the gene expression during this event.

I will measure mRNA processing using qPCR and further analyze with RNA-seq. Measuring the intermediate step between genes and proteins effectively bridges the gap between the genetic code and functional proteins. The number of transcripts can be quantified to provide information on the amount of gene activity under a given set of circumstances.

Word count: 349 (350)

**Abstract:**

*C. neoformans* is an accidental pathogen of mammals. Its natural life cycle is that of an environmental saprophyte. The mammalian lung is an alien environment yet somehow it can circumvent a myriad of defenses to cause disease. Upon inhalation into a host *Cryptococcus* cells transition from spores/desiccated yeast into actively replicating virulent yeast cells. This transformation must be associated with dynamic regulation of mRNA and hence gene expression. What happens to *Cryptococci* when they reactivate within a mammalian host? I would like to investigate this phenomenon in detail during the early stages of infection. I will use quantitative and qualitative methods to examine differential gene expression of *Cryptococcus* under different environmental stimuli. Gene expression profiles represent a snapshot of cellular metabolism at the molecular level and may provide insights into novel therapeutic methods for drug discovery.

Word Count: 136 (150)

**Host Organization**

The University of Edinburgh is one of the world’s top universities, ranked 18th in the world, 5th in the UK and the top university in Scotland. As one of the UK’s leading research facilities Edinburgh University provides a state-of-the-art environment for cellular and molecular biology.

The Centre for Synthetic and Systems Biology, in the Institute for Cell Biology, is a unique inter-disciplinary environment with a track record for multi-disciplinary research. Included is The Genome Foundry (a world class facility for automated DNA design and assembly) and EdinOmics providing expertise in quantitative biochemistry including up-scaling and automation of qPCR. There are facilities for cell imaging (LEAP and Single cell analysis and microscopy groups) and facilities for innovative data analysis and mechanistic modelling.

Dr. Edward Wallace, of the institute for Cell Biology in Edinburgh University, has agreed to be my supervisor for the duration of the fellowship. Dr. Wallace in a renowned RNA scientist and working with him will allow me to build on my existing skills in the field of molecular biology but also learn new techniques including high throughput qPCR, high throughput sequencing (RNA-seq) and library preparation, lab automation, bioinformatics (using R-programming) and working with large data sets. This will provide me with crucial work experience in one of Scotland’s world-class universities and equip me with specific skills in high demand in biomedical research and biotech industries. Rosey Bayne (PDRA/Lab Manager) has 30 years of experience in molecular biology techniques and will be invaluable during my re-training and introduction to new concepts and technologies.

Edinburgh University supports academic and personal development through the Institute for Academic Development. This program provides support, learning and research development opportunities throughout the year including workshops, courses, online resources, networking and advice. These include workshops and courses on academic writing, career management and development, data management, funding opportunities, ethics, teaching and supervising students. During my fellowship I intend to make full use of these resources to enhance my personal and professional qualifications to help advance my career.

**Supervisor:**

Dr. Edward Wallace.

Sir Henry Dale Fellow (Wellcome Trust/Royal Society Early Career Research Fellow).

Institute for Cell Biology, School of Biological Sciences, Edinburgh University.

Project funding remit: The lab is funded by Wellcome Trust/opal Society. Innovative research on fungal pathogens falls under the MRC’s health strategic aim.

**Relevant collaborations:**

Word Count:382 (400)

**Section 5: Retraining program**



**Preparation of cDNA libraries:**

I will prepare cDNA libraries from mRNA extracted from yeast cells grown under the variable conditions described above. I will use a recently described multi-plex method to prepare cDNA libraries 4 for RNA-seq analysis. cDNA libraries provide a snapshot of actively expressed genes and can be used to determine what circumstances trigger expression of various genes.

Workflow:

* Experimental set up
* Extraction of total RNA from sample using the QIAgen plant and fungal RNA extraction kit and QIAgen RNA mini plant kit according to the manufacturer’s instructions..
* Assessment of RNA concentration and quality.
* cDNA Synthesis through reverse transcription using Superscript VI for qPCR analysis.

Random primers will be used to ensure reverse transcription of the entire length of the RNA.

**Quantitative real-time RT-PCR (qPCR)**:

I will use SYBR-green fluorescence based qPCR to analyze gene expression in response to environmental stimuli. qPCR is one of the most powerful and sensitive gene expression analysis techniques available and can be used to perform accurate high-throughput mRNA quantification over a wide dynamic range. Examining gene expression responses to environmental stimuli by qPCR provides a deeper understanding of the molecular mechanisms underpinning physiological change.

Workflow of gene expression analysis using RT-qPCR:

* Experimental design.
* cDNA template preparation for qPCR.
* qPCR to measure expression levels of target genes.
* Data analysis using appropriate normalization methods.

**RNA sequencing and quantitative analysis (RNA-seq):**

RNA-seq uses next-generation sequencing technology to determine the presence and quantity of RNA in a biological sample at a given moment and is very useful to analyze changes in the transcriptome. This technique is able to detect known and un-known features setting it apart from other technologies as this enables novel RNA’s to be discovered without prior knowledge of the sequence. I will use Illumina’s Truseq RNA-seq which is a random primed cDNA synthesis non-stand specific protocol to quantify the changing expression levels in yeast cells under different conditions.

Workflow of gene expression using RNA-Seq:

* Experimental design.
* cDNA template preparation for sequencing from cDNA libraries prepared.
* Sequencing of cDNA library on the XXX instrument.
* Data analysis of resulting short read sequences. I will use xxx to analyze my data.

**Microscopy**:

I will use microscopy to visualize cells before and after exposure to stimuli to correlate morphology with gene expression.

**Bioinformatics and Data Analysis**:

With the advancement of new techniques like qPCR and RNA-seq large data sets are an integral part of the process. Therefore, bioinformatics is becoming an essential tool to interpret and understand the vast quantities of data produced and to extract biological relevance. A number of bioinformatics tools exist process and interpret experimental data.

I will use xxx to analyze qPCR data and YYY to analyze RNA-seq data.

**Lab Automation**:

Lab automation is an important component of the modern lab and is imperative to timely research and data output. There has been a great deal of advancement in instrumentation and methodology to increase the efficiency and effectiveness of research in the lab. For example high throughput screening possible with qPCR and RNA-seq.

Edinburgh University is in the enviable position of having the largest DNA manipulation facility in the Genome foundry specializing in automation.

modular cloning for sysnthetic biology

CRISPR transformation

High throughput sequencing library preparation

Lab automation

Word Count: (400)

**Section 6: Proposed Research plan.**

**Background:**

*C. neoformans* is an opportunistic and facultative pathogen and the causative agent of cryptococcosis. 5,6. *C. neoformans* primarily infects immunocompromised individuals and is one of only a few fungal species that have been shown to cross the blood-brain barrier leading to *cryptococcal* meningitis that is fatal if left untreated 5,6. A global number of 223 000 cases of *cryptococcal* meningitis are estimated to occur annually, with over 70% of cases in sub-Saharan Africa, and an estimated 181 100 resulting deaths 7. Park et al., 2009; estimated the global burdens much higher at almost 1 million global cases and 600 000 deaths.

*C. neoformans* is found in the environment typically associated with pigeon guano, soil and decaying wood. It is a free living fungus with no requirement for mammalian virulence, however, it is able to adapt, survive and proliferate within a mammalian host to cause disease. Both the basidiospore and desiccated encapsulated yeast cells are postulated to act as infectious propagules, and inhalation of these is the primary route of infection.

The natural ecology and disease progression of *C. neoformans* is well characterized, however, few studies have examined in detail the differential gene expression and regulation of *C. neoformans* during early stages of infection. *C. neoformans* must undergo rapid changes in gene expression upon presentation within the alien environment of the lung. Indeed not many fungal pathogens can grow at 37 ⁰C, a characteristic virulence factor of *C. neoformans* and consistent with its role as a human pathogen8. The ability to investigate the genetic response to environmental stimuli is a powerful tool to elucidate the adaptive response/responses required for this accidental pathogen to survive in a hostile environment. I would like to examine in detail what happens when this organism reactivates within this alien environment.

**Aim 1: Determine environmental influences on gene expression in *C. Neoformans* in an *in vivo-like* system.**

Current knowledge of the early events in *C. neoformans* infection are based on research using animal models or *in vitro* culture methods primarily in rich media such as YPD or YPDA. While these culture methods have produced useful data and much has been extrapolated about the infectivity and virulence of *C. neoformans,* they do not accurately reflect the lung environment, where nutrients are likely to be scarce.

**Objective:**

I would like to design an *in vitro* system which more closely mimics *in vivo* conditions. I will establish an inoculation protocol using minimal media in a controlled environment to assess the response of *C. neoformans* to different environmental stimuli. Analysis of differential gene expression over a time-course will produce a snapshot of actively expressed genes and may elucidate the important steps for infection at early time-points. For example, modification of cell wall composition and structure will likely play a crucial step in the initial infection and establishment of *C. neoformans* in a mammalian host.

This work will carry on from investigations previously carried out in Dr. E Wallace’s lab in collaboration with Dr. Elizabeth Ballou at the University of Birmingham designed to dissect the contributions of host factors and temperature in shaping initial growth. In this pilot study a distinct physiological response and differential RNA abundance was documented between the different conditions tested.

**Method:**

I will utilize a culture medium more consistent with the host lung for *C. neoformans* growth prior to re-seeding and examine further environmental stimuli present in the lung environment, such as pH, ionic strength, divalent cations, mucins and nutrient limitation which may have synergistic or antagonistic effects on anti-fungal activity.

I will examine fresh and fixed cells to determine any physiological/morphological changes under each condition. I will extract RNA for further qualitative and quantitative analysis using qPCR and RNA-seq.

**Aim 2**: **Investigate how bacteria influence gene expression in *C. neoformans*.**

In many environments bacteria and fungi coexist and interact in nature competing for space and nutrients. Similar cross-species interactions are expected to occur between the wide ranges of micro-organisms that constitute the human microbiome 9,10.

It has been previously shown that bacterial species can interact with human fungal pathogens.11. Bacteria can stimulate spore germination in several fungi 12–14 and it has been shown that germination of *C. neoformans* is essential for survival in phagolysosomes 15. Bacterial small molecules such as metabolites and extracellular molecules have been shown to affect the morphology and pathogenicity of fungi 16–18. Previously bacterial cell wall components have been shown to modify the morphology of *C. neoformans* from a normal yeast cell to a titan cell 19 Mention previous work done on titan cells here?

**Objective:**

I hypothesize that interaction with bacteria may influence the infectivity and virulence of *C. neoformans* during the early stages of infection. I would like to examine the effect bacteria, residing in the lung, have on infecting yeast or spores.

**Method:**

I will incubate yeast cells with sera containing bacterial components (human, murine, FBS???) to examine the gene expression. I will use microscopy of live and fixed cells to look at the morphology of these cells and correlate this with possible up/down-regulation of cell surface modifications that may enhance infection of this pathogenic fungi. I will extract RNA for further qualitative and quantitative analysis using qPCR and RNA-seq.

**Aim 3**: **Investigate how component of the innate immune system influence gene expression in *C. neoformans*.**

Following inhalation fungi will first encounter the mucosal surface of the lung which is covered in a film of surfactant that maintains surface tension during respiration and contains soluble effector molecules with antimicrobial properties such as the collectins and host-defense peptides.

Collectins (SP-A, SP-D and MBL’s) have previously been shown to bind to *C. neoformans* 22–26. *In vitro* studies suggest SP-A can bind acapsular and capsular yeast cells but with a reduced affinity compared to SP-D and *C. neoformans* was shown to be resistant to SP-A binding once capsule synthesis had been induced 27. In contrast SP-D binds acapsular yeast with a high affinity (and capsular yeast with a lower affinity) causing profound aggregation, increased phagocytosis and enhanced fungal survival 23,24,28,29.

Previous studies, using animal models, have suggested a protective role for SP-D in *C. neoformans* infection 30. In contrast, other studies suggest that SP-D mediates host protection by opsonizing fungi prior to presentation to phagocytes (ref).

**Objective:**

I would like to investigate if SP-D binding affects *C. neoformans* directly by analyzing fungal gene expression before and after binding of purified recombinant SP-D. Direct binding of SP-D may modulate gene expression of surface receptors in a synergistic or antagonistic way and could therefore reciprocally modulate virulence factors, such as capsule production/cell wall rearrangements, thereby enhancing fungal survival. Determining whether binding of SP-D has a direct effect on pathogenicity of *C. neoformans* could pave the way for possible new drug targets aimed at de-regulating virulence instead of directly killing the fungus.

**Method:**

I will prepare recombinant SP-D protein and incubate this with yeast cells in a microtiter plate at various concentrations. I will analyze the plates for agglutination of *Cryptococci* and extract RNA for further qualitative and quantitative analysis using qPCR and RNA-seq.

For each aim I will use a variety of high throughput methods (qPCR and RNA-seq) and functional profiling to identify enriched pathways during the early phase of *C. neoformans* adaptation to the host lung environment. I will analyze the transcriptome of *C. neoformans* over a time-course and examine the transcriptional response to identify the genes that are transcribed under these selective environmental pressures which may lead to micro-evolutionary adaptation of *C. neoformans* in the host lung.

**Challenges:**

Going from yeast cultures to gene expression results involves a multi-step approach including harvesting cells, RNA isolation, removal of genomic DNA, cDNA synthesis and qPCR/RNA-seq. This multi-step approach can introduce inter- and intra-variation which must be normalized in order to make sense of the data produced.

During culture of yeast cells under different experimental conditions the amount of mRNA produced may vary considerably. In order to control for this inter-sample variation I will normalize RNA levels by spiking in a 1:100 ratio of methanol-fixed *Schizosaccaromyces pombe*. This will allow me to compare the effects of environmental stimuli in a qualitative and quantitative manner.

The reverse transcription step of cDNA synthesis has been proposed as the source of the most variability in qPCR experiments. I will use an external RNA control to normalize for external global mRNA changes by spiking extracted RNA samples with equal amounts of control RNA prior to cDNA synthesis. This will allow me to control for the presence of inhibitors of RT and DNA polymerase.

Due to the sensitivity of qPCR it is inevitable that some experimental variation will occur. Normalization of data is especially important when samples are obtained from different conditions and time-courses and will can result in misrepresentation of the expression profiles obtained 31. Therefore, it is essential to normalize target gene expression levels to compensate for intra- and inter- qPCR variations in order to identify real gene variation. This can be done by carefully selecting endogenous/reference controls. I will use 3-4 stably expressed reference control genes and the geometric mean of these will be taken. These will allow me to normalize differences in the amount and quality of starting material as well as in the efficiency of the reaction. It is common to use ubiquitously expressed genes.

Manipulation during the construction of the cDNA library can complicate the analysis of RNA-seq reads where PCR artefacts of short identical reads can be confused for genuine reflection of the RNA present. In order to overcome this limitation I will run 2 or 3 biological replicates and determine whether the same sequences are observed in each sample.

Normalization is an important prerequisite for any quantitative data analysis of gene expression. Different normalization approaches can have significant effects on the distribution and calculation of significant values (*P*-values). A previous study that compared different normalization approaches concluded that it was important to select a method that modifies the original data as little as possible and need to be selected in accordance with the characteristics of the data set to be examined 32,33. What analysis packages will I use?

This project contains some techniques that I am already familiar with but advances in technology have made it necessary for me to update my skills in order to compete effectively in the modern day job market. In order to this I will vastly expand my knowledge of new exciting techniques like RNA-seq while updating existing ones such as qPCR. Learning large scale data set production and analysis will be vital for this project due to the large volume of data that will be produced. The ability to handle and analyze large data sets is becoming an imperative skill in modern biochemistry.

The quality and reproducibility of results can be affected by variations in the workflow. It is very important to check the quality and quantity of RNA following extraction and samples should be handled carefully to prevent RNA degradation. Melting curve analysis will be included to assess the quality of the PCR reaction. Data normalization is a critical step in the qPCR workflow. This corrects for technical variance from the chemical reactions that take place during library preparation and PCR which contribute to inter-ample variance. I will normalize my data using 3-4 stably expressed reference genes.

Word count: 1781 (1500)

**Section 7: Ethical approval and licenses.**

Not Applicable.

**Section 8: Timetable**



**Section 9: References**

**Section 10: Future planning**

Following these pilot studies I would like to further characterize potential components that modulate gene expression in *C. neoformans* and the changes in phenotype/morphology observed. Understanding how bacteria and fungi interact within specific niches in the body and how these interactions affect pathogenesis is critical for identifying potential novel targets for drug development.

Modular cloning and CRISPR

I would like to examine the host-pathogen interaction comparing yeast and spores under previously optimized conditions and look at the cell surface modifications that occur. I would like to compare host epithelial and alveoli macrophages during the infection process and compare the resulting gene expression profiles to delineate what is up or down regulated and from this to infer what may be important at the point of infection. Further understanding of the interaction of opportunistic or primary fungal pathogens will be important in understanding fungal pathogenesis.

Knowing an organism is expressing certain genes provides a lot of information about how an organism is functioning and potentially new insights into which genes (and therefore proteins) are involved in certain behaviors

However, more information is needed than just the mRNA profile to establish function. It may be helpful to work out which proteins are synthesized through proteomics experiments

Few antifungal drugs are available for treatment. And devlopement of new therapies in complicated by the need for pathogen-specific targets. Euk targets not work in euks. Due to their eukaryotic nature and relatively close physiological similarity to human cells pathogenic fungi are notoriously difficult to target for clinical therapy.

Word count: (400)