University of Reading

Investigating evolutionary hypotheses of cancer cell motility

PhD in Biological Sciences

Biological Sciences

Anastasia Wass

September 2019

Declaration of original authorship: Declaration: I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Ana Wass

Acknowledgements

Thank you to all the people who helped make this possible.

Thanks to my supervisors Professor Phil Dash and Dr Louise Johnson for their continuous guidance and feedback throughout this project and to Dr Tiffany Taylor for coming up with the original concept and her help and supervision in initiating the project.

Many thanks also go to my friend Dr Tara Pirie, my mum, Jenny & dad, Martin, my granddad and grandma and last but by no means least my husband, Thomas, who have all supported me throughout this PhD.

Contents

Contents of figures

Contents of tables

Abbreviations

2 D – 2 dimensional

3 D – 3 Dimensional

APS - ammonium persulphate

CellLight - CellLight Nuclear BacMam 2.0 GFP

ChIP - chromatin immunoprecipitation

DMEM - Dulbecco's modified eagle's minimum essential medium

DMSO - dimethyl sulfoxide

ECL Reagent – enhanced chemiluminescence Reagent

ECM – extracellular matrix

FBS - foetal bovine serum

Fiji – image processing software (Fiji Is Just ImageJ)

FITC - fluorescin isothiocyanate

Green CellTracker – CellTracker Green CMFDA

h^2 – narrow-sense heritability

H^2 – broad-sense heritability

HAT - histone acetyltransferase

HDAC - histonedeactylase

MDA – MDA-MB-231

MEM - Eagle's minimum essential medium

MTT - MTT 3-(4, 5-Dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide

PBS – phosphate buffered saline

PCR – polymerase chain reaction

PPC – particles per cell

Qtracker – fluorescent Qdot nanocrystals

R^2 - is the amount of variance shared between the speeds of each group

Red Cell Tracker - CellTracker Red CMTPX

SDS - sodium dodecyl sulphate

siRNA - small interfering ribonucleic acid

TBS - tris buffer saline

TBST – tris buffer saline Tween-20

V_A - additive genetic variation

V_E – environment variation

V_G - genetic variation

V_P - phenotypic variation

Glossary

Adaptation – a process of genetic change of a population owing to natural selection, whereby a feature becomes prevalent in a population as it is a selective advantage. Due to the population becoming better suited to some feature in its environment.

Additive variance - the deviation from the mean phenotype due to the inheritance of particular alleles.

Adenocarcinoma - cancers that are epithelial in cell origin

Alleles - an alternate form of a gene; a single gene can have multiple versions, called alleles.

Aneuploidy - the presence of an abnormal number of chromosomes in a cell.

Broad-sense heritability (H^2) - the proportion of phenotypic variation due to genetic variation.

Cancer – is the name given to a group of diseases involving abnormal cell growth and proliferation.

Carcinogenesis – also known as tumorigenesis. The formation of cancer whereby 'normal' cells are transformed into cancer cells, characterised by changes at the cellular, genetic and epigenetic levels and abnormal cell division.

Carcinoma – a cancer originating from cells in the endoderm, mesoderm or ectoderm of a tissue lining.

Cell lines – a population of cells from a multicellular organism that are able to proliferate indefinitely.

Cell migration – the movement of a cell (or cells) from one location to another.

Cell motility – the ability of a cell to move independently using metabolic energy.

Chromothripsis - multiple chromosomal rearrangements localized to a limited number of genomic regions can be acquired in a single catastrophic event

CpG Islands - regions with a high frequency of C and G nucleotide sites. A region with at least 200 bp with a GC percentage greater than 50% and an observed-to-expected CpG ratio greater than 60%. [(No of C * No of G)/length of sequence].

Cytokinesis - part of the cell division process during which the cytoplasm of a single eukaryotic cell divides into two daughter cells.

Darwinian evolution (Darwinism) – a theory of biological evolution stating all species of organisms arise and develop through natural selection.

Differentiation - The acquisition of cell-specific differences during a multicellular organism's embryonic development or adult life; reflects gene expression and activation of transcription factors.

Dominance variance - caused by interactions between alternative alleles at a specific locus it is the extent to which a trait appears in a population or an individual.

Dispersal – the movement of individuals to different localities.

Dispersal theory – the theory that competition for resources within a population selects for individuals better adapted for dispersal.

Dysregulation – abnormality or impairment in the regulation of a metabolic, physiologic or psychological process.

Environments – anything external to the object of interest that may influence its activity or function.

Environmental variation – differences in the environment that can occur spatially and temporally.

Epigenetics – the study of heritable phenotype changes that do not involve alterations in DNA sequence.

Epigenetic mutations – functionally relevant alterations to the genome that don’t alter the nucleotide sequence of a gene.

Epigenome – a record of the chemical changes to DNA and histone proteins of an individual that can be passed to progeny via transgenerational inheritance.

Evolution – the origin of individuals possessing different states of one or more traits and changes in their proportion over time.

Evolvability – the capacity of a system for adaptive evolution.

Extracellular matrix (ECM) – collection of extracellular molecules such as collagen, enzymes and glycoproteins that give tissues their mechanical properties and provide structural and biochemical support of surrounding cells.

Experimental evolution – the study of evolutionary processes within a population in response to experimental conditions imposed.

Fibrosarcoma - tumours of mesenchymal cell origin from fibrous connective tissue

Fitness – the success in an entity in surviving or reproducing hence it is the average contribution of an allele or genotype to the population.

G1 phase - the first of four phases of the cell cycle that takes place in eukaryotic cell division. The cell synthesizes mRNA and proteins in preparation for subsequent steps leading to mitosis.

G2 phase - the third phase in the cell cycle directly preceding mitosis. A period of rapid cell growth and protein synthesis during which the cell prepares itself for mitosis.

Gene – the functional unit of heredity.

Genetic drift – random changes in the frequency of two or more alleles or genotypes within a population.

Genetic instability – a high frequency of mutations within the genome of a cell line.

Genetic mutations – permanent alteration of the nucleotide sequence of the genome.

Genetic variance – variation in a trait within populations, as measured by the variance due to genetic differences amongst individuals.

Genome – all the genetic material of an individual, includes coding and non-coding DNA as well as mitochondrial or chloroplast DNA.

Genotype – the set of genes possessed by an individual.

Heritability – statistic summarising the amount of variation in a phenotype due to genetic factors.

Heterogeneous – composed of different elements that are not of the same nature.

In vitro – meaning "in the glass", biological studies performed in the laboratory using equipment such as test tubes, flasks and petri dishes.

In vivo – studies performed on whole living organisms such as humans, animals or plants.

Inbred lines - are individuals of a species which are nearly identical to each other in genotype due to long inbreeding.

Immunoprecipitation (IP) - is the technique of precipitating a protein antigen out of solution using an antibody that specifically binds to that particular protein.

Kin selection – a form of selection where alleles differ in their rate of propagation through their influence on the reproductive success of individuals who carry the same alleles by common descent.

Metastasis - the successful migration of cells and the establishment of secondary tumours around the body.

Narrow-sense heritability (h^2) - the proportion of phenotypic variation due to additive genetic variation.

Natural selection – non-random differential survival and/or reproduction of individuals due to variation in their phenotypes.

Phenotype – the composite observable characteristics of an individual including morphological, physiological, biochemical and behavioural properties.

Phenotypic plasticity – the capacity of a genotype to alter or change its phenotype depending on the environment.

Phenotypic variance - measures the degree of phenotypic differences among a group of individuals; composed of genetic, environmental, and genetic-environmental interaction variances.

Population – a group of individuals from the same species occupying a geographic region and exhibiting reproductive continuity from generation to generation.

Proliferation – the reproduction and/or multiplication of similar forms.

Protein – large biomolecules or macromolecules consisting of one or more chains of amino acid residues.

Pseudodiploid - diploid cells with chromosomal translocations

Recombinant inbred lines - an organism with chromosomes that incorporate an essentially permanent set of recombination events between chromosomes inherited from two or more inbred strains.

Resource – a supply from which a benefit is produced.

Response to selection – the change in the mean value of a trait over one or more generations in response to selection.

Robustness - persistence of a characteristic or trait even under environmental perturbations.

Selection – differential survival and/or reproduction or preferential elimination of classes of phenotypically different entities by means of natural or artificial controlling factors.

Selective pressures – a phenomena that alters the behaviour and/or fitness of individuals within a given environment shaping their evolution.

Somatic cell – any diploid differentiated biological cell forming the body of an organism that does not give rise to gamete cells.

Statistical Power - is the measure of a test's ability to accurately detect that the null hypothesis is false. Specifically, power is the probability that a test with the specified assumptions (sample size, difference, standard deviation, alpha level, and type of alternative hypothesis) correctly rejects the null hypothesis when the alternate hypothesis is true.

Trait – a morphological, physiological or behavioural feature.

Tissue – in biology a tissue is a cellular organisational level consisting of similar cell types and their extracellular matrix that together carry out a specific function.

Tumour – abnormal or excessive growth of tissues comprised of heterogeneous cell types predominantly cancer cells.

Tumour microenvironment – the environment around a tumour, including blood vessels, immune cells, fibroblasts signalling molecules and the extracellular matrix.

Variance – the average squared deviation of an observation from the arithmetic mean.

Abstract

Cancer is a disease of evolution. Mutations within a cell lead to the acquisition of cancerous phenotypes. Tumour evolution depends on heritable differences between cells. The extent of heritable variation has not been measured for any trait in cancer cell populations. In this thesis techniques have been developed to estimate the broad-sense heritability (H^2) of cancer cell traits *in vitro* and used to estimate the H^2 of cell motility. Cell motility is a trait related to the cancer hallmark of metastasis. Results show that motility is strongly heritable with H^2 values ranging from 0.77-0.36 across multiple cell generations. H^2 estimates appeared to decrease slightly between more distantly related cells, a trend that could occur due to a decrease in the genetic contribution to motility or an increase in environmental variation. This was tested by treating cells with epigenetic inhibitors and obtaining H^2 estimates. Results showed H^2 estimates were not significantly affected by the application of epigenetic inhibitors with values ranging from 0.95-0.18. Quantification of the amount of environmental variation in *in vitro* cell culture media was attempted using image analysis of fluorescent particles. Variation in particle distribution was found at a range of concentrations, nM – mM. Direct quantitative measures of evolvability in cell traits could have valuable applications to cancer research and tumour treatment.

To understand tumour progression, evolutionary theory can be applied to cancer cells *in vitro* to elucidate the selective pressures driving the evolution of cancer cell traits. In this project experimental evolution techniques have been adapted from microbiology and applied to cancer cell lines *in vitro*. Adaptation of cell lines to low nutrient environments over 12 weeks showed dispersal theory may play a role in the selection of the cancer cell trait motility. Understanding the selective pressures driving the acquisition of cancer phenotypes will have valuable applications clinically in understanding tumour progression.

Chapter 1 - Introduction

1.1 Cancer

Cancer is not a single disease but a collection of distinct subtypes all with different clinical outcomes[1]. Globally cancer affects approximately 17 million people every year with an estimated 9.6 million deaths in 2018 due to the disease[2].

Excessive cell growth leads to tumour formation. Solid tumours can be benign or malignant. Benign tumours are generally slow growing, non-invasive, morphologically similar to their parent tissues and usually enclosed within a fibrous capsule[3]. Prompt diagnosis and treatment typically results in a complete cure. Malignant tumours are normally fast growing, invasive and morphologically abnormal and are not usually encapsulated[3]. Malignant tumours can and do spread around the body[4].

Cancer is the evolution of somatic cells; over time the 'normal' cell population within a tissue may be replaced by cells carrying mutations conferring some cancer phenotypes[3 5]. Carcinogenesis is initiated by genetic and epigenetic mutations altering phenotypes leading to the acquisition of the 'cancer hallmarks'[1 6 7]. Dynamic genetic changes along with clonal expansion and selection are thought to lead to a heterogeneous tumour[6 8 9]. This multi-step process is an example of Darwinian evolution with individual cells being the unit of selection[6]. Tumours contain heterogeneous cell populations that vary in the traits under selection[8 9]. Understanding the cellular heterogeneity within cancer is key to developing prevention strategies and targeted therapeutic interventions[7]. In order to understand how these distinct cell populations evolve, knowledge is needed of their reaction and response to the environment and the role phenotypic plasticity, genetic and epigenetic mutations play in that.

1.2 Evolution and cancer

Cancer is a disease of evolution; tumours evolve through selection acting on mutations within individual cancerous cells and this process is integral to disease progression[10]. Most multicellular organisms have potent tumour suppressive mechanisms, initially necessary for the evolution of multicellularity they also play an important role in allowing large size and longevity[10 11].

Multicellular organisms would not have been able to evolve significant organisation or complexity without mechanisms to prevent competition between individual cells[10 11]. If competition is not prevented, then genetic conflict arises when mutant cell lineages promote themselves at the expense of the organism[11].

Cancer cells represent a breakdown in multicellular cooperation and dysregulation in the control structures imposed during the evolution of multicellularity[12 13]. The genes mutated in cancer are those often used by unicellular organism that have been adapted for multicellularity. Adaptations to multicellularity can occur in multiple ways and the traits used by unicellular organisms are often integral to survival, as such these traits often show great robustness and phenotypic plasticity[11-13]. The high degree of robustness and plasticity found in these traits confer increased adaptability to tumour cells by enabling an effective response to environmental pressures[12]. Cancer is a disease of evolution, applying evolutionary theories to cancer will elucidate the selective pressures driving the evolution of cancer cell traits and aid understanding in tumour progression.

A tumour is a microcosm of genetically heterogeneous cells where each clonal lineage is competing (or occasionally cooperating) [10]. The genetic variability is caused by high mutation rates, multiple rounds of selection, variations in the microenvironment and an increased proliferative rate[10]. Selective pressures such as chemotherapeutic agents or hypoxia will select for cells best adapted to those conditions. If these selective pressures are very strong an evolutionary bottleneck can occur resulting in the selected cells forming the majority of the tumour population[10]. Understanding the evolutionary trade-offs in tumour progression and harnessing competition between cell lineages can be used for disease treatment. Targeted therapies recognise the adaptive nature of cancer populations but in order for them to be effective, understanding is needed of the evolutionary trajectories of cancer in different contexts[10].

Within evolutionary biology understanding of how environmental variation can affect phenotypic plasticity (the capacity of a genotype to alter or change its phenotype depending on the environment[14]) in genotypes is crucial to understanding the evolution and maintenance of biodiversity[15]. These same theories can be applied to cancer biology where the tumour microenvironment may cause or select for increased genetic instability and tumour progression has been linked to genetic heterogeneity[16 17]. Hanahan and Weinberg (2000) predicted that the future of cancer biology would develop into a logical science where the

complexities of the diseases would become understandable in terms of the underlying principles[1]. Using evolutionary biology principles could help to do this. This thesis aims to apply novel evolutionary concepts to cancer biology and in particular to develop and adapt experimental evolution techniques to cancer biology.

1.3 Metastasis

Metastasis is the spread of cancer cells around the body and the growth of secondary tumours[4 18] as such it is the main clinical challenge in treating cancer. Clinically, the evolution of metastasis is one of the most important transitions in tumour progression. Prior to metastasis many solid tumours can be cured surgically and 5-year survival rates are often above 90 %. However, once a cancer has spread to distant sites systemic therapy is necessary, and 5-year survival rates often fall below 15 %[19]. It is currently impossible to predict or detect whether metastasis has occurred until there is a relatively large secondary tumour, and available therapies only target detectable tumours to prevent or reduce their growth or prevent dissemination[20].

As seen in figure 1.1 metastatic tumours are the final stage of a multi-step process termed the invasion-metastasis cascade[21]. The first step is when cancer cells in the primary tumour invade their local surroundings into and through the extracellular matrix. Next is intravasation, where cells move into the blood or lymphatic vessels. These cells then have to survive in circulation before coming to a rest at a distant site and extravasating – moving out of the vessel, into the distant tissue. Cancer cells that have achieved all this still need to adapt and survive in the foreign microenvironment before proliferating and forming metastatic tumours[21]. These events are driven by genetic and/or epigenetic changes within the tumour cell as well as alterations in the tumour microenvironment[21]. Whilst metastasis is composed of multiple cell traits, cell migration is integral to cancer metastasis. Cells must migrate out of the primary tumour and both into and out of the circulatory system[21].

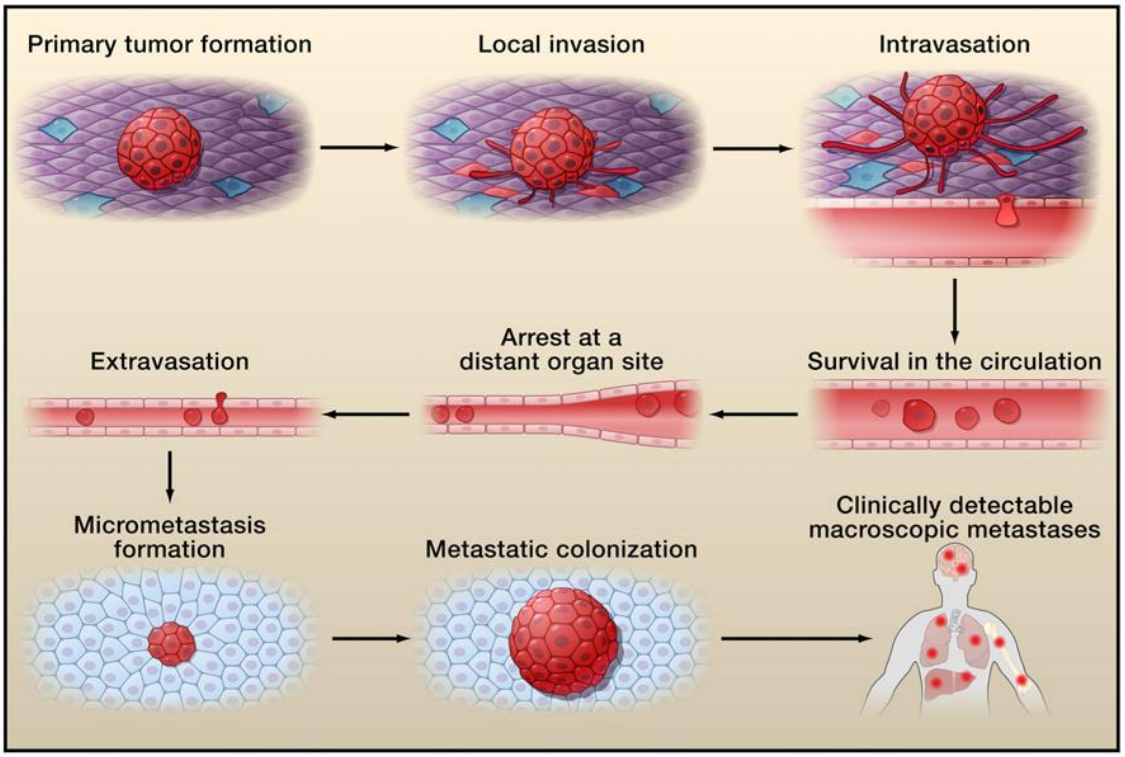


Figure 1.1: The invasion-metastasis cascade; taken from Valstyan and Weinberg (2011)[21]. Cancer cells leave the primary tumour and intravasate into the circulatory system. They travel to a distant site before extravasating into the distant tissue. These cancer cells form micro metastases whilst surviving and adapting to the foreign microenvironment. Once adapted they continue proliferating until a secondary metastatic tumour is formed.

1.3.1 Cell migration

Cell migration is the movement of cells and involves an array of different mechanisms[22] controlled by the cytoskeleton and driven by actin polymerisation[23]. Cell migration can occur naturally during embryonic development, physiological homeostasis and wound repair[24]. Cells can move either individually or as a group[22]. Individual migration can be further classified into amoeboid migration - the cell moves via rapidly alternating cycles of morphological expansion and contraction; and mesenchymal migration – involving proteolytic remodelling of the matrix[22 24]. When undergoing collective cell migration the cell-cell junctions between migrating cells are maintained and they move as connected strands or chords into tissues[22].

Cancer cells have been shown to utilise multiple different types of migration and in solid tumours it is thought to be the underlying cause of cancer metastasis[24]. The role of cell motility in the progression of metastasis has been established experimentally and empirically[24]

making it an important therapeutic target for cancer treatment. However, after years of *in vitro* study there are very few clinical interventions designed to specifically target cell migration[23 24].

1.3.2 The evolution of cancer metastasis

Metastasis is a recurring trait in cancer progression[25], for many types of cancer (such as breast or skin) it is metastasis that causes these tumours to become deadly[2]. However there is a paradox in our understanding of the evolution of metastasis. Genetic and epigenetic mutations generate new clones of cells and those with a survival or proliferative advantage should expand and spread within a tumour. Unlike the other cancer hallmarks (apoptosis resistance, autonomy in growth signals, replicative immortality and sustained angiogenesis[1]) metastasis does not appear to directly enhance cell survival or proliferation[6 25]. If a cell acquired a metastatic phenotype it is likely to be at a disadvantage within the tumour as some of its reproductive fitness is lost to emigration. Therefore it might be expected that cells without a metastatic phenotype would have a fitness advantage and the metastatic phenotype would be driven extinct[6]. The cost of dispersal to individual cancer cells has been demonstrated experimentally, a high number of cancer cells can be detected in patients' blood (a daily average of 10^6-10^7 cells) however few secondary tumours form at distal sites[26-28]. The majority of cells that disperse must either die or return to the primary tumour[26], indicating that genetic alterations promoting metastasis do not confer any advantage within the primary tumour as the majority of motile cells either 'waste' energy that could be spent on growth and proliferation or die[6].

Tumours are heterogeneous, both genotypically and phenotypically[9]. The multistep nature of metastasis coupled with the diversity and complexity of cancer cells poses difficulties in experimental design and interpretation. A biopsy taken from a patient or a cell line studied in the lab, may not be representative of the tumour as a whole due to natural selection acting over time and space[9 29]. Tissue studies of patients are complicated by the cellular heterogeneity of a tissue mass and also by the variance in genetic background between individuals[7 20].

Theoretical models and ecological observations seem to agree that populations that are very phenotypically and genetically diverse tend to occupy broader ecological niches[30-32]. Increased diversity seems to reduce intraspecific competition and extinction risk, increase population growth, establishment success and invasiveness and decrease vulnerability to environmental

changes and changes in population size[14]. These traits nearly all apply to metastatic cancer and could suggest that the phenotypic and genetic diversity seen in tumours contributes to the severity of the disease.

There is no single 'metastatic gene', many studies have indicated there are lots of genes or signalling pathways and molecules that can be co-opted or altered to confer a metastatic phenotype[33-35]. This could suggest something in the environment triggers these genetic and epigenetic mutations or there is a selective pressure maintaining the metastatic phenotype in the primary tumour. In order to understand how these distinct cell populations evolve, knowledge is needed of their reaction and response to the environment and the role phenotypic plasticity and genetic or epigenetic mutations play in that.

1.3.3 Dispersal theory and metastasis

In ecology, dispersal theory suggests that competition for resources within a population would select for individuals better adapted for dispersal[29 36 37]. Cell migration is a key part of the metastasis cascade[1], if we apply ecological dispersal theory to cancer this could help to explain why tumours metastasise. Dispersal theory could explain how a metastatic phenotype is maintained within a primary tumour population[6 38].

Despite the energy demands of motility and the uncertain advantage to cell fitness, dispersal allows migration away from areas of resource restriction[29 38]. Cells able to disperse increase their own chances of survival as they may migrate to an area with less competition or higher resource concentration[39]. Kin selection – where selection on individuals favours traits that increase the fitness of close relatives, will also act to increase dispersal when competition is high as dispersing cells reduce competition at the original site increasing survival of their relatives[39]. As cancer cells are clonal a dispersing cells genome should be preserved within the original population, even if the dispersing cell is unsuccessful the increased chance of survival for the remaining clonal population increases and the genome is maintained and propagated throughout the population[39].

Dispersal theory has been tested experimentally by Taylor & Buckling (2010) who used bacterial populations of *Pseudomonas aeruginosa*. They found that bacteria which dispersed further reduced cell-cell competition increasing population fitness even when the cost of dispersal was high[40]. Mathematical models also predict that even in environments with very low resource heterogeneity, adaptations for dispersal will still be selected for as migration

away from resource competition increases offspring survival and genotype frequency is increased if this competition is not against kin[14 41].

Primary tumours have been shown to be highly heterogeneous in their microenvironment[8]. In ecology, resource heterogeneity has been shown to initiate dispersal in the stream salamander *Gyrinophilus porphyriticus,* with low habitat quality selecting for dispersal[42]. Resource limitation has also been closely linked to dispersal of female African buffalos[43]. The resource heterogeneity within a tumour could be driving selection of the motile phenotype promoting dispersal away from resource limitation[19 36 38 44]. Explaining the evolution of metastasis requires cells to vary heritably in their dispersal behaviour[45]. However, heritability - a statistic that summarises the amount of variation in a phenotype due to genetic factors[46 47] has never been directly measured for any trait in cancer cell populations. Before testing ecological dispersal theory, it would be helpful to obtain heritability estimates of the phenotypic variation in the cell trait motility.

1.4 Experimental evolution

Experimental evolution is the study of evolutionary processes within a population in response to experimental conditions imposed[48]. Experimental evolution uses replicate populations of organisms (microorganisms, such as bacteria, are ideal for such experiments as they have large populations and short generation times) [37] to study evolutionary processes in real time[36 37 48]. The environmental conditions under which evolution occurs are carefully controlled allowing monitoring of the effect of specific selective pressures on the trait of interest[37]. This research approach is used to study adaptation, estimate evolutionary parameters and test evolutionary hypotheses[37 48].

Evolutionary theories are usually inspired through studying outcomes of past evolution (such as phylogeny, divergence of species, genome structure and sequence) [48]. Gerlinger et al (2012) performed genomic analysis on multiple tumour regions from primary and metastatic renal tumours. They found substantial intratumour heterogeneity and using phylogenetic reconstruction hypothesised that it was this that allowed tumour adaptation leading to therapeutic failure through Darwinian selection[9]. Signatures of natural selection on gene sequences have been detected in tumours and evolution can be inferred using techniques such as phylogenetic reconstruction of the cancer genome, however these methods can only hint at how selection shaped adaptation[9 49 50]. Few of the evolutionary and ecological theories behind tumour evolution have actually been tested experimentally.

In contrast, experimental evolution techniques offer the prospect of actually observing evolution in real time[37 48]. Adapting experimental evolution techniques to cancer cells in culture would allow direct observation of which phenotypic traits are under selection. These techniques are widely used and have fundamentally changed our understanding of evolutionary processes[30 50-58]. For example, in microbiology *Pseudomonas fluorescens* have been used to study the underlying causes of adaptive radiation and its role in biological diversity[30]. The longest running, and possibly best known, evolution experiment has been conducted by Lenski et al[51] and involves 12 populations of *Escherichia coli* derived from the same ancestral strain that have been continuously grown in glucose limited media since 1988 (now over 60,000 generations)[52]. This experiment has given rise to many new insights and testing of evolutionary theories such as the trade-offs between growth rates and yields[53] and the relationship between the rate of genomic evolution and adaptation[54]. Such techniques have a wide range of uses, not only in bacteria. Barrett et al (2011) used experimental evolution techniques to investigate the effects climate change could have on populations of the stickleback fish, *Gasterosteus aculeatus*, by artificially exposing them to alterations in their environment and monitoring population adaptation over time[55].

Use of experimental evolution techniques and their application to cancer biology will allow testing of evolutionary theories to further understanding of the clinical progression of tumours[36 38]. The application of such techniques will allow the ecological causes of natural selection to be investigated and the genetic basis of evolutionary changes to be determined[36].

In vitro cell cultures are well suited for use in experimental evolution, sharing many similarities with bacteria. Cancer cell cultures are easy to grow, and many cell lines are well studied, they have relatively fast generation times (≈1 day) allowing experiments to run for multiple generations. They can be stored easily and indefinitely, and large populations of cells take up a small volume allowing multiple cell lines and experimental conditions to be cultured simultaneously as well as storage and resurrection of ancestor populations for comparison[38 59]. However, cancer cells in culture are also fastidious in their growth conditions and individual cells do not grow well in isolation. They are also visually indistinguishable from each other, making experiments requiring co-culturing of populations' difficult[60]. Cancer cells are also more complex than the organisms typically used in experimental evolution, they evolved as part of a multicellular eukaryotic organism before undergoing their own individual mutations and evolution to form the cancerous cell and tumour[12]. As a result their genomes are far more complex than unicellular organisms as they include multiple levels of gene regulatory

systems[12 61]. There are also intricate levels of heterogeneity which exist within tumour populations[62 63]. Using adapted experimental evolution techniques allows evolution to be observed in real time as populations of cells adapt to new environmental conditions by natural selection.

1.5 Using cell culture as an experimental evolution technique

Avoidance of cell death is one of the main cancer hallmarks[1] and means that cancer cells are well suited for *in vitro* cell culture as they continue to divide indefinitely[64]. Immortal cell lines are a very important tool for research into the biochemistry and cell biology of multicellular organisms. Cells can be grown indefinitely in culture simplifying analysis of the biology of cells which may otherwise have a limited lifetime[65].

1.5.1 Advantages and limitations to cell culture

There are many advantages to starting with *in vitro* studies using cancer cell lines. Cell lines are derived from human tumours, so carry some clinical relevance. Most cancer cell lines have been well studied and characterised. Studying multiple different cell lines has similar advantages to studying multiple individuals, but in a simpler system with fewer ethical concerns[7]. Many specific pathways and functions remain in cells in culture. A major advantage of *in vitro* cell culture is the ability to precisely control the environment. Variations between repeats can be kept to an absolute minimum and experimental parameters repeated or altered with relative ease.

No matter how carefully controlled and regulated an *in vitro* environment is, it will always be simpler and lack some elements of an *in vivo* environment. This may lead to a more constant cell metabolism *in vitro*. Whilst useful experimentally this may not be truly representative of the *in vivo* tissue or cell behaviour[64].

There are limitations to *in vitro* cell culture. Cells from multicellular organisms do not normally exist in isolation. Those cells that survive the culture process may not represent the multicellular organism they come from[62]. Differences between cells *in vitro* and *in vivo* mostly arise from the dissociation of cells from a three-dimensional structure and their culture on a two-dimensional substrate. When a cell line forms it may represent only one or two cell types from the original tissue hence many cell-cell interactions may be lost[62 63]. Care must be taken

when extrapolating results from *in vitro* experimentation and applying them to an *in vivo* environment[62].

Whilst there are many differences between *in vivo* and *in vitro,* and some limitations to *in vitro* cell culture, there are also many advantages. Provided the limitations of cell culture are acknowledged and understood, it remains an important tool for the study of cell biology from multicellular organisms. Cell culture thus provides an *in vitro* model of a cell or tissue type in a well-defined environment which can be easily manipulated and analysed[64].

1.6 Choosing cell lines for experimental work

Using just one cell line, or even one type of cancer, to test new hypotheses on cancer progression and treatments, limits understanding of how disease progression occurs and risks missing potential treatments for subsets of cancer types[63]. Using multiple cell lines and different types of cancer makes results more robust in their applicability to cancer as a disease[66 67]. Particularly so in this study where theories of evolution are being applied to explain disease progression.

For this study four cancer cell lines were chosen. These cell lines were chosen to represent not only genetically distinct cancers but also different cancer types. Including different cell types and different cancer types will give a broad range of comparison, making results more robust[64].

In 2015, breast cancer was the most common cancer in the UK, accounting for approximately 15 % of all new cancer cases[2]. Two of the four cell lines chosen originated as breast cancer, MDA-MB-231 and MCF7. They were both established from pleural effusions of metastatic adenocarcinomas (cancers that are epithelial in cell origin) [68 69] but are genetically quite different.

MCF7 is the most commonly used breast cancer cell line in the world[70 71]. MCF7 cells are particularly useful for *in vitro* studies as they have retained characteristics of clearly defined mammary epithelia. The MCF7 cell line is very genetically unstable with differences in genotype between batches[72]. This genetic instability may cause MCF7 cells to react unpredictably and the differences in gene expression may result in diverse responses between the cell lines[70]. MDA-MB-231 cells are epithelial breast cancer cells and are a commonly used breast cancer cell line in medical research laboratories[68]. By using two types of breast cancer

cell lines comparisons can be made in how diverse cancer characteristics react to the experimental conditions.

The third cell line chosen was HeLa cells, one of the most commonly used cell lines for studying human cellular and molecular biology[73 74] they have been extensively characterised[73]. Similar to both MDA-MB-231 and MCF7 cells, they are also adenocarcinoma, but originate from cervical cancer[74]. As HeLa cells are so widely used, any results will have meaningful comparisons to relevant cancer literature. When the HeLa cell genome was compared to a human reference genome there were both single nucleotide and structural variants; a high number of these were in genes effecting proliferation, transcription and DNA repair[73 75]. Like MCF7 cells their genetic instability[75 76] may make them react not only unpredictably, but radically different to the other cell lines chosen.

Lastly the HT1080 cell line is different to the others chosen as they originate from a fibrosarcoma[77]. Fibrosarcomas are tumours of mesenchymal cell origin from fibrous connective tissue. They do show genetic abnormalities, but these are relatively stable and well characterised to the cell line[77].

1.7 Project aims

The aim of this project is to test whether an evolutionary approach to natural selection in cancer is appropriate and to adapt experimental evolution techniques and use these on cancer cell lines *in vitro* to study cancer metastasis.

For experimental evolution techniques to be successful natural selection must be able to act, natural selection relies on heritable variation in a cell trait. As such the first step was to measure the heritability of the cell trait motility. Cell motility is a readily quantifiable trait which has been used as one representation of a cell's metastatic potential. Whilst heritably of cell traits has often been assumed, testing of this assumption is a fundamental pre-requisite in order to quantify the evolvability of motility in cancer cell lines. When measuring heritability of cell motility, the hypothesis was that there would be heritability of cancer cell traits and that these experiments would provide proof of concept.

The second step involved adapting the experimental evolution techniques commonly used on microbes to cancer cells in culture. Whilst cancer cell lines and bacteria share many similarities modifying methods to make them suitable for cancer cell cultures was no trivial process.

Once adapted, these techniques could be used to test whether dispersal theory applies to cancer. Selective pressures could be applied to cancer cells in culture and their motility measured over time. As metastasis is a composite trait a range of methods were used to measure multiple cell phenotypes. If dispersal theory does apply to cancer cells then it would be expected that competition for resources promotes increases cell motility.

Chapter 2 – General methods

2.1 Cell lines

Cell culture is the process by which cells of interest (in this case cancer cells) are isolated from living tissue and grown under carefully controlled conditions outside of their natural environment[64 67].

Multiple cell lines have been used throughout this project and are detailed below. These cell lines represent not only genetically distinct cancers but also a broad range of cancer types. Cell lines were chosen that allowed comparisons to current results in the cancer literature, all the cell lines selected have been well documented and are widely used in *in vitro* studies. Where the same cell lines have been acquired from different sources they are numbered sequentially in the order in which they were obtained.

2.1.1 HeLa cells

HeLa (ATCC® CCL-2™) cells are cervix epithelial cells taken from a 31-year-old black female in 1951[74]. Similar to MCF7 and MDA-MB-231 cells, they are an adenocarcinoma. The HeLa cell line has been reported to contain HPV-18, have a low expression of p53 and to be highly motile and fast proliferators[73].

Characterisation of HeLa cells chromosomes have shown high levels of aneuploidy and many large structural variants[73]. They have been shown to have substantial chromosomal aberrations, including evidence of chromothripsis (where a limited number of genomic regions have undergone multiple chromosome rearrangements) [76].

- HeLa cells were obtained from Public Health England - Lot 14G005 acquired at passage 3.

2.1.2 HT1080 cells

The HT1080 (ATCC® CCL-121™) cell line was isolated in July 1972 from a fibrosarcoma of the connective tissue in a 35-year-old caucasian male[78]. They are also highly motile with a high rate of proliferation[77]. HT1080 cells are different to the other cell lines as they originate from a fibrosarcoma, not a carcinoma[77]. They show genetic abnormalities, but these are relatively

stable and well characterised to the cell line. HT1080 cells have a modal chromosome number of 46 and show signs of pseudodiploidy (diploid cells with chromosomal translocations)[77].

- HT1080 cells were obtained from Public Health England - Lot 14A031 acquired at passage 3.

2.1.3 MCF7 cells

MCF7 (ATCC® HTB-22™) is a breast cancer cell line. It was established at the Michigan Cancer Foundation and is derived from the epithelial mammary gland cells of a metastatic pleural effusion obtained from a 69-year-old caucasian woman. They are generally non-motile, oestrogen receptor expressing cells[69].

The MCF7 cell line is classified as Luminal A based on its expression of both oestrogen and progesterone receptors and its absence of human epidermal growth factor 2. Expression of Ki67 (a proliferation marker) is low[70] and the cell line has been shown to be very genetically unstable with differences in genotype between batches[72]. This project used three separately sourced MCF7 cell lines.

- MCF7-1 cells were obtained from ATCC - HTB-22, acquired at passage 17.
- MCF7-2 cells were obtained from Public Health England - Lot 13K023 acquired at passage 15.
- MCF7-3 cells were obtained from Public Health England - Lot 14I018 acquired at passage 13.

2.1.4 MDA-MB-231 cells

MDA-MB-231 (ATCC® HTB-26 ™) cells are epithelial mammary gland cells. They were derived in 1973 from a pleural effusion of a metastatic mammary adenocarcinoma from a 51-year-old caucasian female[68]. They are a highly motile breast cancer cell line[68].

The cell line is characterised as a claudin-low breast carcinoma. It is triple-negative (lacks oestrogen and progesterone receptors and does not express human epidermal growth factor 2), has down-regulation of claudin-3 and claudin-4, low expression of Ki67 and increased expression for markers associated with the epithelial-mesenchymal transition[70 79]. The cell line is aneuploid with near triploid chromosome numbers, it has a modal chromosome number of 64[68]. This project used two separately sourced MDA-MB-231 cell lines.

- MDA-1 cells were obtained from Dr. P Darbre's Lab acquired at passage 4.
- MDA-2 cells were obtained from Public Health England - Lot 13J001 acquired at passage 40.

2.2 Cell culture

All cells were grown as a monolayer at 37 °C in 5 % atmospheric CO_2. The standard culture conditions are listed below for each cell type. The culture medium provides the necessary nutrients, growth factors and hormones for cell growth as well as helping regulate pH and osmotic pressure. Basal media contain amino acids, vitamins, inorganic salts and a carbon source, but these media must be supplemented with serum[64 80 81].

Any deviations are stated in the specific methods section of each chapter. Unless otherwise stated all cell media reagents were obtained from ThermoFisher Scientific Inc (UK). All cell culture work was done in a sterilised tissue culture hood with sterile equipment. Cell cultures were checked every 2 or 3 days using a light microscope. Cells were either passaged or had the media changed as required.

Cell lines have been grown in media following the recommended guidelines from Public Health England or, where cell lines have obtained from an alternative source, culture has been continued in their previous media.

Both Eagle's minimum essential media (MEM) and Dulbecco's modified MEM (DMEM) are commonly used culture media that are suitable for a wide range of mammalian cells[64]. DMEM has a greater concentration of vitamins, amino acids and bicarbonate buffer than MEM[81].

Phenol red is often added to cell media as a colour indicator of pH. It is red at pH 7.4, changing to yellow if the pH drops, or purple if pH increases[64]. Phenol red interferes with spectrophotometric and fluorescent assays[82]. It is also a weak oestrogen mimic so can enhance the growth of cells with oestrogen receptors, such as MCF7 cells[83]. For these reasons phenol red free media has been used throughout this project.

Cells were grown in culture flasks or dishes as specified in the specific methods section of each chapter. Table 2.1 shows the various sizes and cell numbers of culture vessels used throughout this project.

Table 2.1: Information required for cell culture in various culture vessels. Adapted from Useful Numbers for Cell Culture by ThermoFisherScientific.

Name	Surface area (cm^2)	Recommended seeding density	Average cell number at confluency	Volume of media (mL)	Volume of Trypsin (mL)
35 mm Dish	8.8	0.3×10^6	1.2×10^6	2	0.5
6-well plate	9.6	0.3×10^6	1.2×10^6	3	1
12-well plate	3.5	0.1×10^6	0.5×10^6	1	0.5
24-well plate	1.9	0.05×10^6	0.24×10^6	1	0.2
T-25 Flask	25	0.7×10^6	2.8×10^6	5	2
T-75 Flask	75	2.1×10^6	8.4×10^6	10	4
T-175 Flask	175	4.9×10^6	23.3×10^6	35	15
T-225 Flask	225	6.3×10^6	30×10^6	45	20

2.2.1 MCF7-1, MCF7-2 and MDA-1 culture media

Cells were grown in low glucose (1 g/L) DMEM containing 4 mM glutamax, 1 mM sodium pyruvate and 10% foetal bovine serum (FBS).

2.2.2 MDA-2 culture media

MEM containing 10 % FBS and 2 mM glutamine.

2.2.3 MCF7-3, HeLa and HT1080 culture media

MEM containing 10 % FBS, 2 mM glutamine and 1 % non-essential amino acids.

2.3 Passaging cells

Passaging (also known as subculturing or splitting cells) involves removing a small number of cells from the culture and transferring them into a new flask. If passaged regularly most cancer cells can be cultured indefinitely, as this avoids senescence associated with high cell density[84]. As cells are randomly mixed at each passage, replicate samples of each culture are considered similar to each other and the characteristics of the cell line may be perpetuated indefinitely[64].

The various sizes of culture flasks and their average confluent cell number (the number of cells at which the entire surface area is covered) can be seen in table 2.1. When cells reached approximately 80-90 % confluency in their culture they were passaged, this varied depending on cell type and culture conditions but occurred approximately once a week. The average

generation time for a cancer cell *in vitro* is 24 hours meaning on average populations were passaged every 7 days. All cells were removed from the flask and a subset of the population transferred into a new culture. All media was pipetted off and the cells washed in phosphate buffered saline (PBS).

To detach cells from the culture surface enough Trypsin-EDTA (an enzyme that degrades the cell attachment proteins) to lightly cover the cell culture was added and the culture placed into the incubator for 4 minutes. After this time cells were checked under an inverted light microscope to confirm detachment from the surface.

1 mL more than the volume of Trypsin-EDTA of the appropriate medium was then added and the cell solution pipetted thoroughly to ensure mixing, if needed the cell concentration was then calculated using a haemocytometer.

The cell solution was pipetted into new cultures. If passaging to maintain culture, then enough solution was taken for approximately 10-20 % confluency in the new flask. If a specific cell concentration was required, then the volume of cell solution required was calculated from the haemocytometer count.

Fresh medium was then added to the new cell cultures which were clearly labelled before being placed back into the incubator.

Even with short term cultures, heterogeneity in growth rate and differentiation within a population can produce variability between passages. Replication of experiments and comparison of the results is essential for reliable analysis and conclusions. Keeping passage numbers low and within a short range between experiments will reduce genetic drift within the population as this reduces the number of generations and the time frame over which mutations have to accumulate within a population. Maintaining large population sizes reduces bottlenecking and means sample populations are more representative of the cell line.

2.4 Cell counting

The ability to count cells in culture is essential for standardisation of culture conditions, replication between experiments and measuring the effect of experimental parameters. The rate of cell proliferation is often used to determine the response of cells to experimental parameters, such as nutrient levels or drug concentration. Quantifying culture growth is also

an important aspect of routine cell culture to ensure consistency in technique; a change in growth rate may be an indication of contamination or equipment failure. There are many ways to measure cell proliferation, the simplest being to place a culture on a microscope and count the cells[41]. A direct cell count can be obtained using a haemocytometer or an indirect count can be obtained through methods such as tetrazolium reduction assays[64 85].

2.4.1 Haemocytometer

To determine the cell concentration using a haemocytometer, a known volume of cell solution is placed within a defined area of an optically flat chamber. The number of cells are counted using a microscope and the overall cell concentration is calculated from this sample.

The advantages of using a haemocytometer are that it is very easy and cheap to count large samples and conduct multiple repeats. The disadvantages are that it is labour intensive and requires a considerable amount of time, it is also affected by human error as the cell number is obtained by manual counting[64]. The haemocytometer protocol was adapted from protocol 20.1 of Freshney's book[64].

During cell passaging, before seeding the cell solution into culture, 20 µl of cell solution was mixed 1:1 with 0.4 % Trypan blue in an Eppendorf tube. Trypan blue stain is not taken up by live cells, counting only clear cells allows a viable cell count to be obtained. 10 µl of this mixture was added to one side of the haemocytometer. The haemocytometer was then placed under an inverted light microscope on a 10x objective.

A minimum of two large corner squares (one corner square is highlighted in blue in figure 2.1) and 100 cells were always counted. Cells lying on the top or left-hand side of the square boundaries were counted but not those lying on the right or bottom boundary lines. As the volume of one square is known (1 mm^3) the following formula can be used to calculate cell concentration;

$$Total\ Cells/mL = Total\ cells\ counted \times \frac{Dilution\ factor}{Number\ of\ Squares\ counted} \times 10{,}000$$

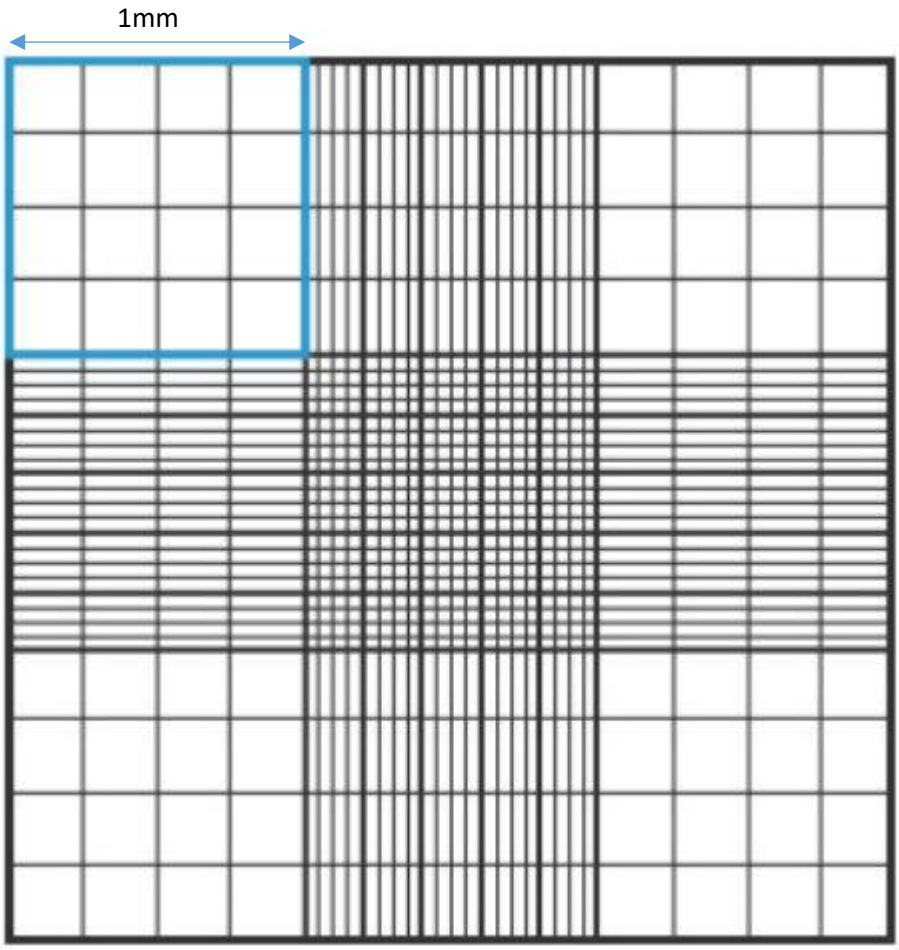


Figure 2.1: Diagram of a haemocytometer grid. Made by Ana Wass using PowerPoint with reference to a Bright-Line Haemocytometer as viewed using an inverted light microscope.

2.4.2 MTT cell viability assay

Viability stains can be used to give a more precise measurement of the live cell concentration. One of the most frequently used methods for measuring cell viability is reduction of MTT 3-(4, 5-Dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide (MTT) [86], a monotetrazolium salt. Tetrazolium salts are a large group of heterocyclic organic compounds that when reduced form formazans (insoluble coloured compounds) [86].

MTT assays provide an indirect measure of the number of viable cells in a culture. They involve the conversion of the water soluble MTT to an insoluble (purple) formazan. Only metabolically active cells convert MTT into formazan. Cells quickly lose the ability to perform this function after death. Formazan accumulates as an insoluble precipitate inside viable cells. This is then solubilized, and its concentration determined by optical density. The optical density being proportional to the number of viable cells present[85].

MTT assays take time and planning to set up but give very quick results and are well suited to high throughput screening as they can test multiple different growth parameters in the same assay[85]. Cells can be plated down at a variety of concentrations and in different media. Plates are read automatically by a plate reader, measuring absorbance and giving precise results each time[87]. The exact mechanism of MTT conversion is not understood but is thought to involve MTT gaining electrons during reactions with reduced nicotinamide adenine dinucleotide, thus MTT assays provide only an indirect measure of cell number[85 86].

The MTT assay was obtained from ThermoFisher Scientific Inc.[88] and their quick protocol option followed (adapted from Twentyman and Luscombe[89]).

A 12 mM MTT stock solution was prepared by mixing 1 mL sterile PBS with 5 mg of MTT powder. This was mixed by vortexing until all the powder dissolved and then filter sterilised using a 0.2 µm filter. The stock solution was aliquoted and stored at 4 °C in centrifuge tubes wrapped in foil to protect the solution from light.

Cells were plated in a 96-well plate (density specified in individual results sections) in 200 µl of medium and left for 24 hours. A negative control of media only, with no cells, was included in every experiment.

The 12 mM MTT stock solution was mixed with the appropriate medium to give 10 µl MTT and 100 µl medium for every 110 µl of solution. This solution was vortexed and 110 µl of

MTT/medium solution was added to each well. The plate was incubated in the dark at 37 °C in 5 % CO_2 for 4 hours.

After incubation, 85 µl of solution was removed from each well and 50 µl of sterile dimethyl sulfoxide (DMSO) added. The plate was then placed on a plate shaker for 20 seconds before being incubated for 10 minutes at 37 °C. After which time the plate was put back on the plate shaker for 10 seconds and then placed on an Emax precision microplate reader and absorbance recorded at 540 nm.

The mean average absorbance value of the control (media only) was calculated and subtracted from each individual wells absorbance value to give the live cell concentration of each well, accounting for optical density differences between media and allowing comparison of results between groups and repeats. These standardised absorbance values have been used in statistical testing of comparison between groups.

2.5 Cell cryogenics

Being able to cryogenically preserve cells means cell lines can be stored long term and defrosted for use at a later date. Cell lines in continuous cell culture are prone to genetic drift. Cryogenic storage of cell lines means passage numbers can be maintained at a low range, reducing genetic drift. Cell cultures are also susceptible to contamination and accidents that could destroy them. Having frozen stock ensures cultures can always be re-initiated[64].

2.5.1 Freezing cells

Methods used in the cryopreservation of cells have been adapted from the techniques outlined by Yokoyama et al (2012) as well as standardised techniques used in the laboratory[90]. Cells were passaged and their concentration calculated using a haemocytometer.

The cell solution was then placed in a 15 mL centrifuge tube and centrifuged at 1500 rpm for 5 minutes to pellet the cells. The supernatant was discarded and the cell pellet re-suspended in a solution made of 10 % DMSO and 90 % FBS. The volume of solution was calculated to give a final cell concentration of approximately $1 \times 10^5 – 10 \times 10^5$ cells/mL.

1 mL aliquots of cell solution were then placed into cryovials. Cryovials were clearly labelled with the name, date, cell line, passage number and medium before being placed into the CoolCell® Freezing Container and the -80 °C freezer for 24 hours. After 24 hours the CoolCell®

container was removed and the cryovials relocated onto aluminium racks before being placed into liquid nitrogen.

2.5.2 Defrosting cells

To defrost cells the required cryovials were removed from the liquid nitrogen storage tanks and placed into a 37 °C water bath for 5 minutes. Pre-warmed medium was added to a culture flask and the defrosted cell solution added to this. The solution was then gently pipetted to ensure cell dispersal. The culture was left for 24 hours to allow cell adherence. After which time the medium was changed to remove all traces of DMSO from culture and the cells checked to ensure viable cells had adhered to the flask.

2.6 Contamination testing

All culture techniques must be carried out using strict aseptic techniques, this is because many contaminants such as yeast or bacteria have a much higher growth rate than animal cells. Even with aseptic conditions care must be taken to avoid cross-contamination between cell stocks. If cross-contamination between cell lines does occur, one cell type may outcompete the other. Experiments would become compromised as it would not be possible to state which cell type was being tested[64 91 92].

Current guidelines strongly recommend routine culture is performed without antibiotics and that approach has been used throughout this project[64]. Adding antibiotics to culture media encourages the evolution of antibiotic resistant organisms, can hide low level contaminants and has anti-metabolic effects[93].

Every time cultures were removed from the incubator they were placed on an optical microscope and checked for signs of infection. These checks included looking at the colour and opacity of the media (yellow/white cloudy media indicates infection) and checking cell growth, the number and morphology of cells.

Contamination of cell lines by mycoplasma remains a major problem in cell culture. Using antibiotics in culture medium masks the infection, so cell lines may be unknowingly contaminated[64 93]. Mycoplasmas are a group of microorganisms between 0.3 and 0.8 µm in size, with no rigid cell wall. Their small size and flexible membrane mean they pass through most commonly used bacteriological filters[94]. They have a very wide range of effects on cell

culture which depend on the mycoplasma species, culture conditions, type of infected cell line and the intensity and duration of the infection. It is estimated that 15 to 35 % of continuous cell lines are contaminated[94].

All filters used throughout this project (0.2 µm) were deliberately chosen to reduce the risk of mycoplasma infection[94]. To ensure contamination free culture, all cell lines were tested every 3 months by lab technicians. This was done by testing the supernatant of the cell culture media for the presence of enzymes produced by the majority of mycoplasma species. If contamination was found in any cultures within the building, all cell cultures had to undergo PCR testing[95].

A mycoplasma infection was detected on one occasion during this project. To ensure no re-occurrence of the infection the source of the problem (mechanical failure of equipment) was identified and fixed. The MCF7-2 cell lines being cultured at the time were destroyed and new cell lines were then obtained from Public Health England.

Whilst not as common as microbial contamination cross-contamination of cell lines can occur. This is where cell culture becomes contaminated with multiple cell lines. Fast growing cell lines (such as HeLa cells) can establish themselves in other cultures and outcompete the original cell line[64]. Obtaining cell lines from reputable sources and practicing excellent aseptic technique reduces the risk[64]. In this project multiple cell lines where never cultured in the same hood unless needed for an experiment (where multiple lines had to be plated on the same well-plate). When this occurred, the cultures used for the experiment where then disposed of. Each cell line culture was checked for characteristics specific to that cell line and experimental results compared between repeats to ensure consistency.

2.7 Microscopy

Optical, or light microscopy is a type of microscope that uses visible light with a system of lenses to magnify small images. An objective lens close to the object being viewed collects light reflected from the specimen and focuses it to a real image. This image is then magnified by a second lens called the eyepiece, giving an enlarged, and inverted, virtual image[96].

Inverted microscopes place the light source above the stage (the platform on which the specimen sits) with the objective and eyepiece lenses below the stage. Inverted microscopes are used to observe living cells which adhere to the bottom of the culture flask[96].

Fluorescence microscopy works by illuminating the fluorescently labelled samples with light of a specific wavelength, the excited fluorophores then emit light of longer wavelengths, which can then be detected[96]. This enables specific regions in the samples, those areas that take up the fluorophores, to be observed.

Time-lapse microscopy is when images seen by a microscope are captured using photography. A series of photos can be captured over time and viewed in sequence, at speed, to create a video of microscopic processes[87].

2.7.1 Time-lapse microscopy

Cells were passaged and their concentration calculated before being plated onto a well-plate. Wells for each cell line or condition were distributed around the plate (non-linearly) and cultures left for 24 hours before being placed onto a Nikon TiE time-lapse system microscope. All cultures had their media changed approximately one hour before being placed on the microscope. An environmental chamber was maintained at 37 °C and CO_2 pumped into the well plate within the microscope to maintain the cells in their culture conditions.

Points were chosen randomly within each well and NIS elements photograph software used to capture images of the same points at set time intervals for the duration of the experiment. Unless otherwise specified a x 4 objective lens was used. NIS software converted these images into a video file for each point, ImageJ and MtrackJ[97-99] were then used to analyse the videos.

Using bright-field imaging to photograph live cells can provide information on cell shape, position and motility. By taking photos at short time intervals (between 15 to 20 minutes), cells can be tracked between frames and cell division noted, allowing cell lineages to be tracked[87]. The exact parameters for each experiment are specified within their specific methods sections.

2.8 ImageJ analysis

Fiji and MTrackJ were used to analyse time-lapse videos[97-99]. Fiji – Fiji Is Just ImageJ, is a distribution of ImageJ[97]. ImageJ is an open source programme inspired by NIH image and designed for processing and analysis of multidimensional images[98 100]. NIH Image, the predecessor to ImageJ, was created by Wayne Rasband at the National Institute of Health in 1987[100]. Its success over the last 30 years is partly due to the thousands of plugins and scripts

that can be added to ImageJ. These add-ons mean it has a wide range of complex functions reflecting its widespread use throughout a range of fields in both science and engineering[97 98 100].

Fiji, also an open source programme. It is completely compatible with ImageJ and supports the installation and maintenance of ImageJ plugins[97]. Fiji parcels all the required constituents of ImageJ into a self-contained package and collects plugins, organising them into categories[97]. This makes Fiji easier to install, run and maintain.

MtrackJ is a plugin designed to facilitate motion tracking and analysis of moving objects in image sequences[99].MTrackJ was written and maintained by Erik Meijering[99] and was developed at the Biomedical Imaging Group Rotterdam of the Erasmus Medical Center in the Netherlands. Using Fiji simplified the installation and maintenance of the MtrackJ plugin as there was only one download and updates were automatic[101].

MtrackJ allowed individual cells to be tracked over the course of the experiment. Tracking cells involves marking their exact position in each frame of the video. As stated previously, videos represented a physical position on a well plate. Each video was labelled numerically. Cell tracks within these videos were also labelled numerically in the order they were created. Each time a cells position was marked this registered as a new point within a track. Points within tracks were also labelled numerically in chronological order. This method allows for identification of the experimental parameters as well as the individual cells location in both time and space. In principle there are no limitations to the number of cells that could be tracked or the number of positions in each track (other than computer memory) [101].

A zoom function made it possible to magnify the field of view and enlarge the object or area of interest. The time frame on videos could also be scrolled through making it possible to observe and assess where the object of interest is going[101].

Pixel calibration had to be checked before any track measurements could be taken. A scale was added so that pixel distances were measured in microns[97]. All x and y co-ordinates (starting at 0, 0 of each individual track) were precisely registered. The tracking and display of these co-ordinates is the inverse of the zooming factor; an image displayed at 400 % has precision of ¼ of a pixel[101]. A time scale was also added for each video so that each frame (t) related to the time measurements in the experiment. Measurements can be obtained of each cell track as well as each point within a track.

2.9 Cell tracking

When tracking cells, video files were opened in ImageJ and then MtrackJ was opened. The video file would be set for the first image in the sequence. New tracks could be initiated by moving the computer mouse to the cell of interest and pressing the left mouse button once. This point is then marked by an overlay shape and the programme automatically moves to the next image in the video. The cell track is formed by repeating this process and marking the nuclues of the cell in every single image in the video[101]. As each image is of the same field of view and represents a new point in time the cell is tracked over distance (measured in microns) and through time (measured in hours). To allow accurate calculation of cell trait parameters all cells used in this project had to be observed for a minimum of 4 points.

For each cell there were four possible outcomes that would finish a track. A cell could divide, it could die, it could move out of the field of view or the video ended before any of the other outcomes. Cells were tracked until one of these outcomes was reached. Tracks where terminated by quickly double clicking (less than 200 milliseconds between clicks) the left-hand mouse button[101]. If a cell became stationary it was possible to move through the images and determine if this was just natural cell movement or whether it was about to divide or die. What happened to the cell, (it's 'fate') was manually recorded and a note made of why the track had been terminated – cell divided, cell died, cell moved off screen or cell did not divide (the experiment ended before any of the other outcomes).

Once a cell had been tracked and a record made of its fate, a new cell could be tracked. The video was returned to the first time frame and a new track initiated following the original process[101]. When a cell underwent division it was (usually) possible to track both daughter cells. The original cell track was terminated and a new track initiated following one of the daughter cells. When the daughter cells track ended and the cell fate had been recorded the video could be taken back to the point in time at which the original division occurred. A new track would be initiated following the second daughter cell. Tracks were labelled numerically in order of creation allowing a manual record to be maintained of which cells were related to each other and how. If either or both of the daughter cells divided the same process would be followed with each of them. Manual recording of cell fate and relatedness allowed subsequent statistical comparisons between groups of cells. All cells present in the field of view at the start of the experiment were treated as the first generation, irrespective of their stage in the cell

cycle. Each cell division and the subsequent daughter cells were treated as a new generation for that family, generations were labelled in ascending numerical order.

For each experiment video files were never processed numerically or in groups related to the experimental parameters. This was to ensure no bias in analysis between repeats, e.g. that video file 1 was always processed at the start of the analysis and in a morning whereas video file 50 was always analysed last thing on a Friday 3 weeks later. Which cell to track was chosen at random and the minimum N for each experimental group was decided in advance.

MTrackJ only allows each track to have one point in each frame. Meaning it would be impossible for a cell to have been accidently measured twice and given two positions at the same time. It is also possible to delete both entire tracks and individual points within tracks and to merge or split existing tracks, this allowed any mistakes during cell tracking to be corrected[101].

Measurements for each track and every point within a track were collected and saved[101]. For cell tracks, the ID number, total number of points, total length and duration of the track, the average distance travelled from the start point and the average distance travelled between each point in time were all extracted and added to the information recorded about cell fate and relatedness. A note was made about which video and hence which experiment the data came from and further information regarding the specific cell line and experimental parameters was added.

For the individual points within each track measurements included; the track number each point came from as well as a point ID number, x, y and t co-ordinates, all measured from the track start point (0, 0, 0), the length of the track from the first point to the current point (inclusive), the distance from the first point of the track to the current point and the distance from the current to the previous point in the track[101].

2.10 Statistics

All statistical tests have been conducted using either IBM SPSS statistics[102] or Minitab[103]. Both are widely used for statistical analysis in many fields and can conduct many different types of statistical tests. Numerous statistical techniques were used throughout this project. All data, irrespective of the type of analysis, was examined to determine the distribution and for the presence of outliers.

2.10.1 Data distribution

Descriptive statistics are a collection of values about the data that will give an idea of its distribution. Commonly these values include the mean, median, standard deviation, range, interquartile range, variance and the minimum and maximum values[104]. A normal distribution describes data distributed in a symmetrical, bell-shaped curve, with the greatest frequency of values in the middle and smaller frequency of values towards the extremes. Non-normal distributions can occur in many ways but are classed as any deviation from a 'normal' distribution. Skewness and Kurtosis values were obtained as part of the descriptive statistics. The skewness value is an indication of the symmetry of the distribution and kurtosis represents the peakedness of the distribution. An ideal normal distribution would give skewness and kurtosis values of 0. If a sample size is large enough, the impact of skewness and kurtosis on the analysis may be reduced[104].

The Lilliefors correction of the Kolmogorov-Smirnov is a non-parametric technique used to assess the distribution of a sample. This test compares and tests for significant differences between the distribution of the data and a normal distribution. The Lilliefors correction means that the normal distribution being compared is not specified. As such a significant result violates the assumption of normality, indicating the data is non-normally distributed[104]. This test was chosen over other normality tests (such as Shapiro-Wilk and Anderson-Darling) as even though it has slightly less statistical power than alternative techniques it makes less assumptions about the data.

Histograms, box-plots and Q-Q plots were used to visualise the data distribution. Q-Q plots are probability plots which compare two probability distributions by plotting their quartiles against each other. Normal Q-Q Plots compared the data values against the expected value from a normal distribution, a straight line suggests a normal distribution. Detrended Q-Q Plots plotted the deviation of the data values from the normal distribution line. These points should not cluster together and should generally lie around zero. Q-Q Plots give an idea of whether data is normally distributed and the skewness and kurtosis of data[104]. Histograms are used to display the distribution of a single continuous variable. The shape of the histogram provides information about the distribution[104]. Boxplots also display the distribution of a continuous variable. They are useful when comparing the distributions of different data sets. Boxplots also allow representation of different groups of data[104].

Where data is non-normally distributed the median and 95 % confidence intervals are used to represent the data. Confidence intervals are estimated from the sample data to calculate the interval of plausible values for the true value of the population parameter. The confidence level (in this case 95 %) represents the frequency of possible confidence intervals that contain this true value. Where data has a normal distribution the mean average along with the standard error have been used to represent the data. The standard error was chosen as this represents the standard deviation of the sampling distribution and data presented in this project is a combination of independent repeats containing multiple mathematical repeats.

2.10.2 Outliers

Using a combination of techniques to assess data distribution allows a more accurate representation of the data to be obtained. The presence and number of outliers and whether they impact on data distribution can then be assessed[104].

Many statistical tests will be affected by the presence of outliers. It is not always feasible to remove outliers from the data sets in this project as they may represent natural variation within a population. The presence of outliers can be checked using histograms and boxplots. Boxplots were used to show the presence of outliers and the distribution of data. Histograms were used to indicate if outliers were affecting the distribution of the data. If the tails of the histogram dropped away in an even slope, the chance that outliers were affecting the distribution was reduced. Comparison of the 5 % trimmed mean (where the top and bottom 5 % of values are removed and a new mean calculated) and the original mean also gave an idea as to whether extreme values were influencing the data distribution[104].

2.10.3 Transforming data

Transforming data involves performing mathematical operations on the data to make the distribution appear more normal. The type of transformation undergone depends on the original distribution of the data. Once transformed, data gains a normal distribution and parametric tests can be used.

Non-normal distributions could occur for one of two reasons; either the nature of cancer cell motility itself has a non-normal distribution or there are outliers present[105]. In either case mathematically transforming data could fundamentally alter it and lead to a misinterpretation of the results[106 107]. In this project outliers cannot be removed from the initial data set as they

represent physiological differences in cell types. Whilst being classed as outlying data the few cells that exhibit an increase or decrease from the rest of the population may represent clinically important subsets. The most common types of data transformation (Log, square root and inverse) act by reducing the relative spacing of scores on one side (the side data is skewed towards) more than the other[105]. This can have important effects if the relative distance or order between different values is important. The minimum value in a data set can also alter how effective a transformation is. As cell lines (and even individual populations) show differences between the cells measured, transformations would affect all the data sets in slightly different ways[105].

None of the data in this project has been transformed, instead results of the normality test along with the descriptive statistics have been given and used to make informed decisions on the most suitable techniques needed to analyse the data and the appropriateness of excluding outliers for each of those techniques.

2.10.4 Non-parametric statistical tests

Non-parametric statistics either make no assumptions about the distribution or leave the distributions parameters unspecified, they are also advantageous when the data is better represented by the median. Non-parametric tests can be used on non-normally distributed data (they compare the median which can be less affected by skew), small N, ordinal or ranked data or when outliers cannot be removed.

Not all outliers could be removed from the data sets and not all data was normally distributed. Non-parametric tests can have less statistical power however the large N seen throughout these experiments minimises the chances of a type 2 error (failure to reject a null hypothesis)[102 103]. Non-parametric versions of between groups analysis, Kruskal-Wallis and Mann-Whitney U tests, convert data into ranks meaning the actual distribution of scores does not matter and makes these test suitable for comparisons of data with different distributions[104].

The Kruskal-Wallis test is the non-parametric alternative to a one-way between-groups analysis of variance. It allows comparison of a cell trait for three or more groups. This technique converts the cell traits of each group into ranks and compares the mean rank of each group[104].

The Mann-Whitney U test is similar to a Kruskal-Wallis test but only allows comparisons of cell traits between two groups. It is the non-parametric alternative to a T-test for independent samples. The test looks for differences between two independent groups by comparing their medians. As with the Kruskal-Wallis test the data is converted into ranks and then tested for significant differences between the two groups[104].

2.10.5 Examining the data

Comparisons were made between experimental repeats (the same experiment repeated at a different time) as well as between mathematical repeats (replicate wells within one experiment). General environmental variation, such as variation in media component concentrations, could be caused by differences between the different wells or between experimental repeats. Statistical analyses were conducted comparing mathematical and experimental repeats for significant differences. No significant differences were found between repeats of any experiments so all replicate data was grouped together for further analysis.

Not all cells were observed for a full cell cycle (cytokinesis to cytokinesis or death). There were four outcomes for each cell, they could die, divide, move off screen or still be present at the end of the video. Comparisons between cells with different fates were done separately for each cell line or group. Any differences between the cell fates may explain the presence of outliers in the data.

For each cell line or group comparisons were also made between the different generations of cells. This ensured any changes over time would be detected. All cells present at the start of the experiment where classed as the first generation, each subsequent division was classed as a new generation. The generations to be examined were capped at three for each cell line, after this point N became very small and generation number varied dramatically between cell lines.

2.10.6 Ordinary least squares regression

Regression is a group of techniques used to explore the relationship between a continuous variable and a number of predictors. Ordinary least-squares regression estimates the unknown parameters in a linear regression model, using the least squares principle: minimising the sum of the squares of the differences between the observed dependent variable in the dataset and

those predicted by the linear function[104 108]. Broad-sense heritability (H^2) was estimated as the slope parameter of an ordinary least squares regression for three cell-cell relationships (see figure 3.2). Where a cell had multiple daughters or cousins their mean value was used. H^2 was estimated using simple unweighted regression of trait values between related cells[106].

Ordinary least squares regression is a parametric test with no non-parametric alternative and many stringent assumptions. Data had to be carefully screened to test for normality, linearity and homoscedasticity. Normality for this technique was checked using a normal probability plot of the regression standardised residuals. Scatterplots of the standardised residuals were used to check for outliers which were then checked using Mahalanobis and Cooks distances. Mahalanobis distance identifies multivariate outliers, it is a measure of the distance between a point P and a distribution D. It measures the number of standard deviations from P to the mean of D. Cook's distance estimates the influence of a data point when performing a least-squares regression analysis. It summarises how much all the values in the regression model change when the observation under question is removed[104]. Any outliers or influential data points identified after using Mahalanobis and Cooks distances were removed from the analysis in order to use this technique. Whilst outliers do represent natural variation in the cell population their presence could be explained by bias in the data identified during initial examinations (section 2.10.5).

2.10.6 Bonferroni correction

Conducting more planned comparisons increases the risk of type 1 errors – rejecting the null hypothesis when it is actually true. To avoid this, where appropriate a Bonferroni adjustment has been applied. This involves dividing the significant p-value (0.05) by the number of comparison tests that will be used. The revised p-value is now used to determine significance[104].

Chapter 3 - The cancer cell trait motility is highly heritable

3.1 Introduction

Tumour evolution depends on heritable differences between cells in traits affecting survival and proliferation. The evolution of metastasis would require cells to vary heritably in their dispersal behaviour and the amount of heritable variation would determine the response to selection[45]. Such assumptions have often been made; for example, models of intratumour competition assume heterogeneous populations with heritable differences between cells[19 109]. However, heritability (a statistic summarising the amount of variation in a phenotype due to genetic factors) [46 47] has never been directly measured for any trait in cancer cell populations. For this reason, the heritability of phenotypic variation in the cancer cell trait of motility has been measured. The cell trait motility was chosen, as despite it having an uncertain fitness advantage, it is a fundamental first step in the invasion metastasis cascade[21].

The broad-sense heritability of motility is estimated for cancer cell lines *in vitro* using simple unweighted regression of cell speed between related individuals and found to be significantly heritable. These estimates are stable across multiple cell generations and between different relationships. These results confirm a central yet untested assumption of cancer evolution, providing proof of concept and a solid foundation for the study of cancer evolution.

3.1.1 Natural selection

Natural selection is the differential survival and/or reproduction of individuals due to variation in their phenotypes. It is a key concept in evolutionary biology first formulated by Charles Darwin[14 110] and combined with Mendelian genetics and developed mathematically by Ronald Fisher[111].

Some of the variation of a phenotype in a population occurs due to random genetic mutations. Natural selection acts on the phenotype and the genetic (heritable) basis of that phenotype changes frequency within a population. As natural selection acts and increases survival and reproduction of individuals with certain traits, the population evolves and adapts to the environment. This is a continuous process as even if the environment remains uniform, genetic mutations will always introduce more variation[14]. Natural selection cannot occur without inherited variation in fitness. Whilst many studies have shown both phenotypic and

genetic variation within a tumour[7 29 112] none have yet measured the heritability of cellular traits in a cancer cell population.

3.1.2 Phenotypic plasticity

Phenotypic plasticity is a change in phenotype, associated with a given genotype, in response to different environmental conditions. These genomic interactions with the environment cause variation in phenotypic traits between individuals[14].

Within evolutionary biology the knowledge of the causes and mechanisms of phenotypic plasticity in genotypes due to environmental variation is considered crucial to understanding the evolution and maintenance of biodiversity[15]. A heterogeneous population can only cope with a certain amount of environmental variation. When phenotypic plasticity is no longer enough to survive the altered environment, selective pressures will drive the evolution of a new phenotype through increased mutations[14 113].

Theoretical and ecological observations tend to agree that populations that are phenotypically and genetically diverse occupy broader ecological niches[14 15]. Increased diversity appears to increase population growth, success of establishment and invasiveness and to decrease vulnerability to alterations in the environment[15]. These observations seem to apply to cancer and suggest that the phenotypic and genetic diversity seen in tumours contributes to the severity of the disease. Phenotypic plasticity would also explain how clonal cells can show a variety of phenotypes in a varied tumour microenvironment.

Whilst cancer progression is accepted to be an evolutionary process involving natural selection[112 114 115] caution must be taken when inferring selection in cancer. Cancer cells develop from a multicellular organism and as such have complex gene regulatory networks capable of providing a range of responses to their environment via phenotypic plasticity that could be mistakenly attributed to natural selection[116].

3.1.3 Cancer evolution

The links between evolution and tumours have long been noted[114]. Cancer is the result of natural selection favouring increased proliferation of cells within an individual[1]. Dynamic genetic changes along with clonal expansion and selection leads to a heterogeneous tumour. This multi-step process has been likened to Darwinian evolution with individual cells being the

unit of selection[6]. Tumour progression has been linked to genetic heterogeneity and the tumour microenvironment may cause or select for increased genetic instability[17]. Within a tumour, heterogeneous cells compete for space and resources, evade the immune system and disperse to colonise new tissues[29].

Tumour progression occurs through genetic and epigenetic changes in the constituent cells which alter phenotypes and leads to the acquisition of the 'cancer hallmarks'[6 7]. These are six alterations to cell physiology; self-sufficiency in growth signals, insensitivity to growth inhibitory signals, evasion of apoptosis, limitless replicative potential, sustained angiogenesis and tissue invasion and metastasis[1].

Metastatic tumours are the final stage of the invasion-metastasis cascade[21], which involves dissemination of cancer cells to distant sites around the body and their subsequent adaptation to these foreign tissues. These events are driven by genetic and/or epigenetic changes within the tumour cell as well as alterations in the tumour microenvironment[21]. As metastasis is a multistage process it is difficult to model *in vitro*. Experimental design can be used to break it down into its constituent phenotypes. Motility is thought to be a key parameter in metastasis as cells must migrate away from the primary tumour into circulation and also migrate from the circulatory system into the new site[21].

Tumour metastasis is a common phenomenon in cancer[25]. From an evolutionary perspective this is puzzling as unlike the other cancer hallmarks metastasis does not directly enhance cell survival or proliferation[1]. Metastatic cells 'waste' some of their replicative energy on invasion which has little to no chance of establishing secondary tumours[19 27].

Metastasis is one of the main clinical challenges in treating cancer and our understanding of how and why metastasis occurs is limited. Most of the current hypotheses rely on natural selection[25 117]. Natural selection is the only process that can lead to adaptation and in order for natural selection to act on a population there must be inherited variation in fitness leading to differential survival or reproduction[14 29].

3.1.4 Heritability

Heritability is a statistic that summarises the amount of variation in a phenotype due to genetic factors[46 47]. If a trait is not heritable then natural selection cannot act on it. However, no-one has tested whether cancer phenotypes are heritable at the cellular level.

Introduced by Sewall Right[118] and Ronald Fisher[119] the concept of heritability and its definition remain key for the response to selection in evolutionary biology. Only if cancer traits are heritable can natural selection act on tumours. If heritable, then estimates will help determine a traits response to natural selection and could aid in predicting tumour responses to therapy.

Heritability estimates allow comparison of the relative importance of genetics on the variation in phenotypic traits within a population[46]. Variance is a measure of the spread of numerical observations. It is usually calculated as the average of the squared deviation from the mean. Heritability estimates the degree of phenotypic variation in a population that can be attributed to genetic variation[14]. It is estimated by comparing individual variation in a specific phenotype between differently related individuals within a population. The heritability is the proportion of this total variation due to genetics and can range from 0 (no genetic contribution to the variability) to 1 (all variation due to genetics) [46 47 113]. Heritability estimates are relative to the genetic and environmental factors in a population[14]. In ecology, heritability has been estimated for various traits in many species with traits associated with fitness in natural populations varying in heritability between 0.1 - 0.7[120-125]. Heritability can be measured in two ways, broad- and narrow- sense heritability.

3.1.4.1 Narrow-sense heritability

Narrow-sense heritability (h^2) is the proportion of phenotypic variation (V_P) due to additive genetic variation (V_A) [126].

$$h^2 = \frac{V_A}{V_P}$$

V_A refers to the deviation from the mean phenotype due to the inheritance of particular alleles[127 128]. As the majority of individuals inherit only one copy of each gene from each parent the amount of the genotype inherited by direct descent can vary based on relatedness, making the narrow-sense heritability the best parameter for predicting selection response in sexually reproducing species[129]. As cancer cells are clonal, daughter cells should inherit the full genome of the mother cell, meaning that the narrow-sense heritability would not be an appropriate estimate [46 47 129].

3.1.4.2 Broad-sense heritability

Broad-sense heritability (H^2) is the proportion of phenotypic variation (V_P) that is due to genetic variation (V_G) [47 126].

$$H^2 = \frac{V_G}{V_P}$$

The broad-sense heritability takes into account all the genetic variation; the additive, dominant, epistatic variation and maternal effects that contribute to a population's phenotype[130]. The maternal effects in this experiment are the mother cells microenvironment and cytoplasmic state[46 129 131]. The additive variation of a particular trait depends not only on the allele's frequency in the population but also the effect of each genotype that includes that allele[128]. In sexually reproducing populations each parent usually passes on one allele per locus however in clonal populations all the alleles passed on should be the same between both mother:daughter and sister:sister cells. Both epistatic and dominant genetic variation involve the interactions between alleles. Dominant variance occurs between alternative alleles at specific loci and epistatic variance involves an interaction between alleles at different loci. Both types of variance can change in a population when inheritance of a specific allele affects expression of another[128]. Genetic variation within cancer cell lines should only arise as a result of mutation meaning the additive, dominant and epistatic variation will vary greatly from that of a sexually reproducing population[46 129].

As cancer cells are clonal a modification of standard parent-offspring regression techniques could be used to estimate broad-sense heritability[46 130]. Estimates could be obtained using simple unweighted regression comparing motility values between clonal cells and their relatives[106]. Whilst cancer cells are clonal this does not mean estimating heritability is simple. Cancer cells arise from multicellular organisms and the complexity of gene regulatory pathways and the sensitivity of cancer cells to their environment need to be considered when interpreting results[38].

3.1.5 Phenotypic variation

Phenotypic variation can occur for several reasons. The source of the variation can determine whether a trait has the ability to respond to natural or artificial selection. The relative importance of genetic and environmental factors to a particular trait can be used to predict the evolutionary dynamics of a phenotype in an entire population[128 129].

Phenotypic variation (V_P) can occur between individuals in a population due to the environment (V_E) as well as their genetics (V_G) and is defined as $V_P = V_G + V_E$ [128 129]. Because of this, heritability can alter due to both environmental and genetic variation.

The main source of genetic variation between individual cells would be expected to arise through maternal effects, cytoplasmic factors present before cell division resulting in transient similarities between sister and mother/daughter cells after division[131]. Genetic variation could also occur through genetic and epigenetic mutations arising during division. As a result, genetic variation may increase over time causing more distantly related cells (cousins) to be less similar to each other. Comparison of H^2 estimates for multiple cell:cell relationships and over several generations can be obtained representing stable differences between clonal cell lineages.

General environmental variation is all the non-genetic sources of variation between individuals experienced by multiple individuals within a population[128]. For these *in vitro* experiments there should be no difference in the environment for families of cells greater than that of the population as a whole[130]. Within a cell family environmental variation can occur both spatially – as cells migrate they move into different environments, and temporally – over time cells use resources and produce waste products, as a result the local concentration gradient will be in constant flux and resource concentrations will decrease.

Heritability can increase when genetic variation is high or when environmental variation is low, what matters is the relative contribution[29].

3.1.6 Aims

Both tumours and cancer cell lines have been shown to be genetically heterogeneous[132-134] and the tumour microenvironment applies selective pressures to tumours[135] yet so far, no studies have investigated the heritability of individual cancer cell traits. The aim of this project is to measure heritability of cell motility, a vital component of metastasis[21] yet one with an uncertain selective advantage, in *in vitro* cancer cell lines. Heritability is a pre-requisite for evolution by natural selection[129], if found these results will provide a fundamental first step towards more detailed research into the role of genes on cell motility[131] and in predicting cancer cells response to selective pressures[20].

3.2 Methods

3.2.1 Cell culture

All cell lines used were lab adapted. As stated in section 2.1, similar cell lines have been obtained from different sources, to distinguish between these they are labelled numerically. All cells were cultured as a monolayer in 5 % CO_2 at 37 °c.

MCF7-2 (passage number 16) and MDA-1 cells (passage number 5) were cultured in 75 cm^2 culture flasks (T-75 life technologies) in 15 mL media due to mycoplasma contamination N = 1. MCF7-3 (passage number between 15-18), MDA-2 (passage number between 41-44), HeLa (passage number between 5-8) and HT1080 (passage number between 5-8) cells were cultured in 225 cm^2 culture flasks (T-225 life technologies) in 22 mL media, N=3.

3.2.2 Time-lapse microscopy

Section 2.7. 1 has details about the overall technique of time-lapse microscopy. For these specific experiments, MCF7-2 and MDA-1 cells were plated onto a 12-well plate with 1000-4000 cells per well. 5000 HeLa, HT1080, MCF7-3 and MDA-2 cells per well were plated onto a 24-well plate. For all experiments six wells per cell line were distributed around the plate in 1 mL media.

NIS elements photograph software and a Nikon TiE time-lapse system microscope were used to select five points within each well. Images were taken (as seen in figure 3.1 panels A-D) every 15 minutes for 70 hours (MCF7-2 and MDA-1 cells) and every 20 minutes for just over 72 hours (HeLa, HT1080, MCF7-3 and MDA-2 cells).

NIS elements photograph software was used to capture images of these exact same points at the set time intervals for the duration of the experiments and then convert these 128,880 images into a video file for each point[87]. ImageJ and MtrackJ[97-99] were used to analyse videos and track individual cells, as shown in figure 3.1. In total 1,932 MCF7-2 and MDA-1 cells were tracked and 9,131 HeLa, HT1080, MCF7-3 and MDA-2 cells.

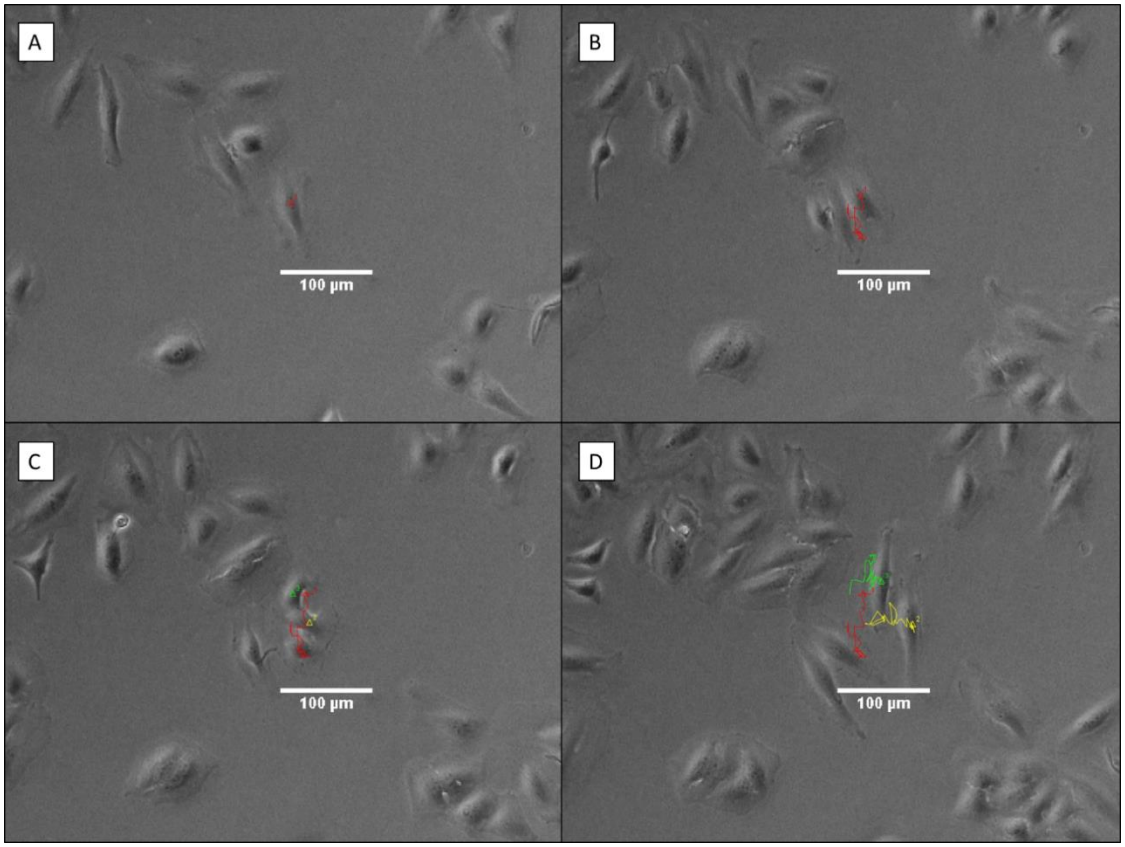


Figure 3.1: Four representative images of HeLa cells on a Nikon Tie time-lapse microscope taken over time (panelled A-D). A – a HeLa cell is chosen and the position of its nuclei marked, MtrackJ automatically numbers the track (1) allowing identification. B – As the HeLa cell moves the position of its nucleus in each frame can be marked. The track length and position change over time and this change is recorded. C – When cell division occurs the first cell track ends. Each daughter cell can then be tracked individually as the first cell was. Each new cell track is also given a unique identification number allowing records to be kept of how cells are related. D – The daughter cells move over time and the position of their nuclei in each frame can be marked. This allows comparison of cell motility between related cells. The images in panels A-D were taken at 20-minute intervals.

3.2.2.1 Large image time-lapse microscopy

An A1-R confocal microscope with super resolution and spectral detection capabilities and a Nikon TE200 time-lapse system were used to try and conduct a large image time-lapse. This involved the same methods as a time-lapse microscopy experiment however NIS elements software was programmed to automatically take multiple photos all bordering the original point chosen. These images could be combined to create a video over time with a much wider field of whilst maintaining the original magnification. These experiments were unsuccessful

due to technical limitations, the software was not able to stitch the images together to form an extended field of view.

3.2.3 Cell tracking

A total of 273 videos were analysed as shown in figure 3.1 using ImageJ and MtrackJ[97-99] this method is outlined in section 2.9 and involves marking the position of the cell's nuclei over time. There were 280 images per video for MCF7-2 and MDA-1 cells and 218 images per video for HeLa, HT1080, MCF7-3 and MDA-2 cells. A total of 11,063 individual cells where tracked with a total of 1,245,411 points within those tracks. As each image is of the same field of view and represents a new point in time the cell is tracked over distance (measured in microns) and through time (measured in hours).

Tracks were labelled numerically in order of creation and a manual record of a cells relatedness (which other cells it was related to and how), generation (all cells present at the start of the experiment where classed as the first generation, each subsequent division was classed as a new generation) and fate (whether it died, divided, moved off screen or was still present at the end of the video) kept throughout the analysis. Comparisons between groups of cells were conducted to ensure consistency within and between experiments and to ensure there was no bias within the sample of cells measured (see section 2.10.5).

Manual tracking of individual cells and recording of generation, fate and relatedness allowed three different cell:cell relationships to be used to calculate estimates of H^2. The three cell:cell relationships can be seen in figure 3.2 and are; Mother:daughter, a progenitor cell to its first generation clonal descendants. Sister:sister, comparing the two cells produced from the same progenitor by cell division. Cousin:cousin, comparing the second generation clonal descendants of one progenitor cell.

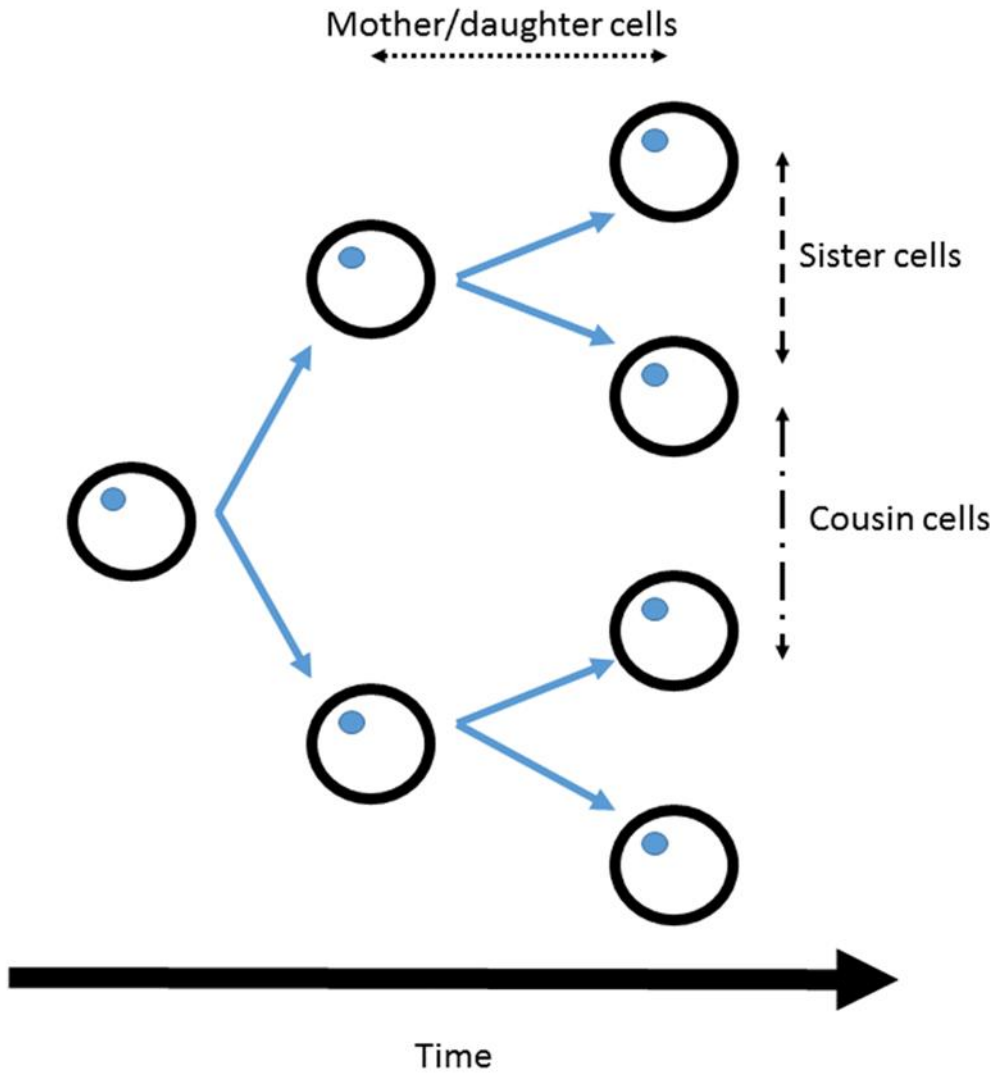


Figure 3.2: Schematic diagram of a cell family over three generations showing the different cell:cell relationships used to calculate broad-sense heritability.

3.2.4 Statistical analysis

All cell track data was collected from MtrackJ[101] and stored in excel files. Manual data was input into excel files throughout data collection. Statistical analysis was performed using Minitab[103] and IBM SPSS[102]. For a full explanation of the techniques chosen see section 2.10.

Cell motility was calculated as a cell's average speed over its lifetime, in microns per hour. Cell speed was measured as the total distance of the path travelled by the cell (in microns), divided by its lifetime (in hours).

The Lilliefors correction of the Kolmogorov-Smirnov test is a non-parametric technique and was used in conjunction with descriptive statistics to determine the distribution of data[104]. Any non-normal results have not undergone transformation as mathematically transforming the original data set could cause fundamental alterations leading to a misinterpretation of the results[106 107].

Significant differences in cell motility between repeats, generations and fates were tested for using Kruskal-Wallis and Mann-Whitney U tests (section 2.10.5). This was done to ensure that any bias in the data was detected; for example, whether cells that moved off screen were significantly faster than those that divided or if cells increased or decreased their speed over time. The generations to be examined were capped at three for each cell line, after this point N became very small and generation number varied dramatically between cell lines. Comparisons were done separately for each cell line. As a result of these examinations some cell groups have not been used in calculations of heritability (see section 3.3.1).

Broad-sense heritability was estimated as the slope parameter of an ordinary least squares regression (section 2.10.6) for three cell-cell relationships (see figure 3.2). Where a cell had multiple daughters or cousins their mean value was used. H^2 was estimated using simple unweighted regression of trait values between related cells[106]. Regression between relatives was not scaled based on relatedness as cancer cells are considered clonal. If significant, genetic variation is occurring between related individuals then only the cousin:cousin regression would need altering as these cells would contain double the amount of genetic variation as mother:daughter and sister:sister cells (two divisions separate these cells as opposed to one division separating the other relationships). Both H^2 and R^2 values have been given in the results. R^2 values show how much of the variance in the dependent variable (cell speed) is explained by the model[104].

3.3 Results

3.3.1 Motility of cell lines

There was no significant difference in speed between the individual wells of each experiment for any cell line (Kruskal-Wallis test; $p>0.07$). As stated in the methods (section 3.2.1) only one experimental repeat could be obtained for the MCF7-2 and MDA-1 cell lines. As seen in figure 3.3 there was no significant difference in speed between the different experimental repeats for HeLa, HT1080, MCF7-3 and MDA-2 cell lines (Kruskal-Wallis test; $p>0.05$). As a result, all the data for each cell line have been grouped together for further analysis.

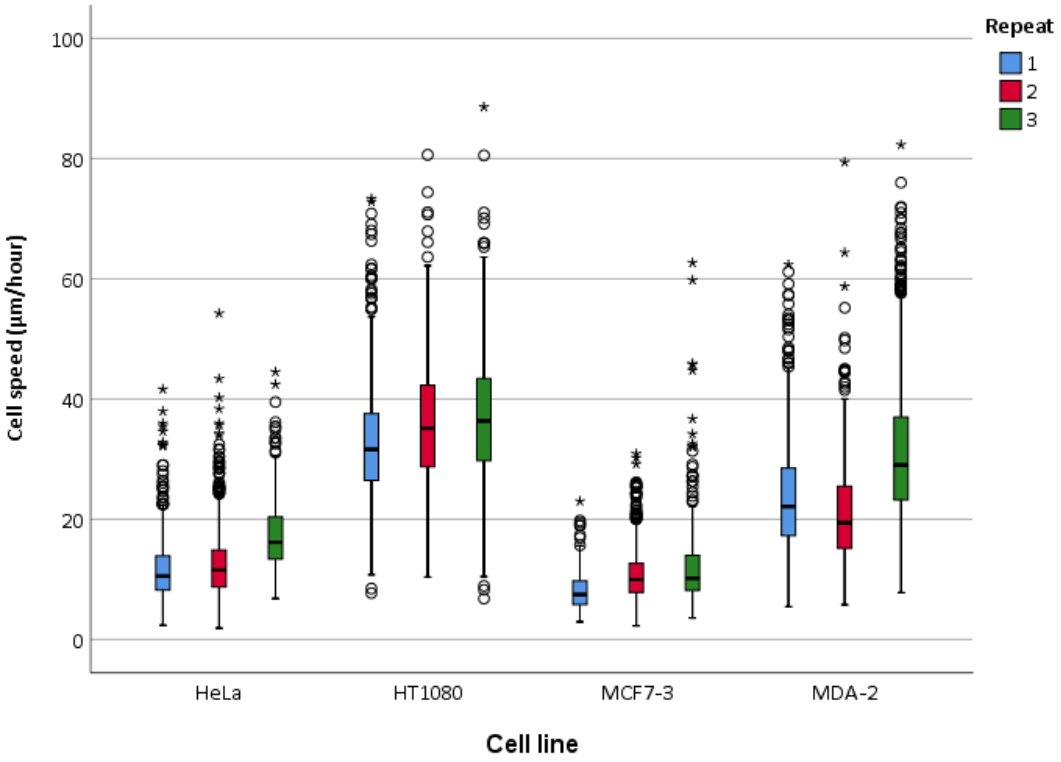


*Figure 3.3: Distribution of cell speeds for HeLa, HT1080, MCF7-3 and MDA-2 cell lines for each experimental repeat. The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. N for each cell line and repeat is; HeLa repeat 1: 1,125, repeat 2: 990, repeat 3: 352, HT1080 repeat 1: 947, repeat 2: 810, repeat 3: 805, MCF7-3 repeat 1: 162, repeat 2: 969, repeat 3: 191, MDA-2 repeat 1: 668, repeat 2: 714, repeat 3: 1070.*

All cell lines had a non-normal distribution of cell speeds (Lilliefor corrected Kolmogorov-Smirnov test; $p<0.001$). There was a significant difference between all cell lines in their speed (Kruskal-Wallis test; $p<0.001$) apart from HT1080 and MDA-1 (Mann-Whitney U test; $p=0.544$). As seen in figure 3.4 and table 3.1; HT1080 and MDA-1 cells had the fastest cell speeds, then MDA-2, MCF7-2 and HeLa with MCF7-3 having the slowest cell speeds. The differences between the MCF7 and MDA-MB-231 groups show that the results of similar cell lines obtained from different sources cannot be grouped together. These groups of cells; MCF7-2, MCF7-3, MDA-1 and MDA-2, will be kept as discrete units and their data tested separately.

Table 3.1: Median speed and N of HeLa, HT1080, MCF7-2, MCF7-3, MDA-1 and MDA-2 cells.

Cell line	**Median Speed (µm/hour)**	**N**
HeLa	11.75	2467
HT1080	34.05	2562
MCF7-2	12.19	1520
MCF7-3	9.76	1322
MDA-1	33.75	411
MDA-2	24.28	2452

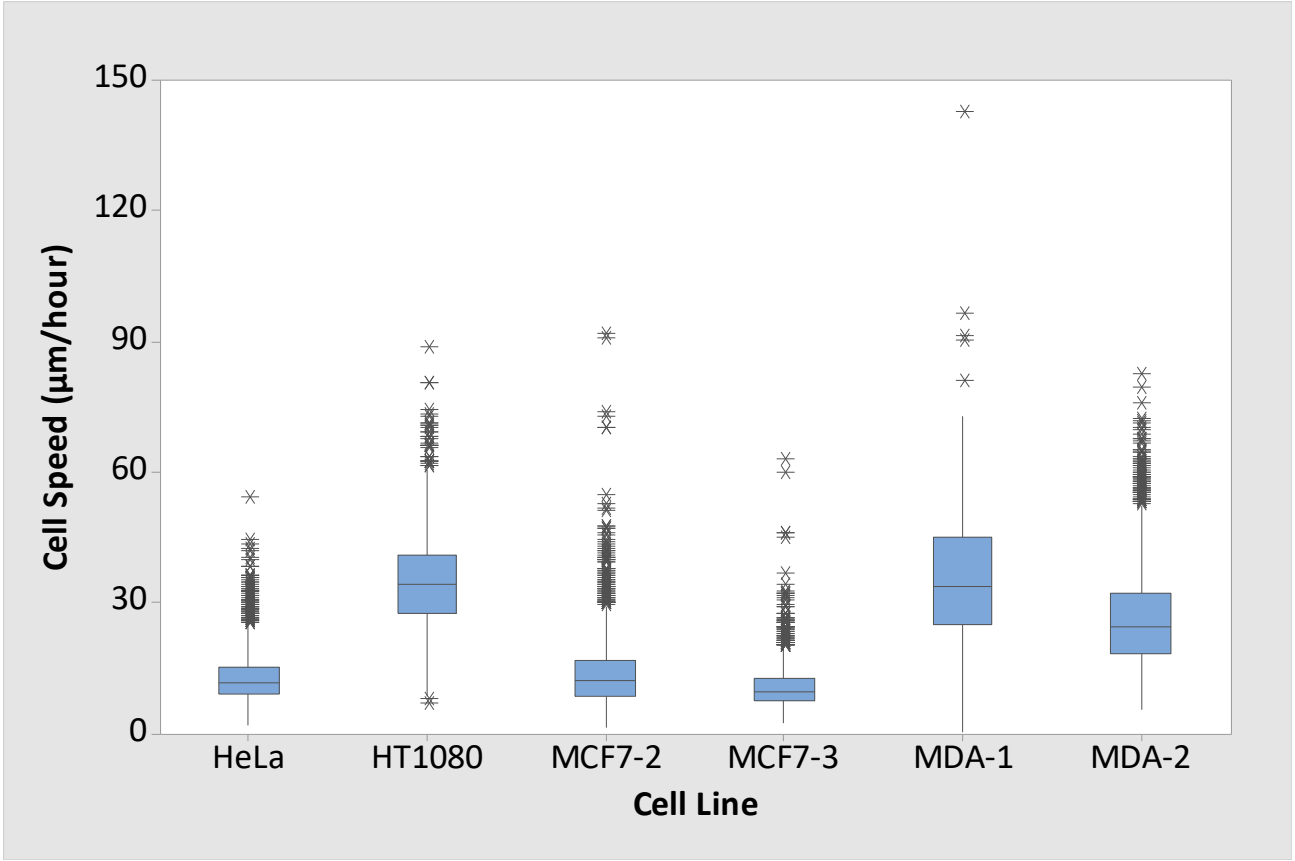


*Figure 3.4: Distribution of cell speeds for HeLa, HT1080, MCF7-2, MCF7-3, MDA-1 and MDA-2 cell lines. The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. Data gathered from three independent experiments for HeLa, HT1080, MCF7-3 and MDA-2 cell lines and one experiment for MCF7-2 and MDA-1 cell lines. N for each cell line is; HeLa 2,467, HT1080 2,562, MCF7-3 1,322, MDA-2 2,452, MCF7-2 1,520 and MDA-1 411.*

Table 3.2 shows the median speed for each of the first three generations measured for each cell line. No significant difference was found in cell speed between the different generations for any cell line (Kruskal-Wallis test, bonferroni correction $p>0.003$), showing that cell speed did not change significantly over the course of the experiment.

Table 3.2: Median speed and N for each of the first three generations of HeLa, HT1080, MCF7-2, MCF7-3, MDA-1 and MDA-2 cells.

Cell Line	**Median Speed (µm/hour) [N]**		
	Generation 1	**Generation 2**	**Generation 3**
HeLa	13.17 [258]	12.61 [495]	11.47 [805]
HT1080	37.94 [227]	37.43 [372]	36.33 [515]
MCF7-2	11.34 [275]	12.5 [455]	11.45 [608]
MCF7-3	9.26 [257]	9.46 [353]	9.58 [374]
MDA-1	35.78 [43]	34.16 [81]	30.17 [139]
MDA-2	24.83 [269]	25.06 [451]	23.60 [655]

Some significant differences were found in cell speed between cells with different fates. Table 3.3 shows HeLa cells that died and MCF7-3 cells that moved off screen had higher speeds than the rest of the population (Mann-Whitney U test; $p<0.001$). These differences in cell fate may explain the presence of outliers or skew heritability estimates as they may represent a bias in the population (see section 2.10.5). As such these cells have not been included in estimates of H^2.

Table 3.3: Median speed and N for each cell fate of HeLa, HT1080, MCF7-2, MCF7-3, MDA-1 and MDA-2 cells.

Cell line	**Median Speed (µm/hour) [N]**			
	Non-divided	**Divided**	**Died**	**Moved off screen**
HeLa	11.42 [1201]	11.73 [1125]	15.32 [130]	12.69 [11]
HT1080	29.71 [484]	34.76 [1216]	29.36 [15]	35.85 [847]
MCF7-2	12.22 [895]	12.16 [625]		
MCF7-3	9.62 [718]	9.98 [718]	8.75 [52]	24.29 [16]
MDA-1	30.47 [226]	36.58 [184]		
MDA-2	22.96 [11260]	25.46 [1129]	23.21 [26]	27.47 [171]

3.3.2 Heritability of motility in cell lines

Cell speed is significantly heritable for all cell:cell relationships in all cell lines. Table 3.4 shows H^2 values range from 0.770 – 0.364 and R^2 values range from 0.59 – 0.13. As stated in section 3.2.4.3 outliers and influential data points have been removed from this analysis. H^2 was estimated using simple unweighted regression of cell speeds between related cells. Figure 3.5 gives an example of how this has been calculated for each of the cell:cell relationships in the MDA-2 cell line.

Table 3.4: Mean trait value, R^2 and broad-sense heritability of motility in HeLa, HT1080, MCF7-2, MCF7-3, MDA-1 and MDA-2 cell lines for three cell:cell relationships.

Cell line	**No. of families**	**Mean speed (µm/hour)**	**Heritability of cell motility H^2 (R^2)**		
			Mother:daughter	**Sister:sister**	**Cousin:cousin**
HeLa	292	12.75	0.706*** (0.5)	0.704*** (0.5)	0.656*** (0.43)
HT1080	214	34.95	0.558*** (0.31)	0.545*** (0.3)	0.364*** (0.13)
MCF7-2	301	14.108	0.573*** (0.33)	0.630*** (0.4)	0.490*** (0.24)
MCF7-3	98	10.73	0.446*** (0.2)	0.564*** (0.32)	0.407** (0.17)
MDA-1	73	35.499	0.760*** (0.58)	0.719*** (0.52)	0.596*** (0.36)
MDA-2	213	26.25	0.770*** (0.59)	0.735*** (0.54)	0.540*** (0.29)

**p<0.01, ***p<0.001

Due to the number of post-hoc comparisons it was not possible to determine significance either between H^2 estimates of the different cell lines or between H^2 estimates of the different families. However as seen in table 3.4 and figure 3.6, in all cell lines the H^2 estimates for the cousin:cousin relationship, whilst still significant, are always lower than those obtained for the sister:sister and mother:daughter relationships. R^2 is the amount of variance shared between the speeds of each group, table 3.4 shows this is also less between cousin:cousin cells than sister:sister or mother:daughter cells. As mentioned in section 3.1.4.2 this could occur for several reasons. Inclusion of multiple different cell:cell relationships and examination over several generations means the significant results seen here represent stable differences between clonal cell lineages.

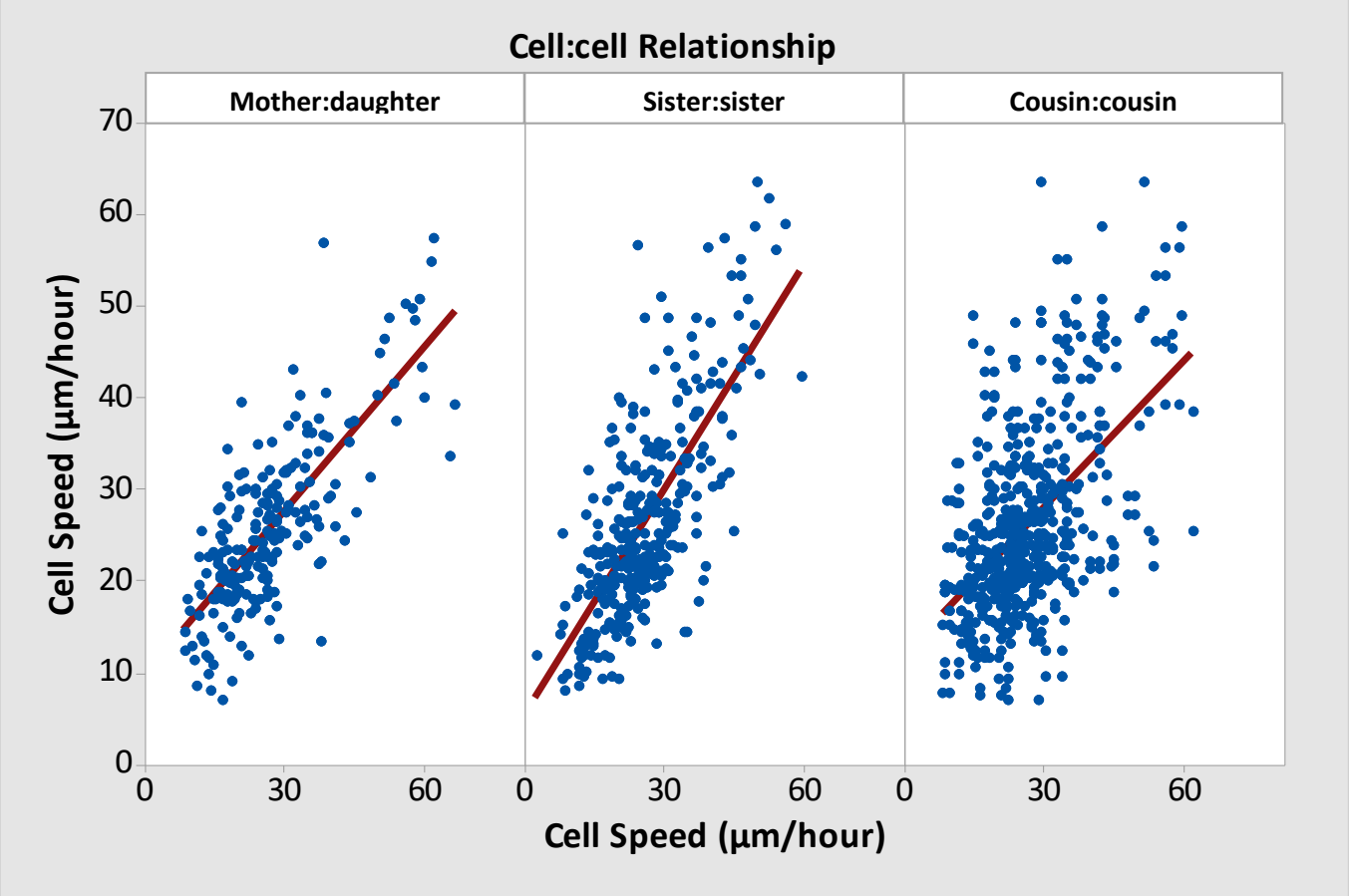


Figure 3.5: Example of the cell:cell regression for motility in the MDA-2 cell line. The slope of the regression line is the estimate of broad-sense heritability. Each point in the mother:daughter regression represents a mother cell and the mean speed of her daughters; in the sister:sister regression each point represents the speed of two cells produced from the same progenitor cell; in the cousin:cousin regression each point represents the second generation clonal descendants of one progenitor cell, where two clonal descendants occur (sister cells) their mean average speed has been used. The N for all cell:cell relationships is 213.

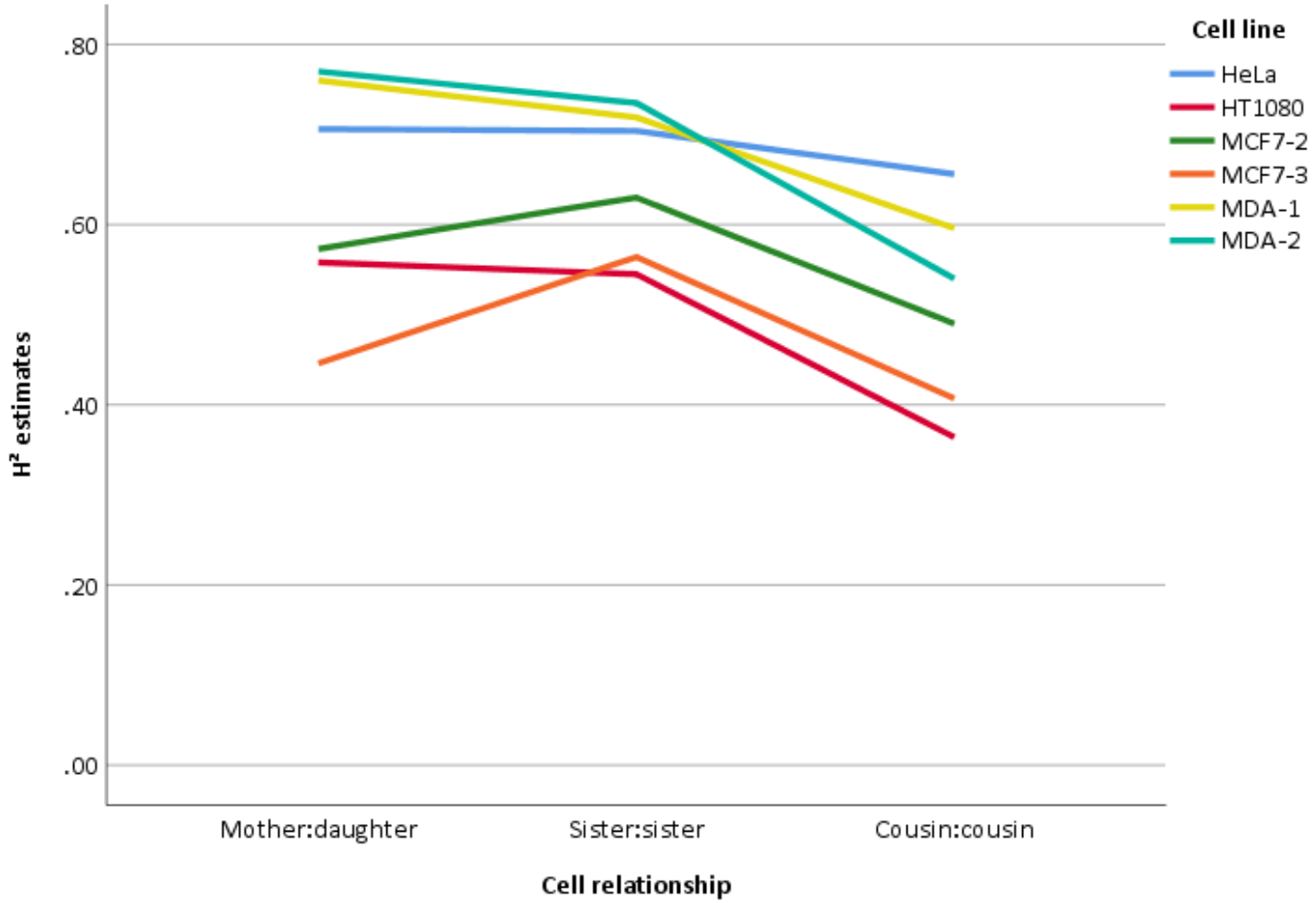


Figure 3.6: Heritability estimates of cell:cell relationships for HeLa, HT1080, MCF7-3, MCF7-2, MDA-1 and MDA-2 cell lines. N for each group is HeLa = 292, HT1080 = 214, MCF7-2 = 301, MCF7- 3 = 98, MDA-1 = 73, MDA-2 = 213.

3.4 Discussion

For all cell lines tested, and for all cell:cell relationships, H^2 of motility is highly significant (table 3.4 and figure 3.5). The novel results seen here identify cancer cell motility as a heritable trait confirming that natural selection can play a role in tumour evolution[17].

3.4.1 Comparing motility between cell lines

As discussed in section 2.10.3 the decision was made not to transform the data in this project[106 107]. Instead extensive analysis was conducted into the distribution of cell motility within and between cell lines. General environmental variation could be caused by differences between the different wells or between experimental repeats[128]. As seen in figure 3.1 when comparing the motility of each cell line between repeats no differences were found, suggesting spatial effects – which could exacerbate similarities between related cells - are minimal[130].

During the experiment there were four outcomes for each cell, they could die, divide, move off screen or still be present at the end of the video. HeLa cells that died and MCF7-3 cells that moved off screen had higher speeds than the rest of the population ($p<0.001$). Unsuccessful attempts were made to track the cells that moved off screen using large image time-lapse microscopy (section 3.2.2.1), this type of microscopy gives a wider field of view and would allow tracking of cells over greater distances[136 137] This was done to try and determine whether the cells that moved off screen were fundamentally faster cells or if they did so by chance, they just happened to be near the edge of the field of view.

Whilst motility is a fundamental prerequisite for metastasis it does not appear to directly enhance cell survival or proliferation[1]. Tracking cells for a longer distance would have helped determine whether their increased motility came at a cost of reduced division or increased chance of death[138]. This cost of motility could also explain why HeLa cells that died had higher motility than other cell fates. If these cell types were inherently faster than the rest of the population then H^2 estimates calculated from the remaining cell fates may not be comparable to the population as a whole[130]. Comparisons were made for the speed of each generation, with no significant differences being found. This showed motility to be consistent over the course of the experiment with no overall increase or decrease in cell speed. Heritability estimates are always specific to the population and environment measured[46], but these results

suggest that the difference in speed between the different cell fates did not affect the overall motility of the population.

Whilst all the cell lines chosen were human cancer cells there are still differences between them (see section 2.1). MCF7 and MDA-MB-231 are genetically distinct breast cancer cell lines[68 70 71]. HeLa cells are a cervical epithelial adenocarcinoma[73 74] and the HT1080 cell line is a fibrosarcoma[77]. There were significant differences in motility between the distinct cell lines (apart from HT1080 and MDA-1) and also within the MDA-MB-231 and the MCF7 cell lines that had been obtained from different sources. Heterogeneity within a cancer cell population along with natural selection and genetic drift may have caused these groups of cells to diverge phenotypically over time[139].

Cancers are known to have elevated mutation rates both in base pair mutations and chromosomal rearrangements[140 141]. There are a range of estimates of the mutation rate for cancer cells from $2 \times 10^{-7} – 4.9 \times 10^{-6}$ mutations per cell per generation[140 141], some studies have estimated that cancer cells within a tumour have a base pair mutation rate per division, 10^5 times higher than non-cancerous somatic cells[140]. Aneuploidy has also been shown to increase the mutation rate due to the imbalance of spindle proteins[142]. Even though cancer cells have been shown to have an elevated mutation rate this may not hold true for cancer cell lines which are cultured in vitro. Haplotype-resolved whole-genome sequencing of HeLa cells indicated that after initial passaging and adaptation to culture conditions there were relatively few point mutations[143]. The range in mutation rates and lack of comprehensive analysis on *in vitro* cancer cell lines makes it difficult to predict the expected mutation rate. However, in highly genetically unstable cell lines mutation rates may be as high as 2 % chromosomal rearrangement and 10^{-4} base pair mutations per cell division[140-143].

It might be expected that cell lines with higher genetic instability would show higher estimates of H^2 as the comparative increase in genetic instability may increase the mutation rate and hence genetic variation. Both the MCF7 and HeLa cell lines are known to be very genetically unstable with many instances of regional copy number increase, extensive genome rearrangement and high levels of chromosome variability[29 76]. If true, then estimates of H^2 would be higher for these cell lines than the others tested. The results in table 3.4 and figure 3.6 show that this is not the case, MDA cells showed the highest estimates of H^2. However, HeLa and MCF7-3 cells were the only cell lines to show a significant difference in motility between the different cell fates. Studies of tumour heterogeneity show that the degree of variability can predict tumour progression[144 145]. This same heterogeneity and variance pose a

problem in the study and treatment of tumours as a biopsy taken from a patient; or a cell line studied in the lab, may not be representative due to natural selection over time[29].

Unfortunately post-hoc comparisons comparing the significance of H^2 estimates between cell lines was not statistically possible due to the amount of tests needed and the increased risk of type 1 errors (see section 2.10.6)[104]. Nor was conducting genetic profiling of the cell lines possible within the scope of this project. The differences in motility within the MCF7 and MDA-MB-231 cell lines suggests that *in vitro* cell lines cultured under similar conditions are likely to have undergone genetic drift and selective pressures causing phenotypic divergence[139].

3.4.2 Implications of finding heritability

Heritability is a pre-requisite for evolution by natural selection[129]. Finding significant heritability shows that in these cell populations in these environments, genetic differences influence the variation in motility between individual cells. This is a statistic that has never before been measured in any cancer cell line. At the cellular level heritability has been demonstrated for proliferation and monoclonal antibody reproduction rates in Chinese hamster ovary cells[146] and in many traits for unicellular organisms such as yeast[147], but never for any trait in cancer cell populations.

Our values of H^2 range from 0.36 to 0.77 (table 3.4). These estimates are at the upper end of H^2 values compared to fitness traits in natural populations[120-125]. Cancer cell lines show greater genetic instability than non-transformed cells[68 69 73]. An increased mutation rate could explain why in general all the H^2 estimates are at the higher end of those seen in natural populations, as mutations would act to increase genetic variation[120-125]. Motility may be a composite trait - more loci influencing the trait means it could display higher genetic variation - explaining the H^2 values[124]. Other studies found that heritability in fitness traits as low as 10 % still provided substantial genetic variance upon which selection can act[124]. These results indicate that cancer cell populations are likely to contain substantial genetic variation for motility upon which natural selection could act[124]. These results provide an essential first step which could lead towards more detailed research into the role of genes on cell motility[131].

Lower estimates of heritability might be expected for a trait more closely linked to cell fitness[123 124]. Cell motility has an uncertain selective advantage[6 19]. Cells that are motile have less energy to use on growth or proliferation, however dispersal reduces competition and if

successful may result in the cell settling in an area with more nutrients. For a fitness trait it would be expected that natural selection would reduce genetic variation as advantageous genes increased in frequency[124]. Metastasis is a multi-stage process resulting in the formation of secondary tumours and motility is thought to be a key parameter as cells must migrate into local tissue, away from the primary tumour, into circulation and out into a secondary site[21]. The relatively high estimates of heritability measured here could support the hypothesis that in this (*in vitro*) environment motility itself is not directly linked to cell fitness and is instead one of many factors in the metastasis cascade. Obtaining heritability estimates for motility has implications in the treatment of metastasis as it increases our understanding of how individual cell traits respond to selective pressures. Better understanding of cancer cell responses to selective pressures, such as chemotherapy treatment, will assist in predictions of disease risk and treatments[129 148].

3.4.3 Factors affecting H^2 in these experiments

Low environmental variation will tend to increase H^2 values. All experiments were performed on cells *in vitro*. A major advantage of *in vitro* cell culture is the ability to precisely control the environment[64]. Cell culture provides an *in vitro* model of cancer in a well-defined environment which can be easily manipulated and analysed[64]. Variations between repeats can be kept to an absolute minimum and experimental parameters repeated or altered with relative ease.

There are limitations to *in vitro* cell culture. No matter how carefully controlled and regulated an *in vitro* environment is, it will always be simpler and lack some elements of an *in vivo* environment[64]. Care must be taken when extrapolating results from *in vitro* experimentation and applying them to an *in vivo* environment[64]. It's important to remember that heritability is specific to a population in an environment so estimates of H^2 *in vivo* would differ from those *in vitro*. The heritability estimates seen here are likely to be higher than those seen *in vivo* due to the less complex environment.

Testing of different cell types and different cancer types makes these results more robust and less specific to particular cell lines or conditions[67]. To reduce the effect of genetic drift and natural selection between the experimental repeats all cells were used within a small range of passage numbers (the number of times each cell line has been cultured). Reducing the amount of time between passages means there has been less time for both genetic drift to occur and natural selection to act[139]. Whilst this does not remove the heterogeneity within the cell population it reduces the variability between repeats. Whilst it's important to recognise the

limitations of these results, low environmental variation could increase H^2 values, finding significant heritability in a cancer cell trait shows that genetic differences have had some influence on the variation in motility between individual cells and is a fundamental first step to adapting experimental evolution techniques across to cancer cell biology[36].

Another factor to consider is the exclusion of some cell fates from heritability calculations. Large image time-lapse microscopy (section 3.2.2.1), whilst unsuccessful in this project due to software errors, could be used to track cells moving out of the field of view and determine whether their increased motility represents a bias in cell fate. The time-consuming nature of data acquisition makes it difficult to simply repeat experiments and increase N. More rigorous statistical analysis could address some of these concerns by performing more detailed comparisons of cell family dynamics[106]. The examination of motility over time within experiments ensures that any change in population dynamics does not affect the results seen here.

3.4.4 Evolvability

No single statistic can adequately describe evolvability (an organism's capacity to generate heritable phenotypic variation) and quantifying evolvability can be difficult as it is not possible to gain a direct measurement[149]. These analyses can only give an indirect clue as to the phenotypic trait under selection.

Measuring the mutation rate within a population may give an indication of the potential evolvability[149 150]. Population genetic models show that in a poorly adapted population, genes increasing the genomic mutation rate can increase in frequency by hitchhiking with the beneficial mutations they produce[150]. Lenski's experiments of *Escherichia coli* populations provided evidence that mutator genes increase in frequency because they increase evolvability[151] however this only applies in poorly adapted populations so any effect would be lost over time[150 151], particularly when using cell culture as cancer cell lines are lab adapted to the *in vitro* culture environment[64].

A popular test to detect natural selection on gene sequences involves comparing the rate of non-synonymous (mutations that alter protein structure and function) to synonymous (mutations that are neutral to the amino acid sequence produced) mutations at a particular locus. The ratio between the mutations is normalised by the total possible number of non-synonymous and synonymous mutations (referred to as dN/dS). Deviations from the

normalized ratio above 1 infer positive selection (more non-synonymous mutations than would be expected randomly) and deviations under 1 infer negative selection (fewer non-synonymous mutations than would be expected randomly) [49]. Applying this type of a statistical assessment to cancer evolution is difficult due to the clonal expansion in cancer cells. A single mutation may occur that drives the cell to undergo clonal expansion, both passenger mutations as well as the driver mutation are increased in frequency making it very difficult to identify the original driver mutation[49].

Heritability is a statistic summarising the amount of variation in a phenotype due to genetic factors[46 47]. It is estimated using parent-offspring regression techniques[106]. Unlike most natural population's cancer cells are clonal meaning the broad-sense heritability is used and estimates obtained via simple unweighted regression[106]. Whilst this statistic is the most appropriate measure of evolvability in this study it still does not give a clear picture as to whether or not motility itself is under selection. Measuring heritability, a fundamental first step for evolution by natural selection, is an essential prerequisite for more detailed research into the role of motility genes in metastasis[131] and predicting the response to selective pressures. As such using heritability as one measure of evolvability paves the way for experimental evolution techniques to be applied to cancer biology[38].

3.4.5 The decrease in heritability between more distantly related cells

These results cannot be considered in isolation. Cancer cells arise from multicellular organisms. They have complex gene regulatory pathways and as such the capacity for intricate, programmed responses to their environment that could be mistakenly attributed to natural selection[116]. The sensitivity of cancer cells to their environment need to be considered when interpreting results[38]. The results in this project show both R^2 and H^2 to be less between cousin:cousin cells than sister:sister or mother:daughter cells (table 3.4, figure 3.6).

R^2 is the amount of variance in cell motility that can be explained by the related cell[104]. Its decrease could be caused by cytoplasmic factors, these can cause transient similarities between sister and mother/daughter cells after division[131]. These similarities occur when daughter cells inherit similar protein states from the cytoplasm of the dividing cell, they last for a few hours after division but fade over time as the cell produces its own cytoplasmic factors[131]. This would cause mother:daughter and sister:sister cells to have less variation between them than cousin:cousin cells and may inflate the estimates of H^2 between closely related cells. This could be further investigated by removing the beginning portion of each cell

track from heritability calculations and hence the removal of any similarities caused by the inheritance of protein states. More in depth and powerful statistical analysis would be required to conduct this type of investigation and whilst lower, estimates of H^2 for cousin:cousin cells are still significant. There are alternative explanations for the variation in heritability estimates and shared variance that can be tested experimentally.

As heritability is a ratio it can decrease for two reasons, if the rate of environmental variation increases or if the rate of genetic variation decreases. The estimated mutation rate of the human genome[152 153] ($1.6 \times 10^{-7} – 3.5 \times 10^{-9}$ base pair mutations per cell division[152 154]), the short time frame of the experiments and the close relationships measured, all suggest that any increase in genetic variation is unlikely to be due to genetic mutations alone. An epigenetic trait is a stably heritable phenotype resulting in changes in a chromosome without alterations in the DNA sequence[47 155]. Epigenetic modifications could occur quickly and would enable a heritably plastic response, meaning the lower H^2 estimates obtained for cousin cells could be due to epigenetic rather than genetic mutations[112 156]. A review by Soo-You (2012) concluded that not only do both genetic and epigenetic mutations play a role in disease progression but they are not mutually exclusive events[157]. Alterations in epigenetic mechanisms can lead to genetic mutations and genetic mutations in epigenetic regulators can lead to an alteration in the epigenome[157 158].

Many studies have looked for a 'metastatic gene' or even a gene expression profile common to metastatic tumours[7]. No-one has yet identified tumour-stage specific gene signatures. A study by Ramaswamy et al (2002) identified a gene expression pattern that appeared to be common in 12 metastatic tumours as well as some primary tumours[33]. The lack of specific metastatic markers suggests that most tumour cells have the genetic potential to be motile but display phenotypic plasticity modulated by epigenetic alterations in response to the microenvironment. In support of this theory dissimilarities in RNA and protein expression have been linked to differences in the metastatic capability of tumour cells[21 155]. There is also evidence that tumours use epigenetic modifications to alter the surrounding tissue to create a more favourable microenvironment[159 160]. Stromal cells removed from a tumour's presence, maintain their phenotypes in *in vitro* culture[7 34]. As tumour cells already use epigenetic modifications to influence and adapt their microenvironment, they could use the same mechanisms to adapt to the microenvironment as well. This theory could be tested experimentally using epigenetic inhibitors. Epigenetic inhibitors would act to decrease the rate of epigenetic modification[161 162]. If true, then the expected results would show a decrease in all

the estimates of H^2 but an elimination of the differences seen between the groups of related cells. If, however increasing environmental variation explained the trend in the results then the differences observed in H^2 estimates between groups of related cells would remain.

Cancer cells are clonal, each daughter cell receives one copy of the mother's genome, hence simple unweighted regression was used to estimate H^2 between related cells, this was not weighted based on relatedness. As discussed in section 3.1.5 genetic variation could arise as a result of mutations or maternal effects, either of which would cause genetic variation to increase with the number of cell divisions. As such it may be plausible to consider weighting the regression between related individuals to take account of this variation. However cousin:cousin cells would have approximately double the genetic variation of mother:daughter and sister:sister cells (two divisions separate these cells as opposed to one division separating mother/daughter and sister cells) and yet the results seen here (table 3.4, figure 3.6) show that H^2 estimates of cousins is not half that of mother:daughter and sister:sister cells. This could suggest that the decrease in H^2 estimates is due to an increase in environmental variation.

These experiments were performed on cells *in vitro*, as stated in section 3.4.3, it is usually assumed that cells in culture will have minimal microenvironmental variation[64]. However, as cells travel further away from each other, the environment could change, increasing environmental variation between cells. Subtle alteration in a cell's microenvironment would act to increase the environmental variation and hence decrease both H^2 and R^2. Whilst any variation in the *in vitro* environment would be less than *in vivo*, the amount of environmental variation and whether this occurs at a cellular level could be investigated experimentally. The tumour microenvironment can vary dramatically, and many cancer cells display phenotypic plasticity[17 163]. In order to understand how distinct cancer cell populations evolve, knowledge is needed of cancer cells reactions and responses to the environment and the role phenotypic plasticity and genetic or epigenetic mutations play in that.

As well as investigating whether epigenetic modifications play a role in the variability in heritability estimates between cell relationships, it would also be possible to investigate whether there is any environmental variation between cells. Examination of media distribution within an *in vitro* culture could determine if the variation in nutrients at a cellular level. High variation in the concentration of nutrients at a cellular level would support the theory that the decrease in both H^2 and R^2 was due to an increase in environmental variation.

Regardless of the mechanisms these results show that motility is a heritable trait and therefore natural selection may be acting within a tumour to select for this trait. By measuring multiple cell:cell relationships the results presented here represent stable differences between clonal cell lines.

Chapter 4 – Investigating the decrease in broad-sense heritability of cell motility between more distantly related cells

4.1 Introduction

In this chapter the lower estimates of heritability for more distantly related cells is investigated. This study has found relatively high and significant values of broad-sense heritability for the cancer cell trait of motility (section 3.3.2). However for all cell lines H^2 estimates from cousin:cousin cells are lower than those obtained for mother:daughter or sister:sister cells. The trend in the results is that measurements of heritability decrease slightly between more distantly related cells. As cancer cells are clonal, H^2 estimates should not vary based on relatedness. As discussed in section 3.4.5 this decrease could occur for two reasons, either an increase in environmental variation or a decrease in genetic contribution to motility[46].

Clonal expansion of cancer cells and the timespan and number of generations measured in these experiments make it probable that any change in the genetic contribution to motility is regulated via epigenetic modifications[164]. This will be investigated using epigenetic inhibitors and their effect on the heritability of cell motility between related cells. The amount of environmental variation is also investigated using fluorescent molecules in culture media. Results are inconclusive as to whether the lower H^2 estimates between cousin cells is due to epigenetic modifications, but results do show spatiotemporal environmental variation at a cellular level.

4.1.1 Epigenetic contributions to evolution by natural selection

The term epigenetics was first coined by Waddington in 1942, originally it referred to how genes interacted with their environment to produce a phenotype[165]. Today, epigenetic effects are changes in gene activity and expression produced by modifications to DNA or chromatin structure without altering the nucleotide sequence of a gene[47 165 166]. These changes can be stably passed to subsequent generations of cells or even organisms[47 167 168].

Many studies have found transgenerational epigenetic inheritance via transmission of molecular factors determining how the DNA is expressed[169], for example Cubas et al (1999) showed flower symmetry in *Linaria vulgaris* was determined via inherited methylation

patterns[170]. Such transgenerational epigenetic inheritance has not only been shown in individual organisms but also between generations of mitotically dividing cells[168]. In sexually reproducing organisms such heritable epigenetic modifications in germline cells can transmit developmentally induced as well as stochastic phenotypes to subsequent generations[168]. In single celled organisms such as bacteria or Archaea transgenerational epigenetic inheritance can lead to the clonal persistence of induced and stochastically generated phenotypic variation[168 171].

In order for epigenetic modifications to contribute to natural selection they would need to affect phenotype (of the trait under selection) [169]. Zhang et al (2018) demonstrated that epigenetic modifications play a role in natural selection using populations of epigenetic recombinant inbred lines of the plant *Arabidopsis thaliana* and comparing their heritable variation in phenology, growth and fitness to both genetic recombinant inbred lines and natural collections[172]. They found similar magnitudes of variance between all lines showing epigenetic modifications can influence phenotypic variance and contribute to adaptation and natural selection[172].

Swift and reversible epigenetic alterations may mediate rapid plastic responses to the environment[173]. Phenotypic plasticity is the capacity to express a different range of phenotypes in response to environmental changes[47]. A plastic response to environmental selective pressures could be mediated through epigenetic mechanisms leading to the inheritance of an environmentally dependant phenotype[169 174]. This is well studied in relation to transgenerational immune priming, where parent's immunological experiences are passed on to enhance offspring immune defence[169]. This mechanism could be particularly important to organisms with limited genetic diversity as it could allow adaptation to environmental perturbation before genetic adaptation occurs[166 173]. The adaptive value of this will depend on the predictability of the environment across generations[169]. Technological advances, particularly the relative speed and ease of gene sequencing, have helped highlight the roles of multiple mechanisms of epigenetic inheritance[169]. Chromatin immunoprecipitation can elucidate the association between specific proteins and DNA regions, when combined with next generation sequencing (known as ChIP-seq) the binding sites of DNA-associated proteins can be narrowed down to base pair resolution[175].

4.1.2 Epigenetics and cancer cells

Cancer cells are clonal meaning changes in the genetic contribution to motility should only arise via mutations. There are many estimates of somatic cell mutation rate, ranging from 2.5 x 10^{-8} mutations per nucleotide site[153] to 1.6 x 10^{-7} – 3.5 x 10^{-9} base pair mutations per cell division[152 154]. The number of generations measured in these experiments make it probable that the majority of any change in the genetic contribution to motility is regulated via epigenetic modifications[164].

Both genetic and epigenetic mutations contribute to cancer progression with widespread alterations to the epigenetic landscape considered one of the hallmarks of cancer[25]. Tumours alter over time as a result of cellular evolution, driven by both genetic and epigenetic mutations, the adaptive value of these mutations is determined by a cells microenvironment[176]. Epigenetic modifications would allow a cell to respond to a changing microenvironment meaning epigenetic modifications could create new heritable variation ultimately leading to adaptation increasing cell survival[166 167].

Environmentally induced epigenetic modifications have been extensively measured in cancer cells. A study by Yang et al (2015) found that heat shock, an environmental stress, increased p53 gene acetylation in HCT116, MCF7 and HT29 cells[177]. Changes in nutrient availability or resource competition could cause cells to alter their motility using epigenetic modifications, as seen in a study by Deng et al (2018) where hypoxia induced transcriptional modifications which increased metastasis in pancreatic cancer cells[176].

Whilst transient, epigenetic modification could be the means by which cells heritably alter their motility[47] any cells that do not remain in the same environment as their relatives could undergo individual epigenetic modifications altering their phenotype to suite their microenvironment. A cell's descendants might undergo individual epigenetic modifications which would act to increase the genetic variation between relatives, reducing heritability between more distantly related cells.

Genetic mutations could also be contributing to the lower H^2 estimates between more distantly related cells; however, the short number of generations measured, and the clonal proliferation of cancer cells mean it is more likely to be epigenetic modifications contributing to differences in heritability[164]. This hypothesis can be tested *in vitro* using epigenetic inhibitors.

4.1.3 The nature of epigenetic modifications

Nucleosomes are the smallest unit of chromatin, as shown in figure 4.1 they are formed from DNA wrapped round an octamer of histone proteins. The histone amino terminal tails protrude and it is these upon which the majority of post translational modifications occur (figure 4.1), including acetylation, methylation, phosphorylation, ubiquitination and sumoylation[179]. The histone code - a complex range of modifications resulting in specific protein binding and gene regulation[47 156], is the combination of these epigenetic modifications and influences chromatin structure, hence gene expression can be influenced by several epigenetic mechanisms[47] (table 4.1).

Table 4.1: Outline of epigenetic modifications, their mechanisms of action and general effect on gene expression.

Name	Mechanism of action	General effect on gene expression
Methylation	Addition of methyl groups (CH_3) to a substrate.	Reduced gene transcription
Acetylation	Addition of an acetyl group (CH_3CO) to a chemical compound.	Increased gene transcription
Phosphorylation	Addition of a phosphoryl group ($P^+O_3^{2-}$) to a substrate.	Reduced gene expression
Ubiquitination	Binding of ubiquitin to protein substrate.	Increased gene expression
Sumoylation	Post-translational covalent attachment of a SUMO protein to a protein substrate.	Alters transcriptional regulation

Protein methylation and acetylation are two of the best studied and characterized of the different types of epigenetic modification[180]. Impaired acetylation has been linked to tumour progression[181] and the methylation rate is generally faster than the rate of genetic mutation making it likely that epigenetic modifications contribute to tumour evolution[29 47 164]. Cancer cells often show abnormal epimutation rates, particularly in methylation leading to loss of activity in tumour suppressor genes[47].

Whilst in general each specific epigenetic modification will be associated with an over-all increase or decrease in gene expression the combination of modifications can result in unexpected and complicated outcomes and result in a range of stably heritable phenotypes[180]. Chromosome condensation can be altered via DNA methylation and histone modification of the N-terminal tail via acetylation, methylation or phosphorylation[47].

As seen in figure 4.1 whilst methylation of both DNA and histones is usually associated with reduced gene expression and histone acetylation with increased gene expression[180] there are still acetyl and methyl groups present on both eu- and hetero- chromatin, it is the overall combination that determines the exact effects on gene expression[182].

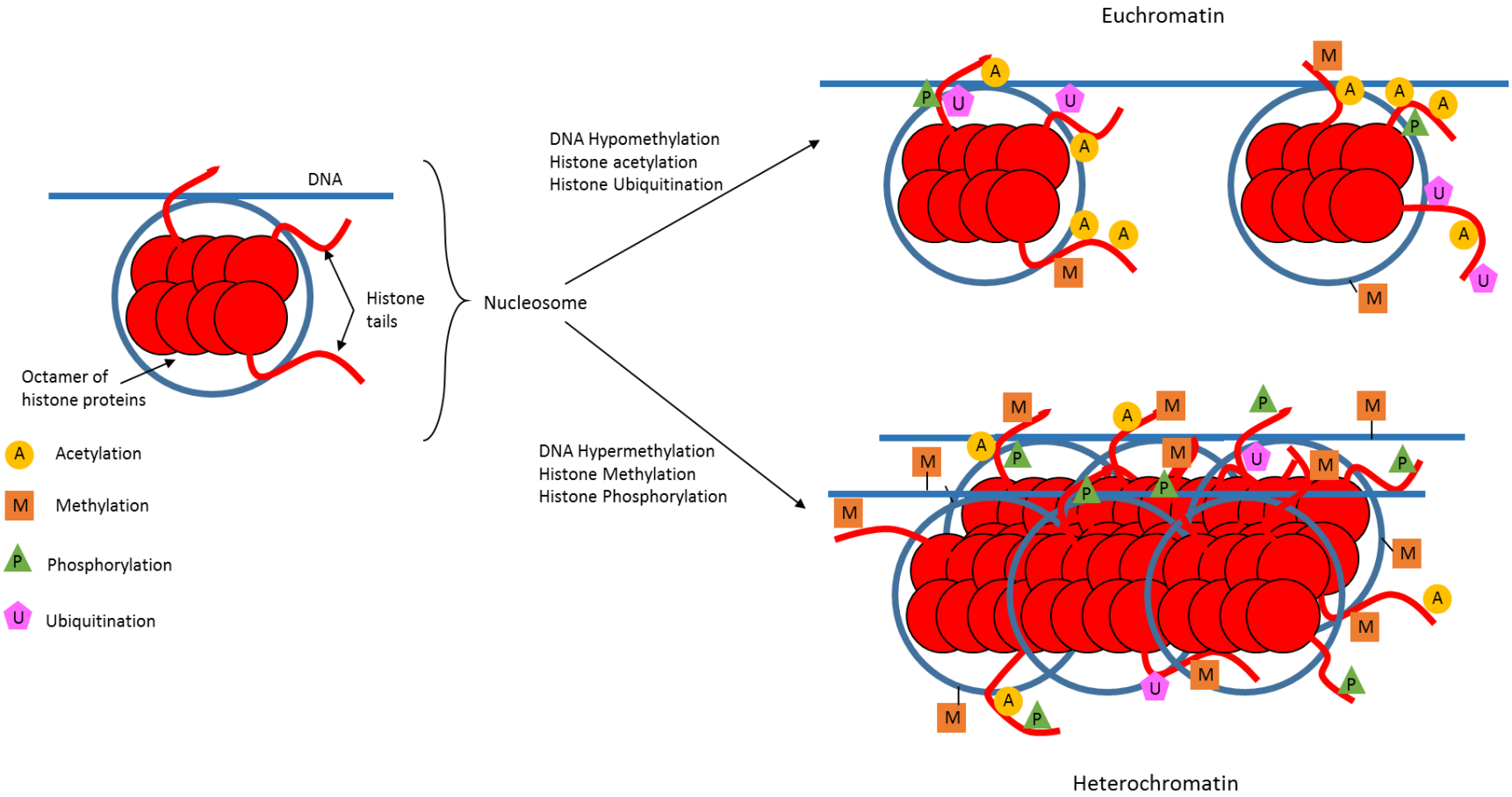


Figure 4.1: Eukaryotic chromatin organisation; modified from Gangisetty et al (2006). Euchromatin is transcriptionally active and characterised with DNA hypomethylation and hyperacetylation of histone N terminal lysine residues. Heterochromatin is transcriptionally inactive and characterised with hyper methylation of DNA and methylation of histone lysine residues[182].

4.1.4 Methylation

Methylation is one form of epigenetic modification that can alter gene expression. It occurs when a methyl group (CH_3) is added to a substrate (either DNA or protein). DNA methylation involves the conversion of cytosine into 5 methyl-cytosine which binds guanine 1.8 times more tightly. This creates heterochromatin (tightly packed regions of DNA) and usually reduces gene expression[29 47].

During cell division daughter cells inherit methylated DNA through enzyme recognition of hemimethylated regions (methylation on only one strand of DNA) [183]. As methylation is not a permanent genetic modification, it may be lost over time through imperfect copying after division[184]. Maintenance of methylated DNA using hemimethylated enzymes, boundary regions and DNA repair enzymes can result in stably heritable epigenetic gene expression[47 181].

DNA methylation patterns in cancer cells are distorted with both hyper- and hypo-methylation of genes shown to be an important mechanism contributing to tumorigenesis[185]. Many intergenic regions of a cancer cell's genome display hypomethylation, thought to contribute to genomic instability, and many tumour suppressor gene promoter regions display hypermethylation leading to loss of activity[185]. This pattern of hypo- and hyper- methylation seem to be pre-programmed with the majority of hypermethylation occurring in CpG Islands marked by polycomb complex proteins (a family of proteins involved in chromatin structure and usually a reduction in gene expression) and hypomethylation being concentrated to Lamin-associated domains (genomic regions in contact with the nuclear lamina, thought to be involved in chromosome organisation) [186 187]. The exact patterns of methylation modifications are most likely linked to the specific developmental history of each cell type just as cancer risk is associated with an individual's lifestyle[186].

Methylation patterns are linked to metastasis, for example circulating tumour cells show hypomethylation of proliferation associated transcription factors and hypermethylation of tumour and metastasis repressor genes[188 189]. The importance of methylation modifications on tumorigenesis has resulted in the epigenome being investigated as a potential therapeutic target as well as its use in cancer diagnosis and biomarkers[185 190 191].

Demethylases are enzymes which remove methyl groups from DNA and histones[192]. Through control of DNA and histone methylation they regulate the chromatin state at specific loci hence altering transcriptional regulation[193]. Demethylases have been investigated as new

targets for their use in drug discovery[194]. In this project GSK J4 has been used to investigate the link between methylation rates and motility.

4.1.4.1 GSK J4 demethylase inhibitor

GSK J4 is an ethyl ester derivative of GSK J1 and a selective inhibitor of the H3K27 histone demethylase JMJD3, as such GSK J4 increases methylation[162]. JMJD3 (also known as KDM6B) belongs to the Jumonji family of demethylase enzymes[192 195]. These demethylases are dependent on 2 co-factors, 2-oxoglutarate (also known as α-ketoglutarate) and Fe2+ for enzymatic activity[194 196]. GSK J4 is a competitive inhibitor of both co-factors but not the substrate itself[197].

The Jumonji demethylase family play a crucial role in regulating epigenetic processes via the removal of methyl groups from histone tails[192]. In non-malignant cell types these demethylases possess tumour-suppressor characteristics as overexpression leads to cell cycle arrest[198]. Jumonji demethylases have been implicated in various human cancers with a wide range of roles, in some cancers they were shown to be mutated, deleted or down regulated whilst in others they have been shown to be upregulated, over expressed or essential for cell survival[195].

JMJD3 accumulation is linked to an increase in the transcription of many genes including BMP-2. BMP-2 binds TGFβ receptors activating SMAD family transcription factors which inhibit cell cycle progression stopping cells from making the G1 to S phase transition[199 200]. Cells in G1 phase of the cell cycle exhibit reduced motility so inhibition of JMJD3 might act to increase cell cycle progression, increasing division rates and population motility[201 202]. JMJD3 also interacts with NF-κB to increase matrix metalloproteinases and growth factor genes[203]. Matrix metalloproteinases are known to play a major role in cell migration as well as proliferation, angiogenesis and apoptosis[204]. Inhibition of JMJD3 may therefore also be expected to reduce cell migration[205]. Cell signalling pathways are complex, ubiquitination and phosphorylation also play a role in SMAD activation and action[200] and metalloproteinases are regulated at multiple levels including mRNA modulation and protein activation[205]. This makes it difficult to predict the effects of JMJD3 inhibition on population motility. Despite the unknown effects on cell phenotype, addition of GSK J4 should reduce demethylase activity. The reduction in novel epigenetic modifications is expected to cause an overall decrease in heritability estimates and eliminate the differences seen between H^2 estimates of cousin-cousin cells and mother-daughter & sister-sister cells.

4.1.5 Acetylation

Acetylation is another epigenetic modification that regulates gene expression it occurs when an acetyl group (CH_3CO) is added to a chemical compound[164]. Acetylation of histone N-termini removes the positive charge reducing its interaction with the negatively charged DNA phosphate group[47 156]. This results in euchromatin (lightly packed regions of DNA that are the transcriptionally active form of chromatin) and generally increases gene expression[164].

Histone protein acetylation is catalysed by histone acetyltransferase (HAT) or histonedeactylase (HDAC) activity[196]. HDACs are a class of enzymes that remove acetyl groups from the lysine residue on the N-terminal of a histone[184 206]. This allows the histones to wrap the DNA more tightly and reduces gene transcription[206].

Acetylation plays an important role in translational protein modification, in particular the lysine residues in the N-terminal of histone proteins. In addition to histones, other proteins such as p53, E2F, a-tubulin and MyoD have also been shown to be deacetylated by HDACs. Many non-histone proteins have been identified that are substrates for at least one type of HDAC illustrating the complex function of HDACs in many cell processes[206].

HDAC expression between normal and malignant cell types has not been well characterised but acetylation rates have been shown to interact with many oncogenes and tumour suppressor genes[181 206]. Dysregulation of HAT enzymes leads to aberrant HDAC activity causing alterations in gene expression and other processes such as cell growth, differentiation and apoptosis which have all been shown to contribute to tumour evolution[181 206].

Inhibition of HDACs are being investigated for their use in cancer treatment. HDAC inhibition should act to increase gene expression however the complexities of cell signalling and the aberrant nature of cancer mean inhibition of specific enzymes may not result in a change of phenotype[207 208].

4.1.5.1 MS-275 HDAC inhibitor

MS-275 (also known as Entinostat) is a benzamide compound that acts to increase acetylation[209] and has been shown to have anti-tumour activity *in vivo*[184]. It is a selective inhibitor of HDAC1 and HDAC3[184]. These enzymes are considered important for tumour cell cycling, are mostly localised in the nucleus and play a role in cell survival and proliferation[206 210]. HDAC1 and HDAC3 regulate histone deacetylation through multi subunit complexes.

Various proteins form a complex with HDAC enzymes to initiate repression of acetylation and DNA transcription[206]. Inhibition of HDAC1 and HDAC3 causes cell cycle arrest and inhibits cellular proliferation hence stopping tumour growth[207 211].

HDACs are classified according to their sequence homology, HDAC1 and HDAC3 are structurally similar subfamily members of the class 1 enzymes and have zinc dependent active sites[196]. MS-275 inhibits HDACs 1 and 3 by binding to their zinc containing catalytic domain resulting in hyper acetylation of nuclear histones and generally increasing gene expression[209]. It has also been shown to induce accumulation of p21 and gelsolin, increases of which have been linked to a decrease in cell motility, and to increase apoptosis[201 202 212 213].

Many cancer cells have aberrant signalling pathways[214] and structural alterations[213] making it difficult to predict the exact effects that an inhibitor such as MS-275 will have. Addition of an epigenetic inhibitor such as MS-275 should reduce the amount of epigenetic modifications occurring, if the lower estimates of H^2 between more cousin cells is due to acetylation causing a decrease in the genetic contribution to motility then the overall heritability of motility would be expected to decrease but the differences between more distantly related cells would reduce.

4.1.6 Environmental variation

Lower H^2 estimates of cousin cells could be due to a decrease in the genetic contribution to motility and/or an increase in environmental variation[46]. In general *in vitro* culture is considered homogenous, especially when compared to an *in vivo* environment, however over time cells will use resources and produce waste products which accumulate in the culture media and could cause fluctuations in the microenvironment[64]. As cell division occurs the population of cells increases resulting in more cells being present in the culture, this could lead to an increase in the amount of environmental variation over time. Cells which migrate away from their relatives may experience differences in nutrient concentration which may also act to increase the environmental variation between more distantly related cells. Very little has been done to investigate the amount of heterogeneity within a cell culture as when conducting *in vitro* experiments due to the repeatability of experimental conditions and the simplicity in comparison to *in vivo* environments the environmental variation within an experiment is usually considered negligible[64].

Environmental heterogeneity is an important factor in heritability. If the environmental variation is high this may cause an increase in phenotypic variation decreasing H^2 estimates[130]. Likewise, if environmental variation is low the variation in phenotype may decrease, increasing estimates of H^2 hence environmental variation may affect H^2 estimates irrespective of any changes in genetic variation[130]. A tumour microenvironment can vary dramatically over both space and time and many cancer cells display phenotypic plasticity[17 163]. Phenotypic plasticity in cancer cell traits means that environmental variation will have an important effect on heritability measurements[215]. Understanding the types of environmental variation is crucial to interpreting results[128 129].

General environmental variation is all the non-genetic sources of variation between individuals experienced by multiple individuals within a population[128]. *In vitro* experiments have carefully controlled and repeatable environments, there may be differences at a cellular level but there should be no difference in the environment within a family of cells greater than that of the population as a whole[130]. Specific environmental variation is the deviation from the population mean due to the microenvironment experienced by an individual. If high, then this can weaken the correlation between genotype and phenotype[128].

When culturing cells *in vitro* it is often assumed that a uniform environment is maintained at a cellular level throughout the culture. Some environmental variation between cells will stem from other surrounding cells but little is known about the amount of variation within the media of the culture. Culture medium cannot be completely homogenous throughout cell culture; however, it is not known how much heterogeneity there is within a flask over short periods of time and whether culture media can be considered homogenous or heterogeneous at a cellular level. Variation in an *in vitro* culture would always be less than *in vivo* however whether or not there is variation *in vitro* at a cellular level could be tested experimentally by examining the distribution of fluorescent particles within culture media.

Fluorescent dyes have long been used to enable quantification of cellular processes[216]. Fluorescence is the absorption and emission of light by a substance and is a three-stage process; a photon with energy is absorbed by a fluorescent molecule (known as a fluorophore), creating an excited electron state. The fluorophore undergoes a conformational change dissipating some of its energy. The third step is emission of a photon with the remaining energy, returning the fluorophore to its ground state. The lower energy of the

emission photon means it has a longer wavelength and allows detection and quantification of the fluorescent signal[217].

Fluorescein is a manufactured organic compound and dye and isothiocyanate is the chemical group, –N=C=S. FITC is a derivative of fluorescein containing an isothiocyanate group and has excitation and emission spectrum peak wavelengths of approximately 495 nm/519 nm[217]. Dextran is a complex branched glucan which coupled with a fluorescin isothiocyanate (FITC) is often used to image and characterize concentration gradients[218-220].

Addition of FITC-Dextran to the media would allow images to be collated displaying the concentration and distribution of particles within *in vitro* culture media. Whilst addition of fluorescent particles would not model specific nutrients within the media, it could be a fundamental first step to quantifying the amount of environmental variation in *in vitro* cell culture at a cellular level. Any variation may not be on the same scale as that in an *in vivo* environment but might indicate whether in this project an increase in environmental variation could be contributing to the reduction in heritability.

4.1.7 Aims

In the first set of experiments cells will be exposed to GSK J4 and MS-275 to inhibit HDAC and demethylase enzymes. The motility and heritability of these cells can then be measured and compared to a control group. If epigenetic modifications are responsible for a decrease in heritability between more distantly related cells, then inhibition should show an overall reduction in the heritability measured but eliminate the difference seen between more distantly related cells.

The second set of experiments involve the addition of fluorescent particles to culture media and measurement of the distribution to enable quantification at the cellular level of particles suspended in culture media. Whilst imaging fluorescent molecules in culture media will not determine the specific distribution of all supplements it can help establish the uniformity of *in vitro* cultures at the cellular level[221] and whether or not environmental variation is increasing between more distantly related cells.

4.2 Methods

4.2.1 Cell culture

Four lab adapted cell lines all obtained from Public Health England were used for all experiments; HeLa[73], MCF7 (MCF7-3)[69], MDA-MB-231 (MDA-2)[68] and HT1080[77]. All experiments were repeated three times with passage numbers between 40-47 for MDA-2, 14-19 for MCF7-3, 4-9 for HT1080 and 4-8 for HeLa cells. Cells were cultured as a monolayer in 5 % CO_2 at 37 °c in 25 cm^2 culture flasks (T-25 life technologies) containing 7 mL of media.

4.2.1.2 Epigenetic inhibitor media

Stock solutions of 10 µM drug in DMSO were made from 10 mg MS-275 and GSK J4 (Selleckchem). Solutions were filter sterilised using 0.2 µm filters, then aliquoted, labelled and stored at -80 °c. Stock solutions were defrosted 30 minutes prior to use in a 37 °c waterbath.

Cells cultures were exposed to a range of drug concentrations. Concentrations of MS-275 were 10, 30, 50 and 100 nM. Concentrations of GSK J4 were 10, 30, 60 and 100 nM. To achieve the required drug concentration 10 µM stock solutions were added to culture media. The volumes needed were calculated using the formula below where C = Concentration, V = Volume and the numbers denote the two different solutions. Drug media was vortexed before being added to cells ensuring thorough drug dispersal.

$$C_1V_1 = C_2V_2$$

To control for the presence of DMSO in the drug solutions, controls were used containing the same volume of DMSO as in the treatment groups but without any drug. This was used in every experiment along with a control of standard media (not containing drugs or DMSO). Differences were seen between the control group and the DMSO control hence the DMSO control has been used for comparison of drug effects on cell traits.

4.2.1.3 FITC-Dextran media

A 50 mg/mL stock solution of fluorescein isothiocyanate–dextran (FITC-Dextran) (Sigma) average molecular weight of 70, 000, was made in PBS and then filter sterilised using a 0.2 µm filter. Stock solution was made approximately 10 minutes before the start of each experiment

and was not stored between repeats. Each stock solution was made in a centrifuge tube wrapped in tin foil and shielded from the light.

FITC-Dextran solution was added to culture media to give concentrations of 1 mg/mL, 1 µg/mL and 1 ng/mL. Solutions were vortexed to ensure thorough mixing and dispersal of fluorescent particles.

4.2.3 Fluorescently staining cells

1 L stock solution of 4 % Paraformaldehyde was made in a fume hood by mixing 40 g paraformaldehyde with 800 mL PBS. The solution was gently heated to 60 °c then placed on an automatic electronic stirrer and drops of 5 M NaOH slowly added until all the paraformaldehyde had dissolved. The solution was left to cool to room temperature and made up to 1 L with PBS. Drops off HCl were slowly added until pH 7.4 was reached, this was tested using an electronic pH meter. The stock solution was made into 50 mL aliquots, clearly labelled and stored at -20 °c. Each aliquot of stock solution was defrosted in the fridge before use.

A 0.1 % Triton solution was made using Triton X-100 (sigma) and PBS. The required volume of solution was made by pipetting Triton X-100 into PBS before gently heating and stirring the solution to allow dissolution.

To make a goat serum solution the required volume of goat serum (lifetechnologies) was added to 0.1 % Triton solution i.e 10 µl goat serum and 90 µl 0.1 % Triton to create 1 mL total volume 10 % goat serum solution.

HT1080 cells were plated at 1 x 10^4 cells per dish onto nine 35 mm glass dishes (IBIDI). 0.5 mL standard culture media was added and they were left for 16 hours in an incubator in 5 % CO_2 at 37 °c.

Media was removed from the cells and 1 mL of 4 % paraformaldehyde added to each glass dish. These were left for 12 minutes before removing and adding 1 mL 0.1 % Triton for 5 minutes. This was removed and 1 mL per well of 10 % goat serum solution was added. The dishes were then incubated at 4 °c for 16 hours.

Each dish was then washed three times for five minutes in 1 mL PBS. 0.5 mL of Phalloidin546 (lifetechnologies) at a 1:1000 dilution in 2 % goat serum was then added to each dish, this selectively stains F-actin present in cells. The dishes were covered and left for 1 hour at room

temperature. Three more five-minute PBS washes followed then one drop of DAPI vectashield (VectorLabs) was added to each dish, DAPI binds A-T rich regions of DNA and vectashield acts as an antifade mounting medium. The fluorescently stained HT1080 cells were taken straight to the A1R confocal microscope.

4.2.4 Fluorescence microscopy

0.5 mL of FITC-Dextran media (at concentrations of 1 mg, 1 µg and 1 ng /mL) was pipetted onto eighteen ibidi glass dishes (six dishes per concentration). Half of the dishes contained fluorescently stained HT1080 cells, the other half contained no cells at all. Due to the lack of an efficient environmental chamber no live cells could be used.

A Nikon A1R confocal microscope with NIS elements photograph software was used to take pictures of five randomly chosen points within each dish. All images were taken using a x 4 objective and included all fluorescent channels. All images were processed using ImageJ[97-99].

Two fluorescence time-lapse videos of FITC-Dextran media at 1 µg/ml, each 1 hour in duration, were also taken using the A1R confocal microscope. The time-lapse was not multi-point and images were taken every 5 minutes for 1 hour using a x 4 objective. The videos were unsuccessful due to technical issues with the microscope and problems with photo bleaching degrading the fluorescent signal over time.

4.2.5 Time-lapse microscopy

Details outlining the overall technique can be found in section 2.7.1. Cells were plated at 5000 cells per well onto a 24-well plate in 1 mL media. Wells were divided equally between each cell line and the number of treatment groups. One hour prior to the experiment cell media was changed. All previous media was removed and the stated concentration of drug, DMSO or control media added.

A Nikon TiE time-lapse system microscope and NIS elements photograph software captured three points within each well. Images were then taken at a x4 objective of these points every 20 minutes for 72 hours.

The time-lapse experiment was repeated three times for both GSK J4 and MS-275. NIS elements photograph software captured a total of 93,312 images which were then converted into a video file for each point[101]. ImageJ and MtrackJ[97-99] were used to analyse these videos

and track individual cells. The total number of cells tracked in each experiment was; 14,029 cells in GSK J4 media and 22,464 cells in MS-275 media.

4.2.6 Cell tracking

In total 36,493 cells were chosen at random from 432 videos each containing 216 frames giving a total of 2,041,597 individual points. ImageJ and MtrackJ[97-99] were used to follow the movement of the central point in each cell between successive frames[101] (as outlined in section 2.9).

At the start of the experiment all cells within the field of view were regarded as the first generation, with every new cell division and the subsequent daughter cells regarded as a new generation in that family. Cells were tracked until the end of the experiment or until they moved off screen, died or divided, known as their fate. A manual record was kept of each cell's fate, which generation it belonged to and how it was related to other cell tracks. This manual record was kept throughout the analysis and allowed comparison between groups of cells to ensure there was no bias in motility within the sample of cells measured (see section 2.10.5).

From the data gathered three different cell:cell relationships could be used to calculate estimates of H^2 (shown in figure 3.2, section 3.2.3).

4.2.7 Statistical analysis

For a full explanation of the techniques chosen see section 2.10. All statistical analysis of cell tracking was performed using Minitab[103] and IBM SPSS[102]. Cell track data was extracted from MtrackJ[101] and stored in excel files, manual data was then inputted into these excel files throughout data analysis and collection.

Cell motility was calculated as a cell's average speed in microns per hour. Cell speed was measured as the total distance of the path travelled by the cell (in microns), divided by its lifetime (in hours).

The Lilliefors correction of the Kolmogorov-Smirnov test was used in conjunction with descriptive statistics to determine the distribution of the data[104]. To avoid fundamentally altering the original data, which could lead to a misinterpretation of results, non-normal results have not been mathematically transformed[106 107]. Non-parametric tests were used to

check for differences in cell motility between drug concentrations, this was done separately for each cell line.

Differences in motility between cell fates, generations, experimental and mathematical repeats were tested for using Kruskal-Wallis and Mann-Whitney U tests[104]. This was done to ensure there was no bias between the samples of cells measured.

The broad-sense heritability was estimated using the slope parameter of an ordinary least squares regression of trait values between related cells for the three different cell:cell relationships[106]. When a cell had multiple daughter cells or cousin cells their mean value was used and H^2 estimated using simple unweighted regression of trait values between related cells[106]. Both H^2 and R^2 values have been given in the results. R^2 values show how much of the variance in the dependent variable (cell speed) is explained by the model[104].

4.2.8 Processing fluorescent images

Images of FITC-Dextran media were opened in ImageJ[97 98] with only the green colour channel. The images were processed individually, and the X and Y values transformed and data binned (the combining of a cluster of pixels into a single pixel) to 4x4, this was done to increase the signal level. Images were then heatmapped, using the Heatmap Histogram plugin[98], to clearly display the FITC-Dextran distribution.

Images containing the fluorescently labelled HT1080 cells in FITC-Dextran media were opened as composite colour files in ImageJ[97 98]. Composite colour superimposes all three colour channels into the same field of view. If needed images could be opened using the default colour mode allowing visualisation of each colour channel separately. The three colour channels were the blue; DAPI stained nuclei, red; phailloidin546 (bound to cellular F-actin) and green; FITC-Dextran.

4.3 Results

4.3.1 Effect of MS-275 on HeLa, HT1080, MCF7-3 and MDA-2 cells

4.3.1.1 Motility of cell lines exposed to MS-275

As seen in figure 4.2 there was no significant difference in speed between the different experimental repeats at any concentration of MS-275 for HeLa, HT1080, MCF7-3 and MDA-2 cell lines (Kruskal-Wallis test; $p>0.05$). As a result, all the data for each cell line have been grouped together for further analysis.

MS-275 is HDAC inhibitor, it acts to increase acetylation[184 209]. Five concentrations (chosen based on the literature[184 207 209 210 211]) were tested to determine if there was a dose response on cell motility or if there were any detrimental consequences that might preclude its use in this project. As seen in figure 4.3 and table 4.2, there is a significant difference in motility within each cell line between MS-275 concentrations (Kruskal-Wallis tests $p<0.001$). However, these differences in motility between groups of cells are very small, the range in the median values of cell motility are; HeLa, 1.21 µm/hour; HT1080, 10.03 µm/hour; MCF7-3, 3 µm/hour; MDA-2, 5.04 µm/hour. There is no correlation for higher MS-275 concentrations to result in lower cell speed.

Apart from the 100 nM group of MCF7-3 cells, all concentrations of MS-275 in all cell lines had a non-normal distribution of cell speeds (Lilliefors corrected Kolmogorov-Smirnov test $p<0.003$). As MS-275 does have an effect on motility but there is no correlation between the concentration and the size of the effect the decision was made to use all concentrations of MS-275 when measuring heritability.

Table 4.2: Median speed (µm/hour) and N of cell lines for different MS-275 concentrations (nM)

MS-275 Concentration (nM)	**Median Speed [N]**			
	HeLa	**HT1080**	**MCF7-3**	**MDA-2**
0	11.18 [1052]	36.91 [1728]	30.66 [828]	26.45 [663]
10	10.46 [1260]	31.37 [1371]	29.51 [883]	28.53 [705]
30	9.97 [610]	34.98 [1047]	32.51 [885]	26.39 [622]
50	10.27 [782]	34.08 [1062]	31.73 [758]	23.49 [429]
100	10.2 [1014]	26.88 [961]	30.58 [766]	28.49 [526]

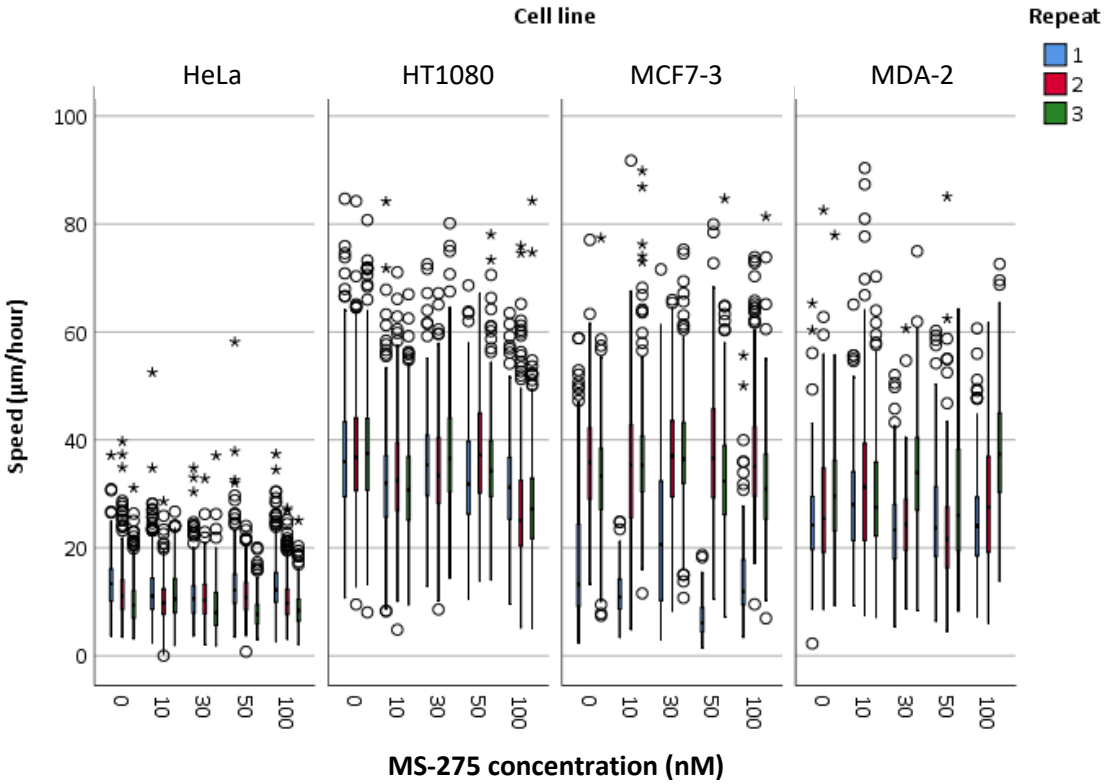


*Figure 4.2: Distribution of HeLa, HT1080, MCF7-3 and MDA-2 cell speeds at MS-275 concentrations (nM) for each experimental repeat. The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. Data gathered from three independent experiments. N for each cell line, MS-275 concentration and repeat is; HeLa 0 nM repeat 1: 335, repeat 2: 428, repeat 3: 289, HeLa 10 nM repeat 1: 394, repeat 2: 344, repeat 3: 522, HeLa 30 nM repeat 1: 317, repeat 2: 122, repeat 3: 171, HeLa 50 nM repeat 1: 304, repeat 2: 269, repeat 3: 209, HeLa 100 nM repeat 1: 357, repeat 2: 384, repeat 3: 273; HT1080 0 nM repeat 1: 388, repeat 2: 477, repeat 3: 863, HT1080 10 nM repeat 1: 432, repeat 2: 310, repeat 3: 629, HT1080 30 nM repeat 1: 285, repeat 2: 392, repeat 3: 370, HT1080 50 nM repeat 1: 332, repeat 2: 292, repeat 3: 438, HT1080 100 nM repeat 1: 145, repeat 2: 463, repeat 3: 353, MCF7-3 0 nM repeat 1: 272, repeat 2: 267, repeat 3: 289, MCF7-3 10 nM repeat 1: 266, repeat 2: 342, repeat 3: 275, MCF7-3 30 nM repeat 1: 393, repeat 2: 176, repeat 3: 316, MCF7-3 50 nM repeat 1: 128, repeat 2: 317, repeat 3: 313, MCF7-3 100 nM repeat 1: 158, repeat 2: 375, repeat 3: 233, MDA-2 0 nM repeat 1: 204, repeat 2: 220, repeat 3: 239, MDA-2 10 nM repeat 1: 270, repeat 2: 205, repeat 3: 230, MDA-2 30 nM repeat 1: 261, repeat 2: 138, repeat 3: 223, MDA-2 50 nM repeat 1: 215, repeat 2: 148, repeat 3: 108, MDA-2 100 nM repeat 1: 221, repeat 2: 148, repeat 3: 157.*

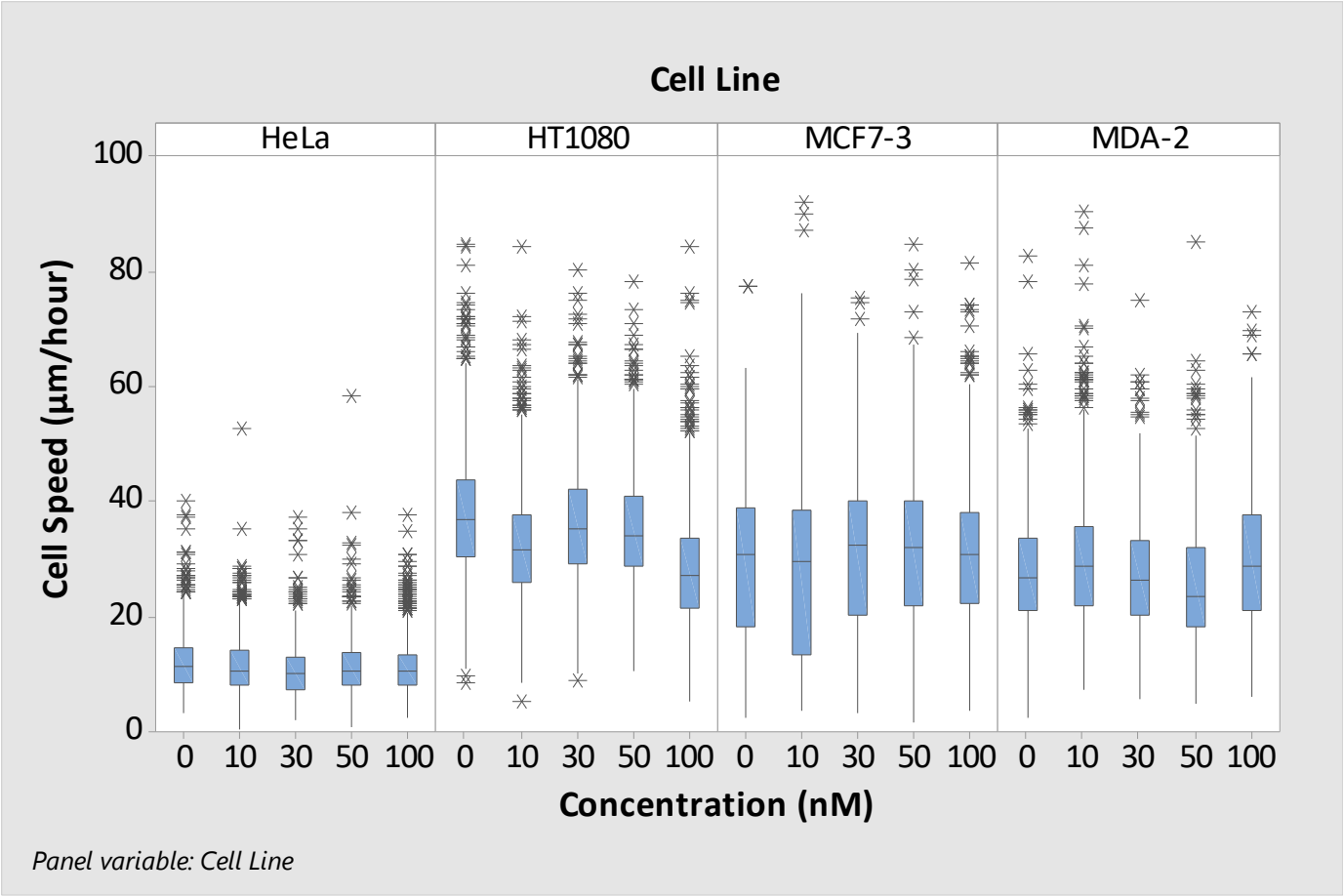


*Figure 4.3: Distribution of cell speeds for HeLa, HT1080, MCF7-3 and MDA-2 cell lines for MS-275 concentrations (nM). The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. Data gathered from three independent experiments. The N for each cell line is; HeLa 4,870, HT1080 6,434, MCF7-3 4,236 and MDA-2 3,047.*

HeLa cells in 100 nM and MCF7-3 cells in 50 nM & 100 nM MS-275 had significant differences in cell speed between generations (Kruskal-Wallis tests $p<0.001$). As seen in table 4.3, for both concentrations of MCF7-3 the 1st generation had significantly lower speeds and for HeLa cells the 1st generation had significantly higher speeds (Mann-Whitney U tests $p<0.001$).

There were also significant differences in motility between cells that had different cell fates (Kruskal-Wallis tests $p<0.003$). Upon further investigation it was found that the differences in motility between cell fates could account for the differences in motility between generations. For MCF7-3 cells non-dividing cells were significantly slower than cells that moved off screen or died (Mann-Whitney U test $p<0.001$), a higher proportion of this cell fate belonged to generation 1 which could explain the overall differences in cell motility. In HeLa cells those that moved off screen (of which there were a higher proportion in the 1st generation) had a significantly higher motility than other cell fates (Mann-Whitney U tests $p<0.001$). By excluding the first generation of cells and only using relationships between their clonal descendants to calculate the broad-sense heritability, the effect on estimates of H^2 caused by variation between generations and cell fates will be minimised.

Table 4.3: Median speed (µm/hour) and N for each of the first three generations in HeLa, HT1080, MCF7-3 and MDA-2 cell lines for five concentrations of MS-275; 0, 10, 30, 50 and 100 nM.

MS-275 Concentration (nM)	Cell line	HeLa			HT1080			MCF7-3			MDA-2		
	Generation	1	2	3	1	2	3	1	2	3	1	2	3
0		14.82 [78]	13.14 [152]	12.33 [259]	44.72 [77]	41.13 [138]	40.06 [213]	39.44 [105]	27.76 [151]	29.37 [180]	30.95 [95]	28.75 [148]	28.34 [216]
10		13.06 [85]	12.99 [169]	11.54 [314]	35.66 [70]	33.24 [121]	32.78 [181]	24.37 [103]	25.60 [166]	25.67 [218]	34.62 [95]	32.46 [147]	29.18 [214]
30		13.03 [51]	12.34 [108]	10.54 [177]	39.23 [100]	38.99 [161]	37.20 [244]	28.97 [88]	28.09 [138]	29.89 [208]	29.00 [95]	27.79 [137]	26.86 [204]
50		13.57 [86]	12.41 [166]	10.83 [286]	38.99 [82]	37.57 [138]	36.42 [199]	17.61 [151]	31.61 [94]	37.01 [118]	29.70 [94]	25.89 [116]	25.03 [142]
100		14.25 [83]	12.54 [167]	10.97 [296]	34.57 [64]	34.65 [106]	29.63 [148]	25.57 [116]	27.42 [144]	31.01 [154]	32.25 [96]	30.13 [125]	28.59 [156]

Table 4.4: Median speed (µm/hour) and N for each cell fate in HeLa, HT1080, MCF7-3 and MDA-2 cell lines for five concentrations of MS-275; 0, 10, 30, 50 and 100 nM.

Concentration of MS-275 (nM)	HeLa				HT1080				MCF7-3				MDA-2			
	Non-divided	Divided	Died	Moved off screen	Non-divided	Divided	Died	Moved off screen	Non-divided	Divided	Died	Moved off screen	Non-divided	Divided	Died	Moved off screen
0	10.7 [538]	11.7 [498]	12.5 [49]		34.2 [756]	38.8 [875]	36.1 [14]	38.2 [180]	28.0 [345]	32.2 [369]	32.1 [22]	34.0 [111]	24.4 [317]	28.2 [294]	37.1 [9]	29.2 [65]
10	10.1 [651]	10.9 [608]	10.4 [50]		31.2 [555]	31.7 [687]	30.5 [31]	31.2 [174]	21.0 [377]	30.3 [399]	33.0 [17]	36.8 [113]	25.8 [352]	30.0 [320]	29.5 [14]	30.6 [51]
30	9.3 [298]	10.5 [282]	10.5 [44]		33.0 [368]	35.8 [482]	35.4 [82]	38.7 [135]	30.7 [326]	32.6 [414]	33.1 [23]	36.5 [157]	24.6 [297]	26.6 [274]	35.2 [6]	32.8 [67]
50	9.7 [369]	11.0 [358]	10.5 [79]		33.4 [438]	34.8 [504]	28.5 [27]	35.4 [121]	11.6 [214]	33.5 [311]	14.3 [15]	36.2 [236]	23.8 [225]	22.3 [172]	26.4 [7]	24.4 [39]
100	9.6 [507]	10.6 [475]	12.2 [57]		25.4 [417]	27.6 [470]	26.0 [28]	31.4 [90]	26.3 [293]	32.3 [335]	27.2 [24]	33.9 [135]	26.7 [232]	29.2 [218]	28.4 [16]	33.2 [71]

4.3.1.2 Heritability in cell lines exposed to MS-275

Cell speed is significantly heritable for all cell:cell relationships in all cell lines and at all concentrations of MS-275. Table 4.5 shows H^2 values range from 0.888 - 0.188 and R^2 values range from 0.78 - 0.04. As stated in section 4.2.7.3 outliers and influential data points have been removed from this analysis. H^2 was estimated using simple unweighted regression of cell speeds between related cells, figure 4.4 shows how H^2 has been calculated between sister MCF7-3 cells exposed to each concentration of MS-275.

Due to the number of post-hoc comparisons it was not possible to determine significance either between or within MS-275 concentrations for estimates of H^2 of each cell line.

As seen in table 4.5 and figure 4.5 in all other groups apart from MDA-2 cells in 30 nM MS-275 the H^2 estimates for the cousin:cousin relationship are lower than those obtained for the sister:sister and mother:daughter relationships. In the 30 nM group of MDA-2 cells the H^2 of cousin:cousin cells is slightly higher than that of mother:daughter. There is no general trend of heritability estimates correlated to the MS-275 concentrations.

R^2 is the amount of variance in cell motility explained by the concentration of MS-275, table 4.5 shows this is less between cousin:cousin cells than sister:sister or mother:daughter cells for all groups apart from MDA-2 cells in 30 nM and HeLa cells in 0 nM.

Inclusion of multiple different cell:cell relationships and examination over several generations means the significant results seen here represent stable differences between clonal cell lineages.

Table 4.5: Mean trait value, R^2 and broad-sense heritability of motility in HeLa, HT1080, MCF7-3 and MDA-2 cell lines for MS-275 concentrations (nM) for three cell:cell relationships.

Cell line	MS-275 Concentration (nM)	No of families	Mean speed (µm/hour)	Heritability of cell motility H^2 (R^2)		
				Sister:sister	Mother:daughter	Cousin:Cousin
HeLa	0	119	11.94	0.538*** (0.29)	0.620*** (0.38)	0.534*** (0.29)
	10	151	11.31	0.618*** (0.38)	0.601*** (0.36)	0.481*** (0.23)
	30	87	10.48	0.661*** (0.44)	0.548*** (0.30)	0.501*** (0.25)
	50	138	11.08	0.658*** (0.43)	0.594*** (0.35)	0.447*** (0.20)
	100	142	10.98	0.464*** (0.22)	0.611*** (0.37)	0.397*** (0.16)
HT1080	0	105	37.52	0.571*** (0.33)	0.483*** (0.23)	0.360*** (0.13)
	10	90	32.1	0.570*** (0.33)	0.454*** (0.21)	0.288*** (0.08)
	30	122	36.14	0.485*** (0.24)	0.566*** (0.32)	0.469*** (0.22)
	50	98	35.09	0.615*** (0.38)	0.428*** (0.18)	0.408*** (0.17)
	100	74	28.3	0.631*** (0.4)	0.506*** (0.26)	0.188* (0.04)
MCF7-3	0	90	29.42	0.863*** (0.75)	0.813*** (0.66)	0.740*** (0.55)
	10	102	28.26	0.858*** (0.74)	0.839*** (0.70)	0.807*** (0.65)
	30	102	30.7	0.863*** (0.75)	0.888*** (0.79)	0.832*** (0.69)
	50	59	30.87	0.731*** (0.54)	0.613*** (0.38)	0.596*** (0.36)
	100	77	30.5	0.726*** (0.53)	0.752*** (0.57)	0.696*** (0.48)
MDA-2	0	104	27.84	0.630*** (0.4)	0.688*** (0.47)	0.527*** (0.28)
	10	105	30.09	0.658*** (0.43)	0.552*** (0.31)	0.389*** (0.15)
	30	102	27.56	0.686*** (0.47)	0.530*** (0.28)	0.567*** (0.32)
	50	68	25.99	0.596*** (0.36)	0.646*** (0.42)	0.417*** (0.17)
	100	75	30.22	0.533*** (0.28)	0.679*** (0.46)	0.427*** (0.18)

*p<0.05 ** p<0.01 *** p<0.001

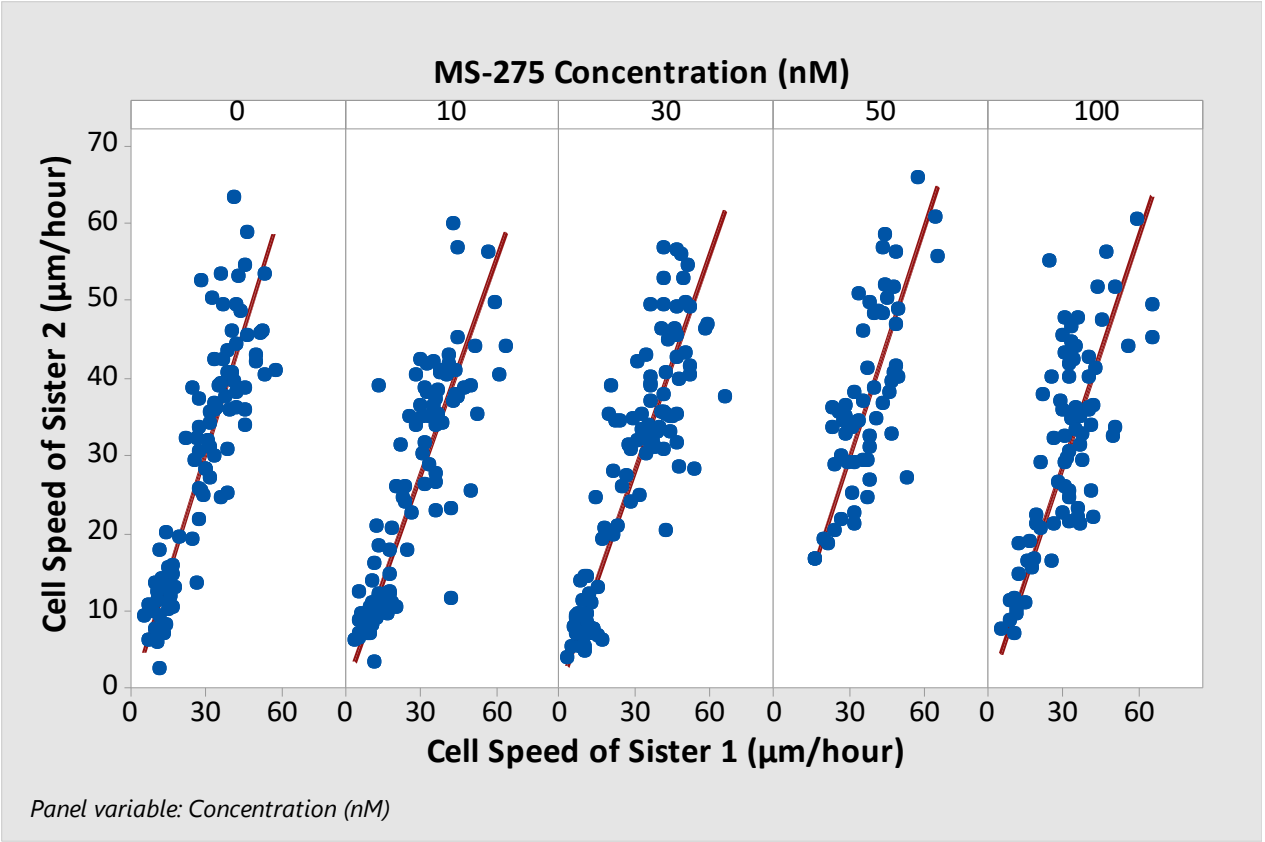


Figure 4.4: Example of the sister:sister regression for motility of MCF7-3 cells for MS-275 concentrations (nM). Each point represents the speed of one pair of sister cells, the slope of the regression line is the estimate of broad-sense heritability. As the cells are clonal, the same method is applicable to other cell:cell relationships. N for each group is 0 nM = 90, 10 nM = 102, 30 nM = 102, 50 nM = 59 and 100 nM = 77.

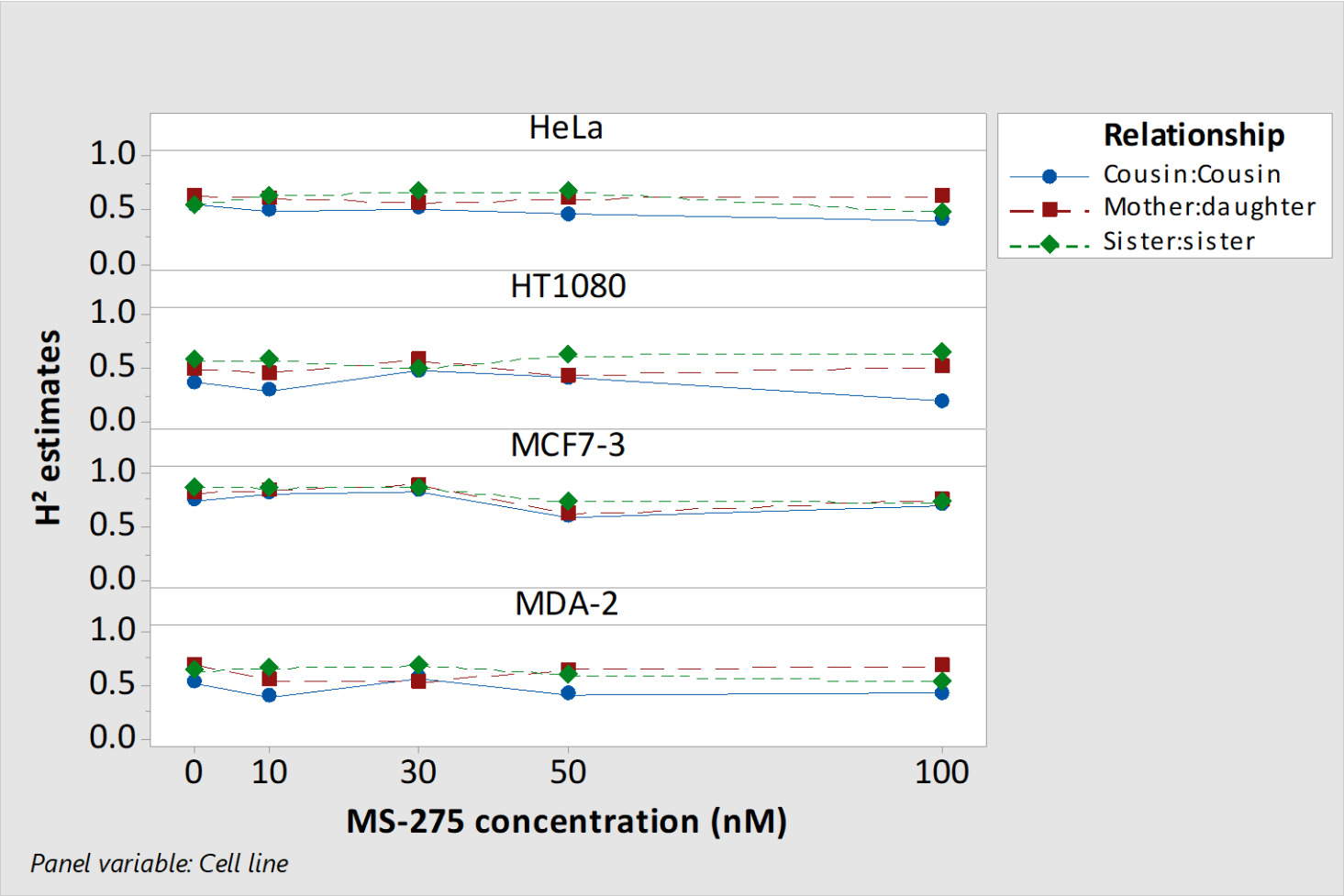


Figure 4.5: Heritability estimates of cell:cell relationships for HeLa, HT1080, MCF7-3 and MDA-2 cell lines at MS-275 concentrations (nM). N for each group is; HeLa; 0 nM = 119, 10 nM = 151, 30 nM = 87, 50 nM = 138 and 100 nM = 142. HT1080; 0 nM = 105, 10 nM = 90, 30 nM = 122, 50 nM = 98 and 100 nM = 74. MCF7-3; 0 nM = 90, 10 nM = 102, 30 nM = 102, 50 nM = 59 and 100 nM = 77. MDA-2; 0 nM = 104, 10 nM = 105, 30 nM = 102, 50 nM = 68 and 100 nM =75.

4.3.2 Effect of GSK J4 on HeLa, HT1080, MCF7-3 and MDA-2 cells

4.3.2.1 Motility of cell lines exposed to GSK J4

As seen in figure 4.6 there was no significant difference in speed between the different experimental repeats at any concentration of GSK J4 for HeLa, HT1080, MCF7-3 and MDA-2 cell lines (Kruskal-Wallis test; $p>0.05$). As a result, all the data for each cell line have been grouped together for further analysis.

GSK J4 is a selective JMJD3 inhibitor, it acts to increases methylation[161]. Five GSK J4 concentrations, chosen based on the literature[192-195], were tested to determine a dose response and investigate any effect on cell motility.

There were significant differences in motility within each cell line between the GSK J4 concentrations (Kruskal-Wallis tests $p<0.001$). As can be seen in figure 4.7 and table 4.6, these differences in motility between the groups of cells are very small, the range in median values within cell lines is; HeLa, 2.05 µm/hour; HT1080, 7.85 µm/hour; MCF7-3, 21.87 µm/hour; MDA-2, 13.08 µm/hour. There is trend for higher GSK J4 concentrations to result in lower cell speed.

Figure 4.7 shows the median value and distribution of motility for each concentration of GSK J4 in each cell line. All data had non-normal distribution of cell speeds (Lilliefors corrected Kolmogorov-Smirnov test $p<0.001$). As with MS-275, GSK J4 does impact cell motility but there is no correlation between concentrations hence all concentrations of GSK J4 will be used when calculating estimates of heritability.

Table 4.6: Median speed (µm/hour) and N of cell lines for different GSK J4 concentrations (nM)

GSK J4 Concentration (nM)	**Median Speed [N]**			
	HeLa	**HT1080**	**MCF7-3**	**MDA-2**
0	13.83 [934]	43.86 [663]	36.06 [304]	37.74 [275]
10	12.62 [1073]	42.13 [709]	21.21 [444]	24.66 [560]
30	11.78 [679]	36.01 [747]	27.8 [411]	25.68 [497]
60	12.59 [577]	38.39 [579]	43.08 [367]	27.68 [482]
100	12.72 [727]	39.32 [549]	26.64 [498]	26.98 [609]

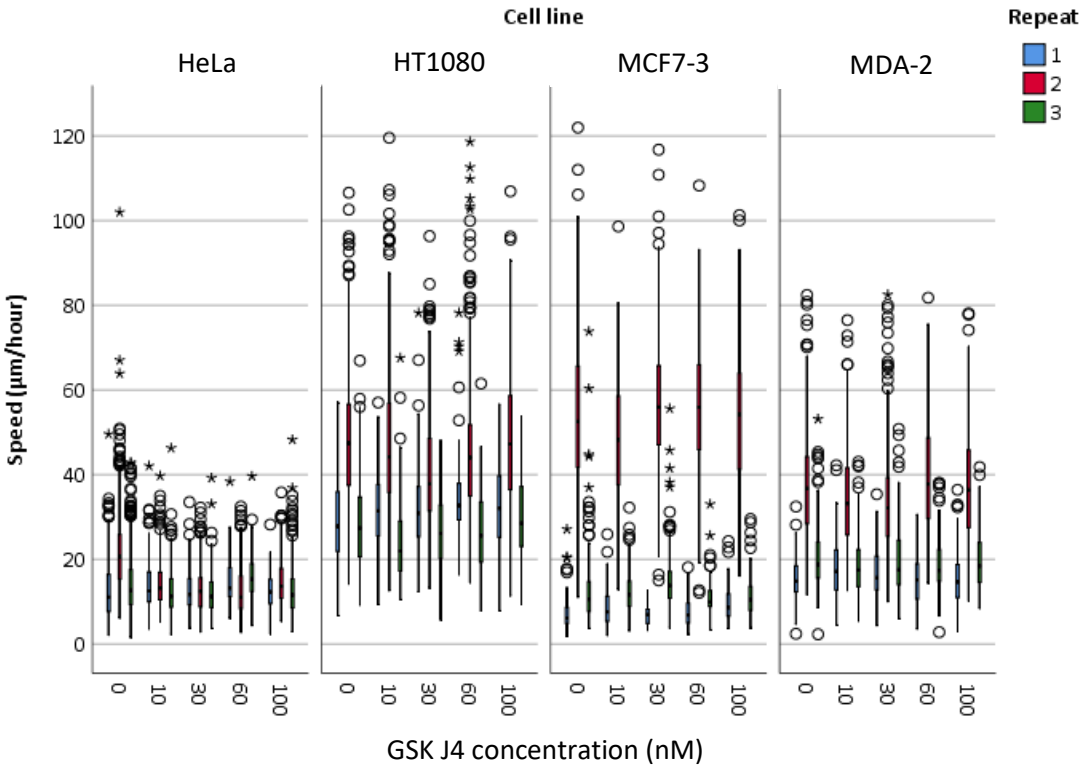


Figure 4.6: Distribution of HeLa, HT1080, MCF7-3 and MDA-2 cell speeds at GSK J4 concentrations (nM) for each experimental repeat. The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. Data gathered from three independent experiments. N for each cell line, GSK J4 concentration and repeat is; HeLa 0 nM repeat 1: 286, repeat 2: 706, repeat 3: 603, HeLa 10 nM repeat 1: 200, repeat 2: 549, repeat 3: 324, HeLa 30 nM repeat 1: 142, repeat 2: 305, repeat 3: 232, HeLa 60 nM repeat 1: 53, repeat 2: 357, repeat 3: 167, HeLa 100 nM repeat 1: 109, repeat 2: 339, repeat 3: 279; HT1080 0 nM repeat 1: 113, repeat 2: 875, repeat 3: 186, HT1080 10 nM repeat 1: 95, repeat 2: 574, repeat 3: 40, HT1080 30 nM repeat 1: 92, repeat 2: 567, repeat 3: 88, HT1080 60 nM repeat 1: 83, repeat 2: 409, repeat 3: 87, HT1080 100 nM repeat 1: 145, repeat 2: 317, repeat 3: 87, MCF7-3 0 nM repeat 1: 74, repeat 2: 330, repeat 3: 174, MCF7-3 10 nM repeat 1: 32, repeat 2: 211, repeat 3: 201, MCF7-3 30 nM repeat 1: 45, repeat 2: 203, repeat 3: 163, MCF7-3 60 nM repeat 1: 49, repeat 2: 235, repeat 3: 83, MCF7-3 100 nM repeat 1: 68, repeat 2: 259, repeat 3: 171, MDA-2 0 nM repeat 1: 103, repeat 2: 521, repeat 3: 115, MDA-2 10 nM repeat 1: 132, repeat 2: 297, repeat 3: 131, MDA-2 30 nM repeat 1: 130, repeat 2: 286, repeat 3: 81, MDA-2 60 nM repeat 1: 90, repeat 2: 274, repeat 3: 118, MDA-2 100 nM repeat 1: 133, repeat 2: 360, repeat 3: 116.

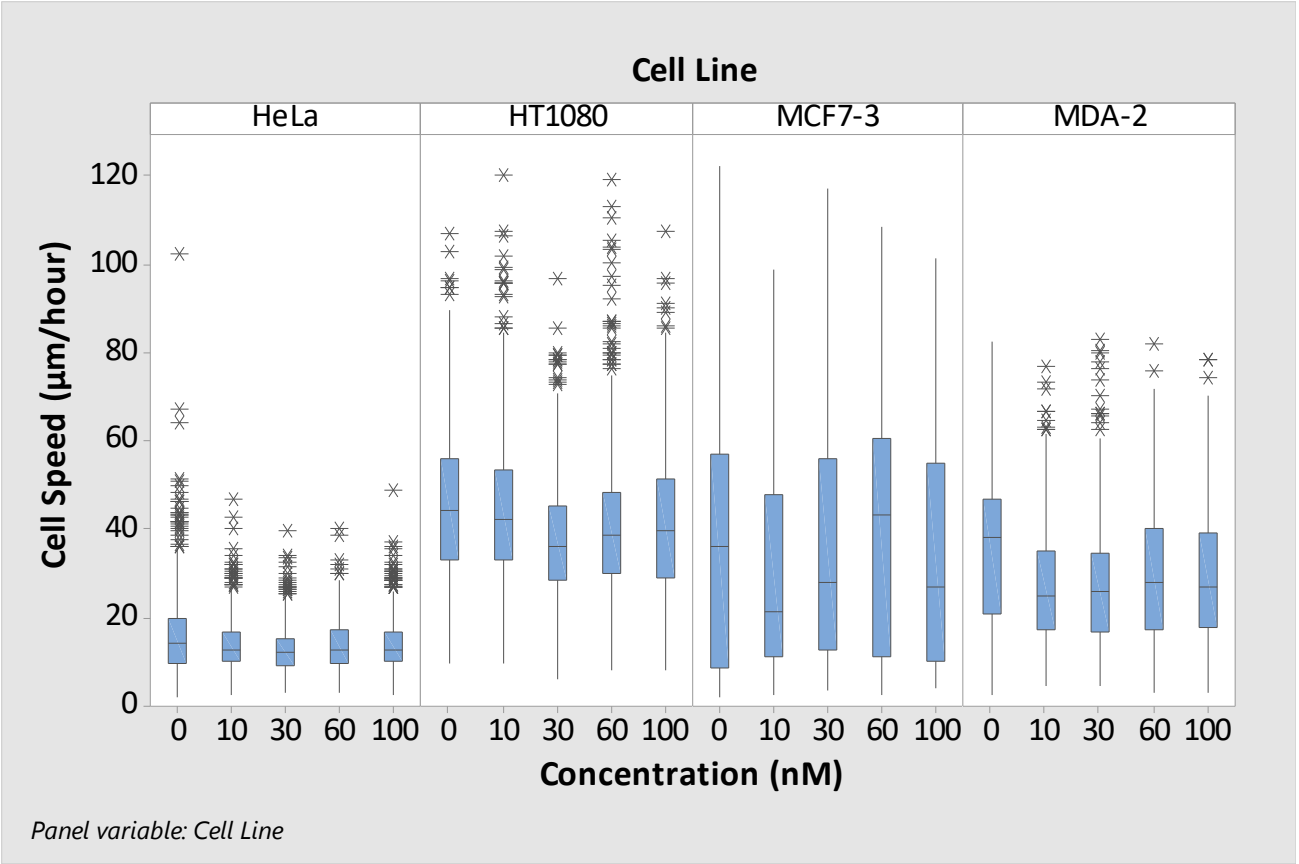


*Figure 4.7: Distribution of cell speeds for HeLa, HT1080, MCF7-3 and MDA-2 cell lines for GSK J4 concentrations (nM). The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. Data was gathered from three independent experiments. The N for each cell line is; HeLa 3,990, HT1080 3,247, MCF7-3 2,024 and MDA-2 2,423.*

HeLa cells in 10 & 30 nM and HT1080 cells in 0 & 10 nM GSK J4 had significant differences in cell speed between generations in (Kruskal-Wallis tests $p<0.001$). As shown in table 4.7 for both concentrations of HT1080, motility in the 2nd & 3rd generations were higher than the 1st generation and at both concentrations of HeLa cells motility in the 1st generation was higher than the 3rd generation (Mann-Whitney U tests $p<0.004$).

There were also significant differences in motility between cells that had different cell fates (Kruskal-Wallis tests $p<0.001$). As shown in table 4.8 all cell lines showed some differences between cell fates but there was no consistent pattern correlated with GSK J4 concentration. Upon further investigation it was found that the differences in motility between cell fates may account for the differences in motility between generations. In HT1080 cells those that moved off screen (of which there were a higher proportion in generations 2 and 3) were significantly faster than those that did not divide (of which there were a higher proportion in generation 1). For HeLa cells those that did not divide were significantly slower than those that did, meaning the increase in speed between generations could show selection over time for increased motility. (Mann-Whitney U tests $p<0.001$).

Table 4.7: Median speed (µm/hour) and N for each of the first three generations in HeLa, HT1080, MCF7-3 and MDA-2 cell lines for five concentrations of GSK J4; 0, 10, 30, 60 and 100 nM.

GSK J4 Concentration (nM)	**Cell line**	**HeLa**			**HT1080**			**MCF7-3**			**MDA-2**		
	Generation	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
0		14.66 [94]	15.19 [159]	13.59 [279]	31.45 [84]	43.61 [77]	46.22 [83]	41.82 [120]	40.99 [100]	29.20 [56]	32.71 [58]	25.83 [78]	36.86 [66]
10		14.79 [93]	14.17 [187]	12.54 [321]	32.45 [85]	38.09 [84]	40.28 [104]	29.48 [104]	25.12 [143]	18.17 [109]	26.74 [71]	23.26 [125]	22.01 [199]
30		13.86 [77]	11.97 [150]	11.08 [247]	32.95 [91]	32.54 [82]	37.23 [97]	39.26 [118]	23.99 [156]	25.48 [109]	28.73 [74]	23.38 [133]	21.59 [159]
60		15.59 [60]	14.27 [112]	12.57 [202]	31.92 [79]	31.44 [78]	37.05 [82]	37.80 [97]	43.16 [118]	38.78 [110]	26.99 [65]	27.86 [106]	24.15 [160]
100		16.07 [63]	12.68 [122]	12.55 [217]	32.04 [79]	37.19 [104]	41.20 [118]	21.30 [98]	15.03 [150]	16.33 [138]	28.34 [93]	25.52 [156]	23.66 [206]

Table 4.8: Median speed (µm/hour) and N for each cell fate in HeLa, HT1080, MCF7-3 and MDA-2 cell lines for five concentrations of GSK J4; 0, 10, 30, 60 and 100 nM.

GSK J4 Concentration (nM)	HeLa				HT1080				MCF7-3				MDA-2			
	Non-divided	Divided	Died	Moved off screen	Non-divided	Divided	Died	Moved off screen	Non-divided	Divided	Died	Moved off screen	Non-divided	Divided	Died	Moved off screen
0	12.53 [427]	14.29 [436]	19.23 [63]	10.59 [8]	36.85 [243]	47.82 [298]	27.87 [23]	50.32 [99]	11.02 [133]	26.46 [93]	39.24 [10]	58.61 [68]	31.84 [113]	41.41 [113]	23.45 [9]	41.31 [40]
10	12.08 [518]	12.97 [506]	15.83 [44]	19.61 [5]	39.16 [268]	44.06 [326]	24.58 [34]	49.09 [81]	13.69 [195]	18.15 [171]	43.23 [23]	50.06 [55]	22.64 [274]	25.42 [245]	18.9 [4]	37.18 [37]
30	10.96 [316]	12.12 [305]	13.34 [57]		33.13 [297]	37.56 [350]	26.2 [22]	40.61 [78]	13.83 [174]	26.54 [144]	41.24 [13]	56.16 [80]	22.49 [219]	27.06 [220]	29.78 [5]	29.49 [53]
60	12.06 [276]	13.68 [265]	12.58 [26]	17.51 [10]	37.12 [247]	41.02 [265]	29.42 [21]	44.5 [46]	29.68 [161]	41.81 [137]	40.69 [9]	60.51 [60]	23.99 [224]	28.11 [213]	41.49 [6]	45.55 [39]
100	12.16 [336]	13.50 [339]	13.11 [43]	14.41 [9]	36.42 [161]	40.76 [241]	24.16 [20]	44.62 [127]	14.38 [237]	26.64 [202]	28.54 [7]	52.83 [52]	24.95 [275]	29.16 [262]	18.73 [11]	34.93 [61]

4.3.2.2 Heritability in cell lines exposed to GSK J4

Cell speed is significantly heritable for all cell:cell relationships in all cell lines and at all concentrations of GSK J4. Table 4.9 shows H^2 values range from 0.953 - 0.264 and R^2 values range from 0.91 - 0.07. As stated in section 2.10.6 outliers and influential data points have been removed from this analysis. H^2 was estimated using simple unweighted regression of cell speeds between related cells, figure 4.8 shows how H^2 has been calculated in MDA-2 cells at varying concentrations of GSK J4 for the mother:daughter relationship.

As seen in table 4.9 and figure 4.9 there is a general trend of H^2 and R^2 values for cousin:cousin cells to be lower than that of mother:daughter and sister:sister cells. However, this is not true for all groups; HeLa cells in 100 nM, HT1080 cells in 0 nM and MCF7-3 cells in both 50 and 100 nM GSK J4 all have cousin:cousin values that are greater than at least one other cell:cell relationship.

Due to the number of post-hoc comparisons it was not possible to determine significance either between the concentrations of GSK J4 or between the different cell:cell relationships for the estimates of H^2 of each cell line. There does not appear to be a clear pattern in either H^2 or R^2 values correlating with GSK J4 concentrations.

Table 4.9: Mean trait value, R^2 and broad-sense heritability of motility in HeLa, HT1080, MCF7-3 and MDA-2 cell lines for GSK J4 concentrations (nM) for three cell:cell relationships.

Cell Line	GSK J4 Concentration (nM)	No of families	Mean speed (µm/hour)	Heritability of cell motility H^2 (R^2)		
				Sister:sister	Mother:daughter	Cousin:Cousin
HeLa	0	135	15.63	0.799*** (0.64)	0.783*** (0.61)	0.717*** (0.51)
	10	159	13.62	0.656*** (0.43)	0.487*** (0.24)	0.479*** (0.23)
	30	123	12.56	0.580*** (0.34)	0.633*** (0.40)	0.459*** (0.21)
	60	101	13.70	0.669*** (0.45)	0.774*** (0.6)	0.673*** (0.45)
	100	107	13.79	0.390*** (0.15)	0.560*** (0.31)	0.438*** (0.19)
HT1080	0	42	45.20	0.823*** (0.68)	0.749*** (0.56)	0.811*** (0.66)
	10	52	44.38	0.656*** (0.43)	0.738*** (0.54)	0.518*** (0.27)
	30	48	37.6	0.824*** (0.68)	0.697*** (0.49)	0.601*** (0.36)
	60	42	41.19	0.733*** (0.54)	0.768*** (0.59)	0.485*** (0.24)
	100	59	41.33	0.728*** (0.53)	0.720*** (0.52)	0.264** (0.07)
MCF7-3	0	28	35.52	0.891*** (0.79)	0.909*** (0.83)	0.885*** (0.78)
	10	83	29.1	0.913*** (0.83)	0.870*** (0.76)	0.865*** (0.75)
	30	53	34.96	0.897*** (0.81)	0.910*** (0.83)	0.858*** (0.74)
	60	55	39.11	0.898*** (0.81)	0.917*** (0.84)	0.921*** (0.85)
	100	69	33.06	0.953*** (0.91)	0.884*** (0.78)	0.909*** (0.83)
MDA-2	0	31	35.62	0.764*** (0.58)	0.806*** (0.65)	0.662*** (0.44)
	10	70	27.03	0.816*** (0.67)	0.677*** (0.46)	0.565*** (0.32)
	30	80	27.32	0.874*** (0.77)	0.824*** (0.68)	0.687*** (0.47)
	60	80	29.84	0.833*** (0.69)	0.803*** (0.65)	0.802*** (0.64)
	100	103	29.80	0.760*** (0.58)	0.786*** (0.62)	0.702*** (0.49)

**p<0.01, ***p<0.001

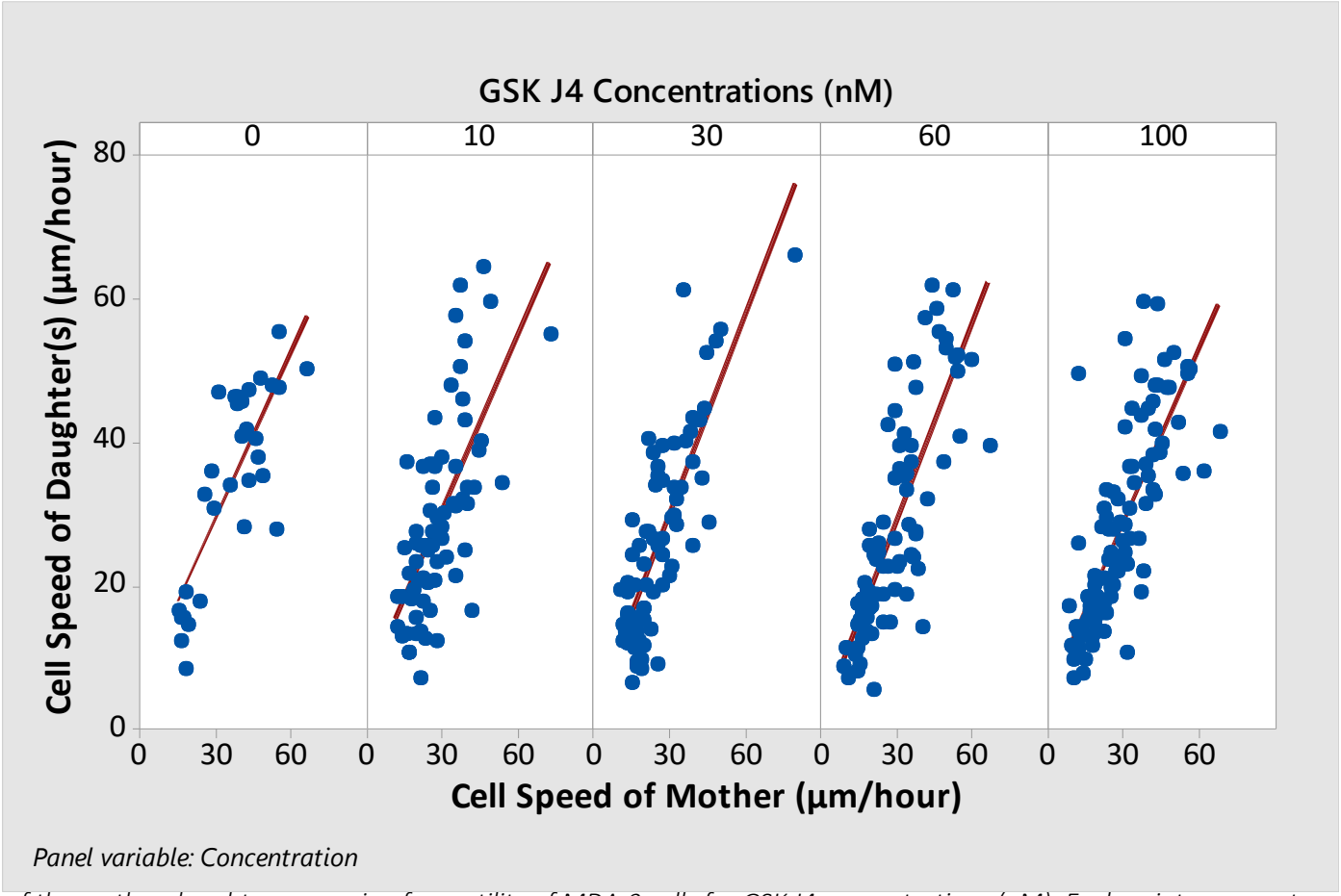


Figure 4.8: Example of the mother:daughter regression for motility of MDA-2 cells for GSK J4 concentrations (nM). Each point represents a mother cell and the mean speed of her daughters, the slope of the regression line is the estimate of broad-sense heritability. As the cells are clonal, the same method is applicable to other cell:cell relationships. N for each group is 0 nM = 31, 10 nM = 70, 30 nM = 80, 60 nM = 80 and 100 nM = 103.

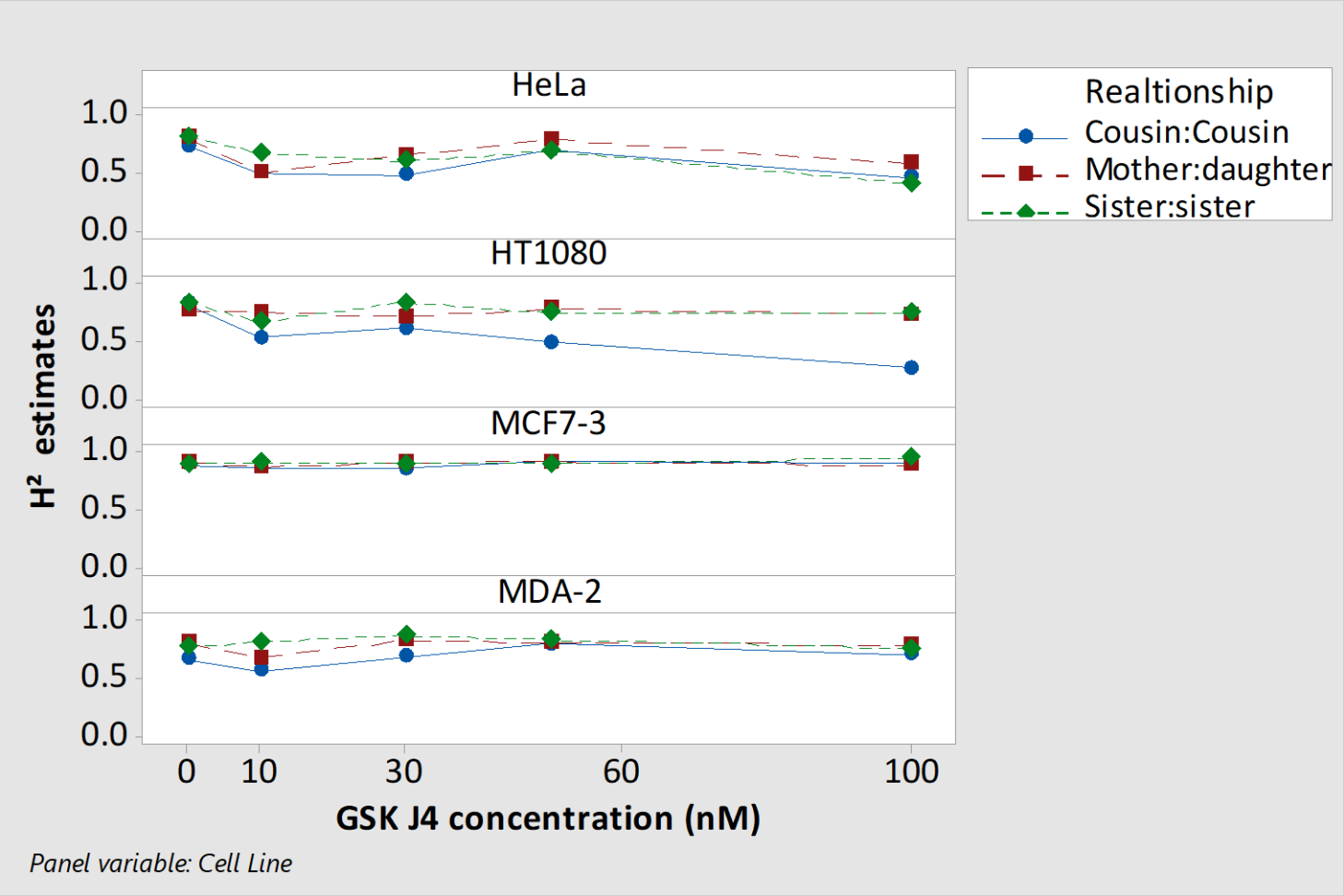


Figure 4.9: Heritability estimates of cell:cell relationships for HeLa, HT1080, MCF7-3 and MDA-2 cell lines at GSK J4 concentrations (nM). HeLa; 0 nM = 135, 10 nM = 159, 30 nM = 123, 60 nM = 101 and 100 nM = 107. HT1080; 0 nM = 42, 10 nM = 52, 30 nM = 48, 60 nM = 42 and 100 nM = 59. MCF7-3; 0 nM = 28, 10 nM = 83, 30 nM = 53, 60 nM = 55 and 100 nM = 69. MDA-2; 0 nM = 31, 10 nM = 70, 30 nM = 80, 60 nM = 80 and 100 nM = 103.

4.3.3 FITC-Dextran distribution

The distribution of FITC-Dextran is not uniform throughout the media for any of the concentrations tested. Figure 4.10 shows heatmapped images for the three concentrations of FITC-Dextran in media; 1 mg/mL, 1 µg/mL and 1 ng/mL. The images show that particle distribution in media is not uniform for a range of concentrations. Figure 4.11 shows the distribution of FITC-Dextran media with fluorescently labelled HT1080 cells present it demonstrates that there is variation in particle distribution throughout the media at a cellular level. These images do not model the specific distribution of all particles within culture media but do demonstrate that particles within media are not uniformly distributed.

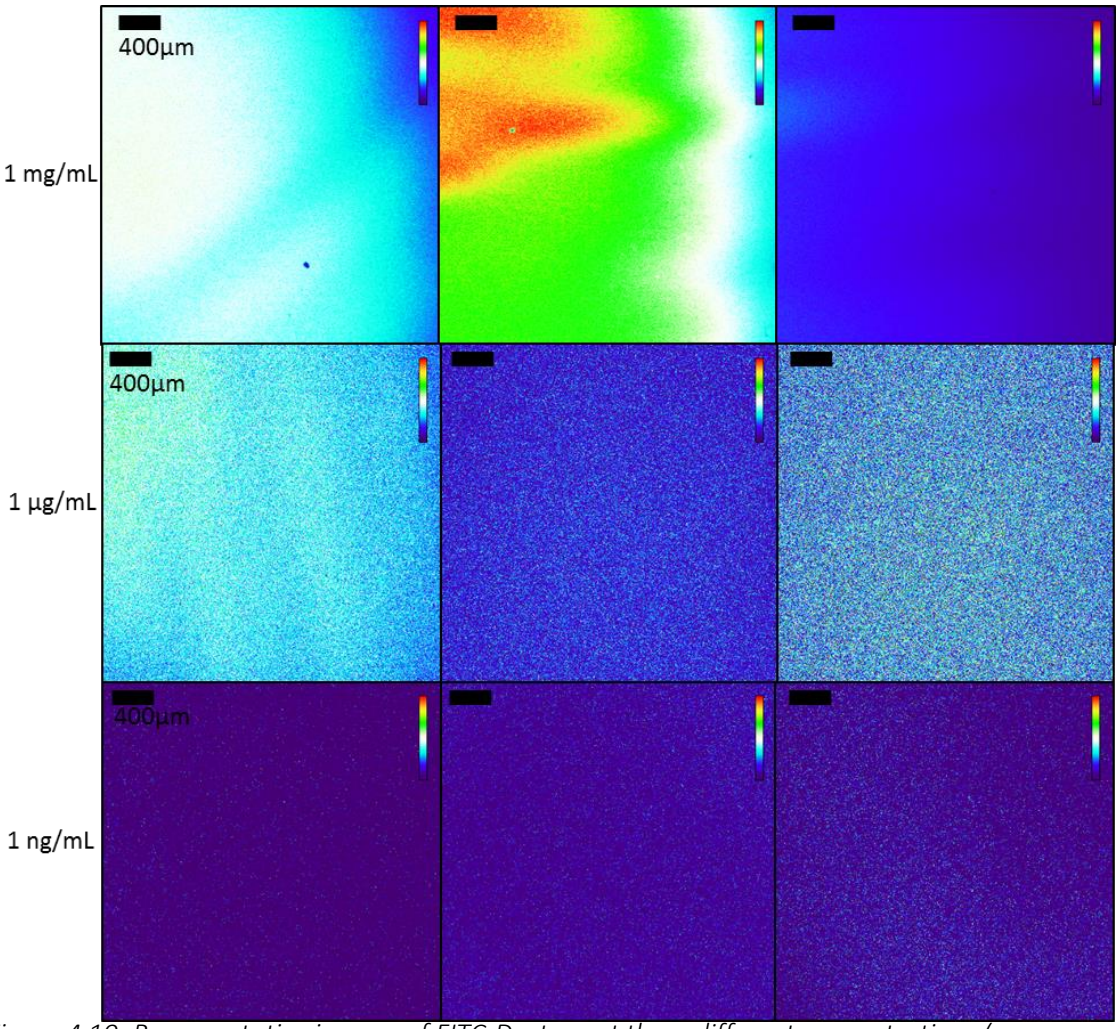


Figure 4.10: Representative images of FITC-Dextran at three different concentrations (mg, µg and ng /mL) in culture media. Each row represents one of the three FITC-Dextran concentrations, with each column being a repeat. Scale bars at 400 µm are shown in black on the top left corner of each image and the heatmap indicator bar is also scaled to 800 µm. Heatmap indicators show the highest density of particles (red) to the lowest (purple).

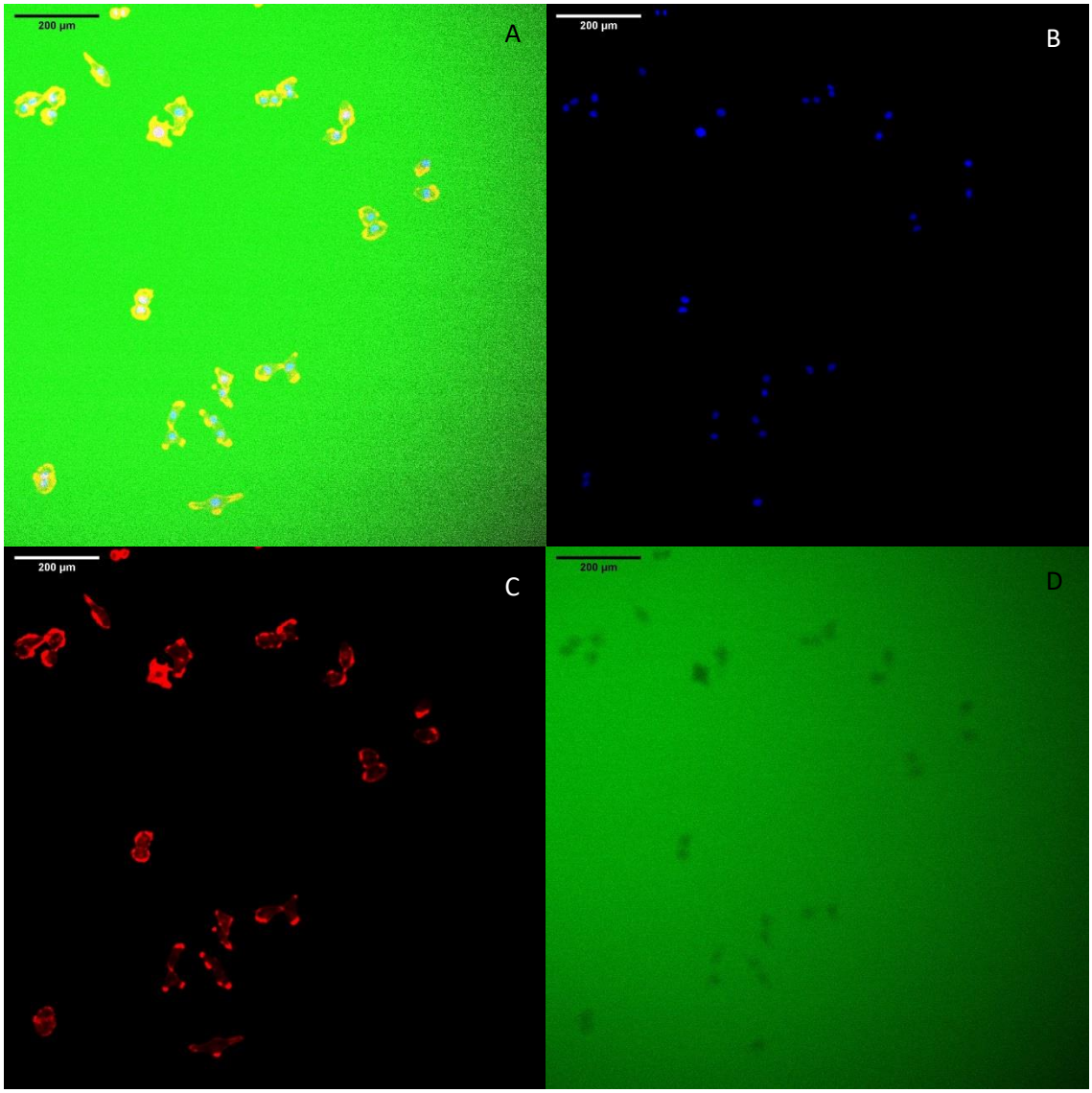


Figure 4.11: Fluorescently labelled HT1080 cells in 1 mg/mL FITC-Dextran media. Each panel (A-D) is one fluorescent channel of the same image, a scale bar representing 200 µm is located in the top left hand corner of each image. A) Composite image containing all colour channels. B) blue channel:DAPI staining A-T rich regions of DNA C) red channel: phalloidin546 staining F-actin D) green fluorescent channel: FITC-Dextran.

4.4 Discussion

Epigenetic modifications are not permanent and can be quickly altered. As epimutations occur at a higher frequency than genetic mutations[164] and can result in stably heritable phenotypes[47 167 168] they may play a role in regulating cell phenotypes such as motility. If true, then this might have explained the decrease in H^2 estimates seen in chapter 3 as these cells would have accumulated more epigenetic modifications since their shared ancestor increasing their dissimilarity to each other.

If epigenetic modifications are responsible for lower H^2 estimates between cousin cells, then it was hypothesised that inhibition of such modifications should result in a) an overall reduction of heritability estimates and b) eliminate differences in H^2 between cell:cell relationships. Addition of MS-275 and GSK J4 do not appear to affect cell motility or heritability estimates. At all concentrations of MS-275 and GSK J4 for all cell lines tested and all cell:cell relationships measured, H^2 of motility is still highly significant (tables 4.5 and 9, figures 4.4 and 4.8). Heritability estimates between cousin:cousin cells remain lower than estimates between mother:daughter and sistser:sister cells. These results suggest that no individual epigenetic modification examined in isolation, plays a major role in regulating heritability of cell motility.

Environmental variation was also investigated to see whether this could account for the lower H^2 estimates. At all three concentrations of FITC-Dextran measured, seen in figures 4.10 and 4.11, it is clear that there is environmental variation in *in vitro* culture at a cellular level. These results suggest environmental variation may explain at least some of the variation in heritability estimates as it suggests the correlation in relative's motility could be due to a correlation in their microenvironment rather than genetic variation.

4.4.1 Cell motility and the effect of epigenetic inhibitors MS-275 and GSK J4

Based on the literature five concentrations of GSK J4 and MS-275 were chosen[161 162 184 197 209 210]. These concentrations were tested to determine whether there was a detrimental effect on cell motility that might preclude the use of these drugs or certain concentrations from this study.

It was predicted that MS-275 would act to reduce cell motility (section 4.1.5.1) however this was not seen in the results (section 4.3.1.1, figure 4.3). The predicted effects of GSK J4 on cell motility were unknown due to the complex nature of cell signalling pathways. For both MS-

275 and GSK J4 there were significant differences between cell motility of different concentrations ($p<0.001$), however as seen in figures 4.3 and 4.7, these were very small and there was no correlation concentration and motility, higher drug concentrations did not consistently cause an increase or decrease in cell speed.

Cell signalling pathways are highly complex and interlinked with cell functions being regulated at multiple levels[200 205]. Many cancer cells have aberrant signalling pathways, abnormal epimutation rates and there is extensive evidence of deregulation of transcription factors and histone modifying enzymes[222]. The intricacy of these pathways and aberrant cancer cell signalling makes it difficult to predict the effects of inhibiting a single molecule on cancer cell phenotype[214]. Inhibition of specific enzymes (in this case HDAC1, HDAC3 and JMJD3) may not be enough to alter gene expression as small alterations in cell signalling pathways could result in the maintenance of the motility phenotype[208].

Acetylation and methylation are two of the best studied epigenetic modifications[180] and methylation is one of the most common types of epimutation[29 47 164]. It was postulated that if epigenetic modifications played a role in regulating cell motility then alteration of either acetylation or methylation rates would have an effect. These results suggest that either epigenetic modifications do not play a role in regulating cell motility or that the inhibition seen in this project was not great enough to have an effect.

MS-275 acts to increase acetylation and GSK J4 to increase methylation. These drugs should generally have opposing effects on gene transcription. Not only was no overall effect on cell motility observed but no differences in motility between the two drugs. GSK J4 and MS-275 treatment may not have resulted in the complete inhibition of their respective enzymes. Higher concentrations of each drug were initially trialled but these resulted in complete cell death. Use of alternative inhibitors such as siRNA, small double stranded RNA molecules which interfere with the expression of specific genes by preventing translation, would ensure complete inhibition of these specific enzymes[223]. MS-275 and GSK J4 are selective inhibitors[162 184] and gene expression can be influenced by multiple epigenetic mechanisms contributing to the histone code[47 172]. Multiple types of inhibitor may be needed to disrupt epigenetic inheritance. Complete inhibition of all acetylation or methylation may show that these epigenetic modifications do play a role in regulating cell motility however this would require further testing as complete inhibition of an epigenetic modification may also result in cell death.

There was no clear dose response and no single concentration had a detrimental effect on cell motility. Thereby all concentrations of drugs were used to measure heritability.

4.4.2 Checking for bias in cell motility

As in section 3.4.1 cells were checked to ensure motility did not alter over the course of the experiment due to confounding factors, this was checked through comparison of generations and cell fate.

All cells present at the beginning of the experiment were classed as generation 1 irrespective of their position in the cell cycle. The decision was made not to perform cell cycle synchronisation (addition of chemicals into the culture media causing the cells to pause at a specific stage in the cell cycle)[224] as this can have negative effects on cell growth and proliferation and would add another element of complexity when factoring in heritability estimates and comparison between groups[224 225].

As seen in tables 4.3 and 4.7, HeLa and HT1080 cells in GSK J4 and HeLa and MCF7-3 cells in MS-275 showed differences in cell motility between generations ($p<0.004$). For these groups of cells, it would appear that motility of the population changed throughout the experiment. These trends could be explained by a difference in motility between cell fates ($p<0.001$) and did not correlate to drug concentration. As shown in tables 4.4 and 4.8, a general trend was seen in the motility of cell fates, those that moved off screen had higher motility and those that did not divide had lower motility. The specific differences between cell fates and the proportion of these cell fates within each generation help explain the differences in average cell motility between generations.

As a result of this analysis generation 1 was excluded from estimates of H^2 and the relationship between their clonal descendants used to calculate heritability (figure 3.2). This an imperfect method of analysis as it leads to the exclusion of certain generations from the results. Another explanation for a change in population motility over time is the action of the inhibitors added to the media. These inhibitors could cause an increase in certain cell fates leading to the differences in motility between generations. The proportion of each cell fate for each drug and concentration was compared and no significant differences ($p>0.07$) found between groups. Future work involving more powerful and rigorous statistical models could eliminate these concerns. Use of statistical software such as R would allow inclusion and comparison of more complex family dynamics making results more reliable when making inferences to population

dynamics. This examination of motility within experiments ensures that any change in population dynamics does not affect the results seen here.

4.4.3 Heritability in the presence of epigenetic inhibitors, MS-275 and GSK J4

GSK J4 reduces demethylase activity, leading to an increase in methylation and MS-275 reduces HDAC activity, leading to an increase in acetylation[162 209]. The hypothesis was that epigenetic modifications may play a role in regulating cell motility and that epimutations were responsible for a decrease in the genetic contribution to motility. Hence inhibition of such epigenetic modifications would result in an overall reduction in heritability estimates but eliminate differences between more distantly related cells.

For both MS-275 and GSK J4 the H^2 of motility is highly significant (tables 4.5 and 4.9, figures 4.4 and 4.8) however the results do not support the hypothesis. Whilst all estimates are significant as seen in figures 4.5 and 4.9 lower H^2 estimates are still observed between more cousin cells. It may be that due to the complexities of cell signalling, inhibition of a wider array of proteins is needed to cause a measurable effect on heritability[208].

The trend for more distantly related cells to have lower heritability does not hold true for all groups. As seen in tables 4.5 and 4.9 and figures 4.5 and 4.9, Hela cells in 100 nM, HT1080 cells in 0 nM and MCF7-3 cells in 50 and 100 nM GSK J4 and MDA-2 in 30 nM MS-275 all had H^2 estimates for cousin:cousin cells greater than that of other related cells. Whilst not a general trend these anomalous results could indicate that the inhibitors are having an effect.

The results seem to suggest that even with epigenetic inhibitors H^2 estimates are still lower for cousin:cousin cells than mother/daughter and sister cells. However due to the number of tests needed and the increased risk of type 1 errors (see section 2.10.6)[105] post-hoc comparisons comparing the significance of H^2 estimates between cell lines, concentrations and types of drug were not statistically possible. It cannot be ruled out that the lower heritability estimates are due to a reduction in the genetic contribution to motility and further investigation would be needed into whether this reduction is due to complex interactions in epigenetic modifications or due to genetic mutations.

If epigenetic modifications are responsible for a reduction in genetic contribution to motility then future work testing a wider range or a combination of epigenetic inhibitors could help to show this. Inhibitors targeting phosphorylation could be of interest as histone phosphorylation

changes chromatin structure by altering the interactions between protein:DNA and protein:protein[226 227]. As it is not only the chromatin structure that is altered but also protein interactions this can induce crosstalk between other chromatin modifications[226]. Using a phosphorylation inhibitor in combination with a methylase inhibitor may have an amplified effect on reducing gene expression and the number of epimutations that occur between generations. Tests such as PCR and ChIP sequencing could help determine whether genetic or epigenetic mutations had occurred however it would only be possible to indirectly compare mutations to heritability estimates as it is not possible to measure families of cells[228 229].

4.4.4 General environmental variation

Environmental heterogeneity is an important factor in heritability. Heritability is not the opposite of phenotypic plasticity. A character may have perfect heritability in a population and still be subject to great changes resulting from environmental variation[130]. Understanding the amount of environmental variation is crucial to interpreting heritability results[128 129]. The results seen here, figures 4.10 and 11, show that the lower H^2 estimates for cousin cells could be due to an increase in environmental variation[46]. FITC-Dextran with an average mol. wt. of 70,000 Da was added to culture media at three concentrations, 1 mg/mL, 1 µg/mL and 1 ng/mL. Figures 4.10 and 4.11 show that dextran is heterogeneous within the media.

Using the fluorescent properties of FITC-Dextran has allowed images to be collated on the heterogeneity within *in vitro* culture media. Whilst the heterogeneity seen in cell culture will not be on the same scale as that seen in an *in vivo* environment this variation in dextran occurs at a cellular level (average cell size between 20 – 80µm), causing fluctuations in the cellular microenvironment. This shows that cells in culture are not in a homogenous environment and that there is heterogeneity which could apply selection for dispersal.

This experiment does not represent the distribution of media components, it merely shows that culture media is not evenly distributed. The molecular weight of the dextran used in these experiments was chosen as it is relatively close to that of albumin (65-70 kDa), one of the most common proteins in serum[230]. This experiment could be repeated with dextran of different molecular weights to check that this heterogeneity is maintained at a cellular level for different compounds.

By comparing the distribution of different concentrations of dextran in media it is possible to show that even in *in vitro* cell culture the cells environment is not uniform. This can be seen in

figure 4.10 where the different fields of view show different distributions of dextran for the same concentration of solution. This experiment has not specified the distribution of solution but visually demonstrated that any liquid will not be uniform in its distribution and, that for a range of concentrations the variation in distribution will vary at a cellular level.

This environmental variation does not appear to have a pattern or vary by a specific amount. This apparent 'randomness' in variation would appear to support the hypothesis that the variation in H^2 estimates are caused by an increase in environmental variation. Whilst these results show that environmental variation will alter over time this would be expected to vary randomly for each generation and between cell types. As heritably estimates are consistently lower in more distantly related cells it would be unlikely that an increase in environmental variation is the sole factor in the variation of H^2 estimates.

Future work could quantify the environmental heterogeneity as it should be possible to use the variogram function, a type of spatial statistics that links the variance of a stochastic process and its degree of spatial dependence[39 231]. Combining information on the amount of environmental variation and the epigenetic and genetic mutations that occur would give in depth analysis on how selective pressures act on cancer cell lines *in vitro*.

It may also be possible to refine the regression analysis to compare related cells that remain within close proximity to each other and compare these H^2 estimates to those of related cells who migrate further away from each other. Whilst not directly measuring the amount of environmental variation it would help elucidate the contribution of the environment to H^2 estimates. This would involve complicated analysis at the tracking stage of the data analysis and quantification of the temporal-spatial variation to allow definition of proximity parameters.

The results suggest that a heterogeneous environment affects cell motility suggesting that dispersal theory could be applied to cancer. Cell motility is a key part of the metastasis cascade, a common phenomenon in cancer[25], dispersal theory could help to explain why tumours metastasise. Culture of cells in different environments and comparisons of their heritability and motility will provide further insight into whether dispersal theory applies to cancer evolution.

Chapter 5 – Applying experimental evolution techniques to cancer

5.1 Introduction

As motility has been identified as a heritable trait (chapters 3 and 4), natural selection can be used to select for cancer cell populations with this phenotype in a new application of the principles of experimental evolution techniques[48].

In this chapter, cell biology methods and experimental evolution techniques are adapted and combined to test whether dispersal theory applies to cancer cells in *in vitro* culture and whether the phenotype of cell motility can be selected for over time. Selection experiments, culturing of cancer cell populations in resource restricted environments, are conducted and multiple phenotypes in both the evolved and ancestral populations measured and compared to see whether adaptation has occurred. Results suggest that dispersal theory does apply to cancer cells with resource restriction selecting for an increase in cell motility. Not all experimental evolution techniques were successfully adapted to cancer cell lines, but these experiments are proof of concept and enable further work to be conducted in this area.

5.1.1 Application of evolutionary theories to cancer biology

The mechanisms of cancer cell metastasis are well studied however few of the evolutionary and ecological theories of metastasis have been tested experimentally. Natural selection is the only process that leads to adaptation. As motility has been shown to be a heritable trait, experimental evolution techniques can be used on *in vitro* cancer cells[19 44]. Applying selective pressures for cell motility should cause a shift in the population's average motility over time. How long this shift takes to appear could vary based on other evolutionary processes such as mutation, random drift, strength of selection and population size. The rate of selection will also depend on the selective pressure applied[113]. Answers to such questions as how a metastatic phenotype is selected for, may eventually lead to new methods of treatment or prevention as well as tests for disease progression or severity[29 36].

This project a novel approach and various experimental evolution techniques adapted and subsequently tested for their use on cancer cells in culture. These modified techniques will permit the study of evolutionary processes in cancer cell culture allowing testing of evolutionary theories establishing causality and increasing our knowledge of disease progression[36 48].

5.1.2 Dispersal theory and cell motility

Dispersal theory proposes that heterogeneity in resources selects for migration as organisms able to migrate will reduce their kin competition and be able to exploit more resources[29 36 39]. Dispersal theory has been modelled mathematically[19 44] and there is experimental evidence of its role in migration in both ecology and microbiology[39 231-233].

Dispersal theory predicts that an individual cancer cell migrating away from the primary tumour would reduce competition between its clonal relatives remaining in the primary tumour, increasing their chance of survival and proliferation of the genotype, thereby indirectly increasing its own fitness even if it is unsuccessful in dispersing[19 39 44]. The small number of dispersing cells that are successful in establishing a metastatic tumour will be able to exploit untapped resources with little competition. Dispersal theory explains how motility may be selected for and metastasis may evolve even when the individual cost of migration is high.

Application of dispersal theory to cancer biology would allow clarification of the ecological causes of natural selection[36]. Metastasis could be considered an evolutionary consequence of this selection[19 44]. Primary tumours are highly heterogeneous in their microenvironment[30] and this resource heterogeneity could create competition with selection favouring motile cells able to disperse away from resource limitation[19 44]. Individual cells may become motile to move around the tumour either away from hypoxic areas or towards areas with higher nutritional resources, metastasis away from the tumour may be a by-product of this dispersal.

5.1.3 Experimental evolution techniques

Experimental evolution of cancer cells *in vitro* will allow testing of multiple different selective pressures and their effects on cell motility. Whilst experimental evolution techniques can test evolutionary hypotheses, cell biology techniques, such as western blots, will be needed to quantify the molecular responses to the selective pressures and elucidate the mechanisms causing a change in cell phenotype. A range of methods will be used to quantify cell phenotype as differences may be seen between the population as a whole and the individuals within that population. A summary of existing experimental evolution and cell biology techniques are shown in table 5.1. The majority of experimental evolution techniques need to be adapted for use in cancer cell culture[36] and there are many methods and techniques

already widely used in molecular cell biology which can be co-opted and used to test evolutionary hypotheses.

Table 5.1: Experimental evolution and cell biology techniques that can be combined and adapted for testing of evolutionary hypotheses in cancer cell lines.

Technique	Method	Outcome
Selection experiments	Long term culture of cancer cell populations under resource restriction then measurement of cell phenotypes.	Resource restriction should select for increased cell motility. Selection experiments should result in populations with increased motility.
Growth assay	Measures the rate of proliferation within a population of cells[14].	Population growth can be used as a measure of fitness and the proliferation rate is often used to determine cell response to experimental parameters[85].
Common garden assay	Populations of cells adapted to a specific environment are placed into a different environment, usually when an evolved population are placed back into their ancestral environment and vice versa.	Helps disentangle the effects of genotype and the environment[234] as information provided on whether changes in phenotype are an adaptive or a plastic response to the environment[14].
Competition assay	Co-culture of two cell populations creating competition for resources. The number of surviving cells in each population or each population's contribution to future generations is measured[14].	Measures the relative fitness of two populations to each other[235] and shows how successful a combination of phenotypes are at ensuring survival in the presence of other cell types[14 51 236].
Fluorescent probes	Fluorescently label specific cell organelles or cellular processes[216 236].	Allows detection of live cells[217] and makes cells in culture distinguishable from each other.
Wound healing assay	Cells are cultured in a confluent monolayer and a scratch made; the rate of wound healing is then measured[237].	Studies the effects of a variety of experimental parameters on mass cell migration and proliferation[237 238].
3D cell culture	Cells 'stick' together and grow in clumps – termed spheroids, which are free floating in the media[239].	Spheroids offer an *in vitro* model that more accurately reflects clinical expression profiles compared to monolayer cultures[240 241].

5.1.3.1 Selection experiments

Selection experiments apply a selective pressure, such as a decrease in nutrients, to cancer cells in culture[36]. If the selective pressure favours dispersal, cell motility would be expected to increase as cells disperse to try to locate or exploit resources[36]. Cryogenic storage of the ancestor population allows comparisons of cell phenotypes to determine whether adaptation has occurred. Culture of cells in a resource restricted environment over time will increase competition and can be used to test whether dispersal theory applies to cancer[36].

Culture medium provides the necessary nutrients, growth factors and hormones for cell growth as well as helping regulate pH and osmotic pressure[64 80 81]. Reducing the concentration of media components will create competition for that resource. Cells in culture have been shown to be susceptible to alterations in pH, glucose and serum. Glucose is an energy source added to culture media at a precise defined concentration[242]. Unlike glucose, serum is a vital source of nutrients which contains many substances whose concentrations and effects are unknown and vary from batch to batch[243]. Reducing the concentration of either will create competition between cells[64 242 243]. Culture pH is not a resource, however hypoxic environments (often found in tumours) that can induce an acidic pH have been shown to increase metastatic progression and capability[244]. One study by Rofstad et al (2006) demonstrated that acidic extracellular pH increased invasiveness *in vitro* and increased metastatic potential *in vivo* through acidity-induced upregulation of proteolytic enzymes and angiogenic factors[245]. As acidic pH increases cancer cell metastasis it may be that lowering the culture pH could also be a selective pressure for cell motility. Measuring and comparing phenotypes, such as cell motility, in both the evolved and ancestral populations will test for adaptations to competition.

5.1.3.2 Growth and common garden assays

It is already possible to conduct some evolution experiments on cancer cell lines using existing cell biology methods. For example, growth assays can be conducted using either MTT assays or time-lapse microscopy to quantify cell growth[86]. Quantification of growth is an important measure of a cell's response to the environment and growth assays are essential in various other techniques such as common garden and competition assays. Common garden assays are also easily amenable to cancer cell culture as it is possible to change the culture environment of a cell population and measure the response. Cells evolved over time under a specific selective pressure (for example a decrease in serum) can easily be placed back into their

ancestral environment (the original culture media they were adapted to) and the growth rates compared to determine the plasticity of response.

5.1.3.3 Competition assays

Competition assays are an important measure of fitness as *in vivo* cancer cells must compete against other (non-cancerous) cell types as well as non-clonal cancer cells. Competition assays require co-culture of cell populations (table 5.1) and whilst different cell lines may have diverse morphologies, in general individual cells in culture are not distinguishable from each other without a fluorescent tag. Fluorescent probes can be used to distinguish between cell populations however the probes used cannot have an effect on cell fitness as this would add another level of selection to the experiments[236 246]. Therefore, before conducting any competition assays a fluorescent probe would need to be found that had no effect on cell fitness.

5.1.3.4 Direct selection of motile cells

Dispersal theory suggests that competition for resources should be a selective pressure for increased motility however cells in culture have no means of escaping from a resource restrictive environment once their motility is elevated. As molecule diffusion is isotropic (uniform in all directions) and unrestricted in free solution, current techniques cannot produce precise, user defined gradients with specified spatial and temporal outlines. This makes it impossible to 'reward' motile cells within these cultures[247 248]. As well as culturing cells in a resource restricted culture and testing for adaptation it may be possible to directly select for cells with altered motility by physically separating those cells with increased motility. These cells could then be 'rewarded' by being re-cultured under the original conditions. Over time an increase in population motility should be selected for and these cells could be compared not only to an ancestral population but also to their slower-moving companions.

5.1.3.5 3D cell culture and collective migration

3D cell cultures (termed spheroids) express chemical gradients of various nutrients, oxygen and catabolites, similar to those found in tumours[239 249]. There is widespread evidence that gradients in various biomolecules play a role in biological processes, including cancer cell metastasis[248]. There is also evidence of differences between collective migration and single cell migration[250], methods such as time-lapse microscopy provide information on individual cell

motility that can be averaged across a population. Alternative techniques such as wound healing and spheroid spreading assays measure whole population motility and have been used to study spreading behaviour of cancers[237 238 251-253]. Comparison of motility between monolayer and 3D culture will allow scrutiny of the motility phenotype to ensure greater relevance to *in vivo* environments[249].

5.1.4 Aims

This project aims to adapt and apply experimental evolution techniques across to cancer cells in culture. The technical aims of this project involve a) testing of fluorescent tags for their suitability for use in competition assays b) in the use of transwell assays and hydrogel layers in selection of motile cells. If successful adaptation of these methods will allow a new approach for studying cancer cell evolution.

The scientific aims involve the use of selection experiments for long term cell culture in a resource restricted environment to see whether cell motility is selected for over time. Experimental evolution and cell biology techniques will then be used to compare between the evolved and ancestral populations to see whether dispersal theory applies to cancer and to investigate the mechanisms behind any alterations in cell phenotype. Using an experimental evolution approach to investigate cancer cell metastasis will allow testing of evolutionary hypothesis.

5.2 Methods

5.2.1 Cell culture

All cell lines used were lab adapted. As stated in section 2.1, similar cell lines have been obtained from different sources, to distinguish between these they are labelled numerically.

In this chapter cell lines have been cultured in media containing different concentrations of nutrients, to distinguish between groups of the same cell line the defining element of the media has been used as a reference (tables 5.4 and 5.5). All cells were cultured in 5 % CO_2 at 37 °C, unless otherwise specified cells have been cultured as a monolayer.

5.2.2 Adaptation of HeLa, HT1080, MCF7-3 and MDA-2 cell lines to serum replacement media

The variation in serum between batches makes standardisation of protocols difficult[64]. Standardisation of media would be necessary before beginning selection experiments. To solve this problem, many researchers have attempted to create a serum free media and adapt cells across to this new media for culture. In this project attempts were made to adapt HeLa, HT1080, MDA-2 and MCF7-3 cell lines to culture in serum free media. The specific media for each cell line (section 2.1) was supplemented with a replacement serum solution (a premade proprietary formula obtained commercially designed to replace FBS in culture media) and varying concentrations of hormones; Rolipram and is a phosphodiesterase-4 inhibitor, it inhibits degradation of cyclic adenosine monophosphate which is an important intracellular signalling molecule[254], Epidermal growth factor (EGF) stimulates cell growth and differentiation[255], Fibroblast growth factor (FGF) is a cell signalling protein that primarily acts as a mitogen[256] and Dexamethasone is a corticosteroid involved in glucose metabolism[257]. The combination of drugs and concentrations were chosen based on the literature[258]. The working solutions of these are given below.

- 1 x serum replacement solution (sigma)
- 1 mM rolipram (santacruz Biotech)
- 500 ng/mL EGF (jena Bioscience)
- 5 µg/mL FGF (PeproTech)
- 1 mM dexamethasone (Sigma).

Initial cultures began with a 90:10 mix of standard media: serum replacement media. All cell lines were maintained for 1 week in each ratio of media during which time they were cultured as normal. The ratios of media were 90:10, 75:25, 50:50, 25:75, 10:90, 5:95 and 1:99. Cultures could not be maintained in a complete serum free media as cell death occurred.

As cultures could not be maintained without the presence of FBS, the serum replacement media was not used for the experimental evolution techniques and variation was reduced through the use of one batch of FBS for all experiments.

5.2.3 Adaptation of HeLa, HT1080, MCF7-3 and MDA-2 cell lines in 10 % and 0.5 % FBS media over 72 hours

For these experiments four cell lines between the stated passage numbers were used; MCF7-3 passages 14-19, MDA-2 passages 40-47, HeLa passages 4-8 and HT1080 passages 4-9 cells (all obtained from Public Health England). These cell lines were cultured in 225 cm^2 culture flasks in 22 mL media containing either 10 or 0.5 % FBS.

For each experiment HeLa, HT1080, MCF7-3 and MDA-2 cells were cultured for 72 hours in 10 % and 0.5 % FBS nutrient media to allow comparison of cellular characteristics whilst adaptation to nutrient deprivation occurred.

5.2.3.1 Direct selection of motile cells

Attempts were made to culture cells in a specific area, allow them to migrate and then extract the migrated live cells for further culture. This would have allowed the direct application of a selective pressure in the form of a resource gradient and direct selection of motile cells.

Within the scope of this project these techniques were unsuccessful. The low N of cells in some experiments and the fastidious growth of cancer cells made ongoing culture for selection experiments impossible.

Some success was gained using the transwell assay (table 5.2) which has been built upon and taken further within this lab group to successfully use collagen layers for migration and selection experiments.

Table 5.2: Methods used to try and physically separate motile cells in culture.

Method	Overview	Experimental Issues
Transwell assay	Chambers with 0.4 µm pores were suspended over a well[248]. The pores allowed cell migration. 0.5 % FBS media was added to the top chamber and 10 % FBS media to the bottom to test whether nutrient gradients select for cell motility.	The number of cells which migrated through to the bottom layer were too few for continued culture.
Hydrogels	Hydrogels made from collagen, gelatin or matrigel[247]. Cells were cultured onto these surfaces or homogenously mixed within the hydrogels. After migration hydrogels could be removed leaving motile cells in culture.	Hydrogel layers disintegrated upon agitation. Trypsin removal of cells which had migrated out of hydrogels was possible but numbers were too few for continued culture.
Cloning rings	A metal ring was placed into the centre of a well and cells cultured inside. Rings could be removed and replaced allowing migration and isolation.	Hydrostatic forces made it impossible to form a seal between the cloning ring and culture surface meaning cells could not be isolated for removal.
Wound healing inserts	These are silicone inserts with wells separated by a 500 µm gap. Cells were cultured in each of the wells the insert removed allowing migration and then replaced to isolate motile cells.	Hydrostatic forces prevented watertight seals from forming between the inserts and culture surface meaning extraction of a specific group of cells was not possible.
Chemotaxis chambers	Chemotaxis chambers contain two wells connected by a thin channel, cells were cultured in 0.5 % FBS media on one side and could migrate to the other well containing 10 % FBS media.	There were issues extracting cells from just one well as both wells are linked and the small number of cells extracted were too few for continued culture.

5.2.3.2 Time-lapse microscopy

HeLa, HT1080, MCF7-3 and MDA-2 cells were plated at 5000 cells per well onto a 24-well plate in 1 mL of the 10 % FBS media. Approximately 1 hour prior to being placed onto the Nikon TiE time-lapse system microscope all media was removed from all wells and replaced with either 10 % or 0.5 % FBS media. All wells were divided equally between each group and cell line and distributed randomly around the plate.

For each of three independent repeats, NIS-Elements photograph software captured images of three points within each well every 20 minutes for 72 hours. A total of 46,656 images were captured and NIS-Elements converted these into 216 video files. These were analysed using ImageJ and MtrackJ[97 98] to track 16,436 individual cells.

5.2.3.3 MTT assays

MTT assays provide an accurate measurement of live cell concentration and were used to quantify cell proliferation (for full protocol see section 2.4.2). HeLa, HT1080, MCF7-3 and MDA-2 cells were plated in 48 wells of a 96-well plate at 5000 cells per well in 200 µl of either 10 % or 0.5 % FBS media, 24 wells per group. 200 µl of each media was also added to 24 wells per plate as a control. Plates were left for 72 hours before performing the MTT assay. Three independent replicates were obtained for each cell line in both media.

5.2.3.4 Western blot protein analysis

Western blotting involves the electrophoretic separation of proteins on an acrylamide gel and their transfer onto a membrane to allow immunochemical visualisation, detection and quantification[259]. The solutions used for this method are outlined in table 5.3.

Cell lysates were prepared following the protocol outlined in Mahmood and Yang (2012) [259] and the total protein concentration calculated using a bradford protein assay and protocol (Bio-Rad) [260]. From each cell lysate the volume required to give 50 µg of protein as well as 2 µl of stain free protein ladder diluted 1:3 with ddH_2O[261] were loaded into separate wells of Bio-Rad pre-cast gels[262]. To separate the proteins, gels were run at 220 V for 30 minutes[259 262] following the Bio-Rad protocol[262].

Separated proteins were transferred from the gel to a membrane using a mini transfer pack on a Trans-blot Turbo Transfer System[263] (Bio-Rad) running a turbo 1 mini TGX programme and following the Bio-Rad protocol[263]. After transfer the membrane was placed onto an EZ image gel doc machine and ImageLab software was used to activate the membrane by enabling reactions between the proteins and trihalo compounds enabling visualisation and quantification of the total protein in the membrane (as seen in figure 5.1)[262 263]. The membrane is washed in TBST and blocked in 5 % BSA for 1 hour before being incubated in 5 mL 1:1000 dilution of primary antibody: 5 % BSA overnight at 4 °C[259 264]. After further TBST washes the membrane was incubated in 5 mL 1:1000 dilution of a HRP conjugated secondary antibody: 5 % BSA for 1 hour at room temperature[259 252 264]. The membrane was washed and a 1:1 mix of ECL reagent (PRIME) solutions applied for 1 minute. A Luminescent image analyser and Image Quant Las 4000 software took images and quantified membrane proteins, as seen in figure 5.2. These results were exported into an excel file and the ratio of each protein was calculated as the band intensity normalized to total protein[259 264 265].

Antibodies (from CellSignalling Technology) specific to four different groups of proteins were chosen based on availability and applicability to the wider project. The specific proteins tested for and the groups they belonged to were; Stress and Apoptosis: P-p53, C-Caspase 3, P-MAPKAPK-2, C-PARP, P-c-Jun, P-SAPK.JNK. HDACs: HDAC3, HDAC1, P-HDAC4, HDAC4, HDAC2, HDAC6. Focal Adhesions: Talin 1, Paxilin, Tensin 2, Vinculin, Arp 3, α-actinin. Actin Nucleation and Polymerisation: N-WASP, WAVE 2, Profilin-1, P-Rac 1, and GapDH.

Table 5.3: Solutions and components used in the western blot procedure.

Name	**Total Volume**	**Chemicals**
Lysis buffer	1 mL	10 µl protein inhibitory cocktail 990 µl RIPA buffer
5 x Tris glycine electrophoresis (5 x Running buffer)	Made up to 1 L with dH_20	15.1 g tris base 94 g glycine 50 mL 10 % SDS
1 x Tris glycine electrophoresis (1 x Running buffer)	Made up to 1 L with dH_20	200 mL of 5 x tris-glycine electrophoresis running buffer
10 % sodium dodecyl sulphate (SDS)	Made up to 500 mL with dH_20	50 g SDS
Transfer buffer	Made up to 1 L with dH_20	2.9 glycine 2.8 g tris base 200 mL methanol
10 x Tris Buffer Saline (TBS)	Made up to 1 L with dH_20	80 g sodium chloride 2 g potassium chloride 30 g tris base Adjusted to pH 7.4 with concentrated hydrochloric acid
1 x Tris Buffer Saline (TBS)	Made up to 1 L with dH_20	100 mL of 10 x TBS
Tris buffer Saline Tween-20 (TBST)	Made up to 1 L with 1 x TBS	1 mL tween-20
5 x SDS gel loading buffer	Made up to 40 mL with dH_20	1.514 g tris base 4 g SDS 20 g glycerol 10 g mercaptoethanol 5 mg Bromophenol Blue
10 % ammonium persulphate (APS)	Made up to 50 mL with dH_20	5 g APS
1 M Tris (pH 6.8)	Made up to 500 mL with dH_20	60.57 g tris base Adjusted to pH 6.8 using concentrated hydrochloric acid
1.5 M Tris (pH 8.8)	Made up to 500 mL with dH_20	90.86 g tris base Adjusted to pH 8.8 using concentrated hydrochloric acid
PBS	Made up to 1 L with dH_20	10 PBS tablets
5 % BSA	Made up to 50 mL with TBST	2.5 g powdered BSA

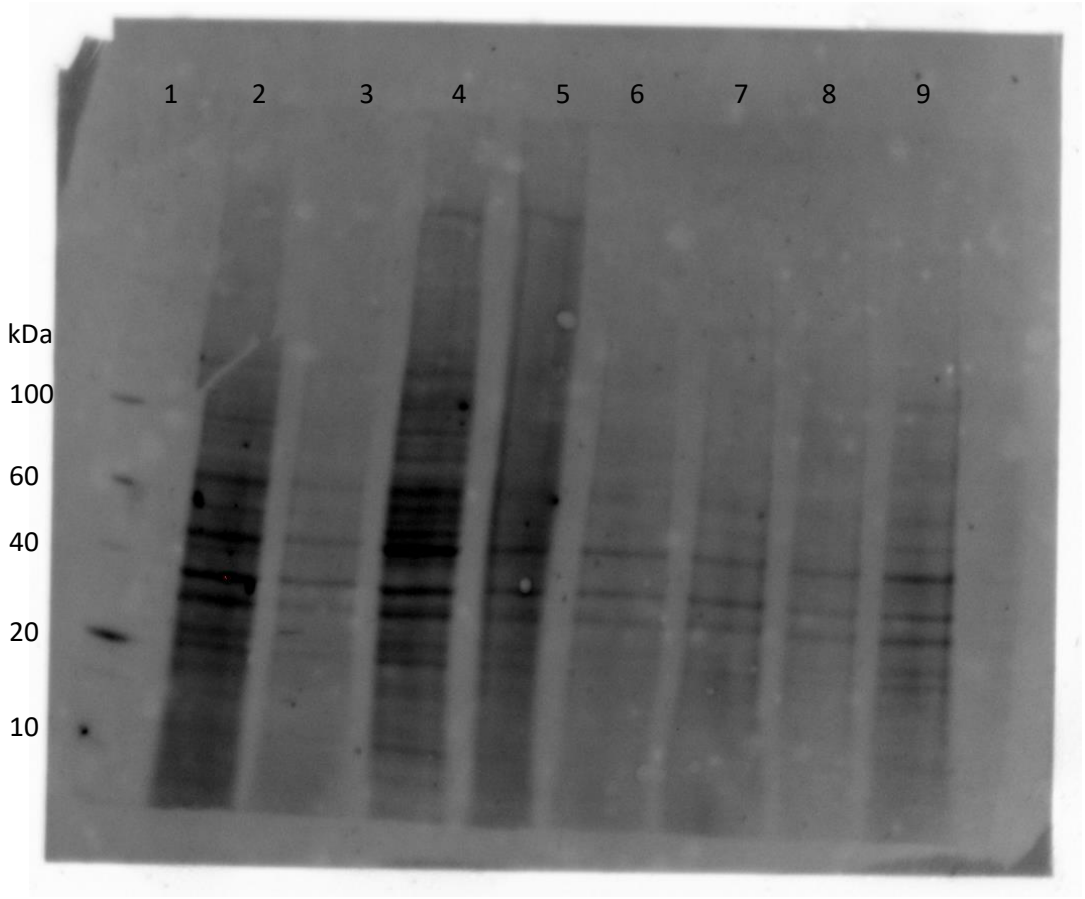


Figure 5.1: Image of total protein transferred onto membrane during western blot experiment. Image taken on an EZ image gel doc machine using ImageLab software. Lanes are labelled 1 to 9 and contain proteins; Lane 1: protein ladder, lane 2: HT1080 10 %, lane 3: HT1080 0.5 %, lane 4: HeLa 10 %, lane 5 HeLa 0.5 %. Lane 6: MCF7-3 10 %, lane 7: MCF7-3 0.5 %, lane 8: MDA-2 10 %, lane 9: MDA-2 0.5 %. The protein ladder molecular markers are noted (in kDa) down the left-hand side.

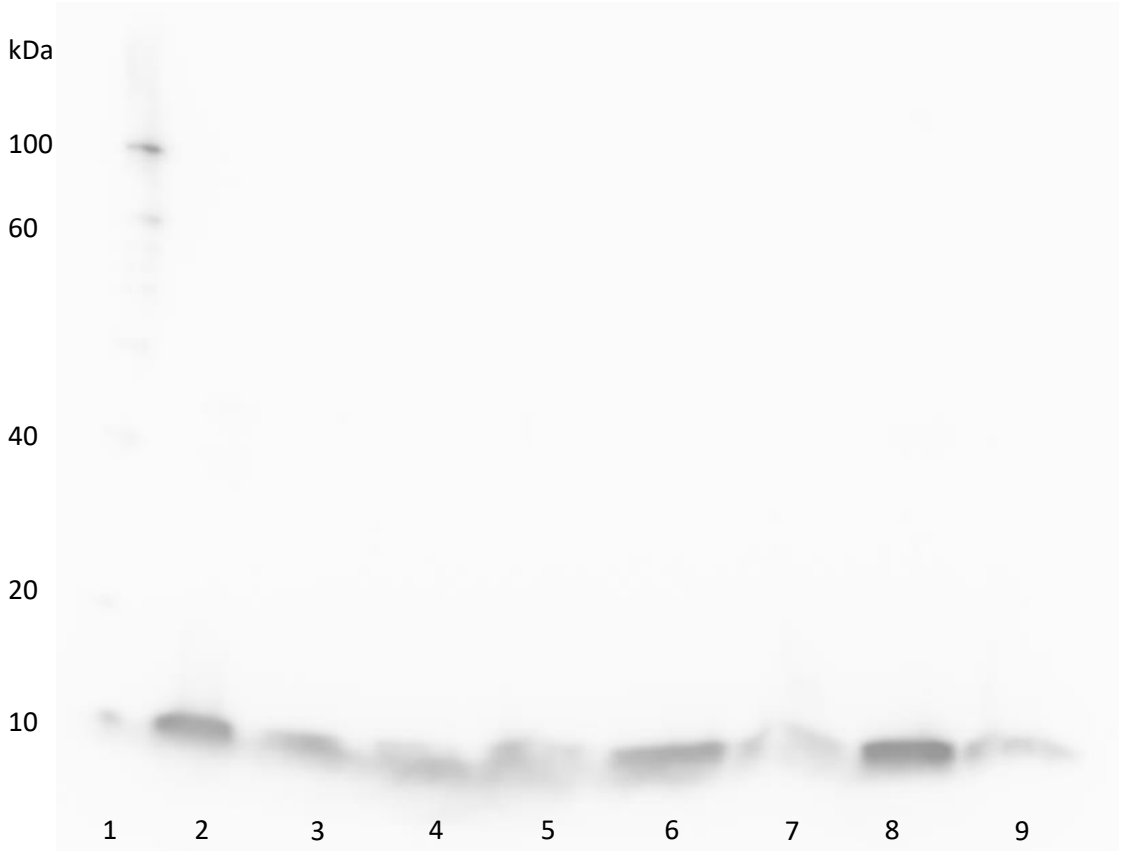


Figure 5.2: Western blot analysis of cell lysates from HeLa, HT1080, MCF7-3 and MDA-2 cells in 10 and 0.5 % media using Profilin-1 Rabbit mAb. Image taken on a Luminescent image analyser using Image Quant Las 4000 software. Lanes are labelled 1 to 9 and contain proteins; Lane 1: protein ladder, lane 2: HT1080 10 %, lane 3: HT1080 0.5 %, lane 4: HeLa 10 %, lane 5 HeLa 0.5 %. Lane 6: MCF7-3 10 %, lane 7: MCF7-3 0.5 %, lane 8: MDA-2 10 %, lane 9: MDA-2 0.5 %. The protein ladder molecular markers are noted (in kDa) down the left-hand side. The dark bands in lanes 2 – 9 are the stained profilin-1 protein.

5.2.4 Measuring adaptation of MCF7-1 cells to a 5 % and 0.5 % FBS nutrient environment after 12 weeks

MCF7-1 cells, acquired from ATCC, were cultured in 25 cm^2 culture flasks in 7 mL media containing either 10, 5 or 0.5 % FBS. MCF7-1 cells were maintained in low (0.5 % FBS) or high (5 % FBS) resource media for twelve weeks (passages 17-29). Three replicate lines were maintained for each treatment group with each line initiated 2 weeks after the previous one. A sample of MCF7-1 cultured in their original 10 % FBS media (passage 17) cells were cryogenically stored to act as an ancestral reference. Every 7 days cells approximately 25 % of the cell population (see table 2.1) were passaged into new cultures, at each transfer all remaining cells were cryogenically stored.

Table 5.4: MCF7-1 cell groups. The reference name and % FBS culture media of MCF7-1 cells split into multiple cultures.

Reference Name	**% FBS Media**	**Adaptation Time**	**Passage Number**
Ancestor	10 %	0 weeks	17
Evolved 5 %	5 %	12 weeks	29
Evolved 0.5 %	0.5 %	12 weeks	29

5.2.4.1 Time-lapse microscopy

The three groups of MCF7-1 cells (table 5.4), were plated at 2000 cells per well onto a 12-well plate in 2 mL of their respective media. All wells were divided equally between each group and distributed around the plate. Approximately 1 hour prior to being placed onto the Nikon TiE timelapse system microscope media was removed in all wells and fresh media of the same type added.

NIS-Elements photograph software captured images using a x10 objective at five points within each well every 15 minutes for 52 hours 15 minutes. This was repeated three times and a total of 37,620 images were captured. NIS-Elements converted these into 180 video files which were analysed using ImageJ and MtrackJ[97 98] and 1,060 individual cells tracked. The quality of the videos along with the smaller field of view due to using a x10 objective resulted in fewer cells being tracked.

5.2.4.2 Common garden assay

All three groups of MCF7-1 cells were plated onto a 12-well plate at 5000 cells per well in 2 mL of their respective media. The cells were then left for twenty-four hours to adhere. After this time all media was removed and replaced with either 5 or 0.5 % FBS media. The number of wells were divided equally between each group and each media resulting in four wells per cell group with two wells of each media.

The well-plate was placed on a Zeiss A1 inverted epifluorescent microscope at x4 objective. Seven points were chosen within each well and images taken of these points every 12 hours for 72 hours. The number of cells in each image was counted using ImageJ[97 98]. The number of cells in the images was counted giving the proliferation rate in both the adapted and the opposing media.

5.2.4.3 Spheroid spreading assays

Corning spheroid multi well plates and flasks (Merck) contain a covalently bound hydrogel layer that inhibits cellular attachment. To begin a spheroid culture, evolved MCF7-1 cells were passaged (section 2.3) from a monolayer culture and pipetted onto spheroid multi well plates. Cells were grown as spheroids for 24 hours in 2 mL of their respective media.

Spheroids were passaged onto standard adhesive 12-well plates very slowly using a wide tipped 1000 µl pipette to keep spheroids whole (6 wells for each group). Half the wells for each cell group had 2 mL of their adapted media added (the media they had been cultured in for 12 weeks), the other half had 2 mL of the opposing media added (the media the other groups of evolved MCF7-1 cells had been cultured in for 12 weeks). When returned to flasks with a suitable surface, spheroids will adhere and individual cells will disperse outwards, eventually forming a monolayer. The spheroids were left to attach to the plate for 4 hours then a Zeiss A1 inverted epifluorescent microscope was used to capture images of six spheroids within each well every six hours for 72 hours. It was not possible to measure the same spheroids between time points due to the magnification adjustments and manual tracking required to maintain spheroids within the field of view. Attempts were made to use a Nikon TiE time-lapse system microscope to observe spheroid spreading however it was not possible to make magnification adjustments over the course of the experiment and automatic stage movement, required for multipoint time-lapse images, caused spheroid detachment and movement.

ImageJ software[97 98] was used to calculate the area covered (in µm²) by cells dispersing from the spheroids and spheroids within wells were randomly paired between photos taken over time points and the change in spheroid area calculated as a percentage increase.

5.2.4.4 Competition assays

Competition assays involve co-culture of two cell lines or groups hence they require populations easily distinguishable from each other. Before competition assays can be conducted fluorescent probes that allow live cell imaging lasting over 48 hours without compromising other cell characteristics must be found. Within this project three distinct types of fluorescent tag were tested; CellLight, CellTracker and QTracker.

5.2.4.4.1 CellLight

CellLight reagents are fluorescent protein-signal peptide fusions that target specific cellular structures and can be used in live-cell imaging. CellLight Nuclear BacMam 2.0 GFP (Lifetechnologies C10602) accurately targets the cell nucleus, the provided protocol was followed[217].

MCF7-1 cells (passage 19-21 in 10 % FBS) were plated in a 6 well-plate in 2 mL media at a concentration of 1 x 10^5 cells per well and left to adhere for 24 hours before transfection. The formula below was used to calculate the volume of CellLight to add to each well. The desired particles per cell (PPC) were 35 (as recommended in the supplied protocol[217]).

$$Volume\ of\ CellLight = \frac{number\ of\ cells\ \times desired\ PPC}{1\ \times 10^8 CellLight\ particles/mL}$$

The media on the cells was removed, each well then had either 2 mL of CellLight media or 2 mL of standard media added, wells were divided equally between the treatment groups. The well-plate was then incubated for 16 hours then placed onto a Zeiss A1 inverted epifluorescent microscope and five points chosen at random in each well. Images of these points were captured every 12 hours for 72 hours and ImageJ[97 98] used to count the cell number. This experiment was repeated three times.

5.2.4.4.2 CellTracker

CellTracker Green CMFDA and Red CMTPX dyes (Lifetechnologies C7025 and C34552), are cytoplasmic fluorescent dyes[217]. The protocol from Lifetechnologies[217] was followed and a range of concentrations tested.

MCF7-1 cells (passage 20-24 in 10 % FBS) were plated in a 12 well-plate in 1 mL media at 2 x 10^4 cells per well and left to adhere for 7 hours. All media was removed and replaced with 1 mL media containing either no fluorescent tag, green CellTracker or red CellTracker to the specified concentrations. All wells were divided equally between treatment groups. The well-plate was incubated for 45 minutes and after this time the media-CellTracker solution was removed and replaced with standard media. The well-plate was kept in the dark at all times and the number of wells were distributed evenly between the groups.

The well-plate was placed on a Zeiss A1 inverted epifluorescent microscope and five points chosen at random in each well. Images of these points were captured every 12 hours for 72 hours. All images were taken twice, once with a fluorescent filter to allow imaging of the CellTracker and once with a light filter to allow imaging of all cells present in the field of view. All experiments were repeated three times and ImageJ[97 98] was used to count the cell number.

5.2.4.4.3 QTracker

QTracker (Life technologies Q25049) are a type of nanoparticle designed to fluorescently label cells[217]. QTracker cell labelling kits use target peptides to deliver Qdot nanocrystals into cell cytoplasms. The QTracker cell labelling kit was obtained from Thermo Fisher Scientific Inc. and their accompanying protocol followed and adapted as stated below[217].

2×10^4 MCF7-1 cells (passage 21-25 in 10 % FBS) were plated onto 35 mm glass culture dishes (ibidi) and incubated for 12 hours in 0.5 mL media. QTracker labelling solutions of 2, 5, 10 and 15 nM were prepared by mixing 1:1 appropriate volume of QTracker component A (2 µM nanocrystals in 50 mM borate buffer, pH 8.3) and component B (carrier, PBS pH 7.2). Solutions were incubated for 5 minutes in the dark at room temperature, the appropriate volume of media was added and the solution vortexed for 30 seconds.

0.2 mL QTracker media was added to each glass dish and the cells incubated at 37 °C in the dark for either 1 or 4 hours. The solution was removed and cells washed twice with 0.3 mL media. 0.5 mL media per dish was added and the cells placed onto a Zeiss A1 inverted

epifluorescent microscope to try and capture images of fluorescently labelled live cells. Despite various concentrations and incubation times being trialled no fluorescence was detected.

5.2.5 Measuring adaptation of MCF7-2 cells after 6 months culture under resource restriction

MCF7-2 cells, obtained from Public Health England, were cultured in 25 cm^2 culture flasks in 7 mL media. These cultures were maintained for six months (passages 23-48) with three replicate lines for each treatment group, each line initiated 2 weeks after the previous one. All the culture media are listed in table 5.5 along with the name used to refer to each group of MCF7-2 cells. Initially (passage 15) MCF7-2 cells were cultured only in the MEM and DMEM control media. This allowed adaptation to the lab conditions as well as increasing cell number. After adaptation to the control media these cultures were then split into the eight treatment groups (shown in table 5.5). Every 7 days cells approximately 25 % of the cell population (see table 2.1) were passaged into new cultures. Cryogenic samples of each culture were stored monthly and samples of MCF7-2 cells (at passage 23) in the control MEM and DMEM media were cryogenically stored to act as an ancestral reference.

All cultures started with $0.7x10^6$ cells. Cells cultured in the DMEM control media were split into the DMEM control, 0 mM Glucose and 1 mM Glucose media. Cells cultured in the MEM control media were split into the MEM control, 0.5 % FBS, 5 % FBS, pH 6.5 and pH 7.0 media. pH was lowered through addition of lactic acid which was added dropwise to media at 37 °C. Media was left in 5 % CO_2 for five minutes then a sterile pH meter used to determine pH.

Due to lab wide mycoplasma infection these cultures had to be destroyed and any results, unless gained prior to a clear contamination test, were not able to be used for analysis. Spheroid spreading assays were conducted but could not be used in the analysis, data for some experimental repeats have also had to be removed.

Table 5.5: MCF7-2 cell groups. Culture media and reference name of MCF7-2 cells split into multiple cultures

Reference name	Basal media	FBS (%)	Sodium pyruvate (mM)	Glutamine (mM)	Non-essential amino acid (%)	pH	Glucose (mM)
MEM Control	Non-phenol red MEM	10	1	2	1	7.0-7.5	5.5
5 % FBS	Non-phenol red MEM	5	1	2	1	7.0-7.5	5.5
0.5 % FBS	Non-phenol red MEM	0.5	1	2	1	7.0-7.5	5.5
pH 7	Non-phenol red MEM	10	1	2	1	7.0	5.5
pH 6.5	Non-phenol red MEM	10	1	2	1	6.5	5.5
DMEM Control	Non-phenol red DMEM	10	1	4	n/a	7.0-7.4	5.5
1 mM Glucose	Non-phenol red DMEM	10	1	4	n/a	7.0-7.4	1
0 mM Glucose	Non-phenol red DMEM	10	1	4	n/a	7.0-7.4	0

5.2.5.1 Time-lapse microscopy

Evolved MCF7-2 cells were plated at 1000 cells per well onto a 12-well plate in 2 mL of their respective media. Approximately 1 hour prior to starting the experiment all media was removed and replaced. All wells were divided equally between each group and distributed around the plate.

NIS-Elements photograph software captured images of four points in each well every 15 minutes for 47 hours 15 minutes. A total of 18,144 images were captured and NIS-Elements converted these into 96 video files. These were analysed using ImageJ and MtrackJ[97 98] to track 2,897 individual cells, N=2.

Large image time-lapse microscopy was attempted; where multiple photos are taken all bordering the original point allowing videos with a wider field of view at the original magnification (see section 3.2.2.1). This was done to obtain heritability data as well as more accurate motility however all attempts were unsuccessful.

5.2.5.2 MTT assays

MTT Assays (detailed in section 2.4.2) measure the live cell concentration. MCF7-2 cells were plated at 1 x 10^4 cells per well in 200 µl of their specified media, each group of MCF7-2 cells were plated in 24 wells and each plate also contained 24 wells with a control of media only. Three independent replicates were obtained for 0 mM Glucose, pH 7 and pH 6.5 and DMEM and MEM Control cells, two independent replicates for 1 mM Glucose cells and only one replicate for 5 % FBS cells.

5.2.5.3 Wound healing assay

Each group of MCF7-2 cells were cultured in three wells of a 6-well plate until they formed a confluent monolayer. A 20 µl pipette tip was used to scratch a mark down the centre of each well and the media changed to remove any free floating cells[237 238]. The well-plate was placed on a Zeiss A1 inverted epifluorescent microscope and images of several points along the wound captured over time. The wound size in each image was measured using ImageJ software[97 98] and the rate of wound closure calculated. The rate of wound closure is the percentage reduction in wound size measured over time. Reduction in wound size is calculated as a percentage to account for the differences in size of the initial wounds.

5.2.6 Cell tracking

A total of 444 videos were analysed using ImageJ and MtrackJ[97 98] cells were chosen at random and tracks follow the central position of each cell in each image[101]. Where cell number was very low, either due to experimental parameters or technical limitations all cells in the field of view were tracked. A total of 18,873 cells were tracked comprising of 1,487,517 points within those tracks.

A cell's generation (labelled numerically in ascending order from the first progenitor), fate (the outcome of the cell – divided, died, moved off screen or still present at the end of the experiment) and its relationship to other cell tracks were manually recorded. Manual recording of these characteristics allowed subsequent comparisons between cell groups to ensure there was no bias within the population of cells measured.

Three different cell:cell relationships could be used to calculate estimates of H^2. As stated previously (chapters 3 and 4) these were mother:daughter, progenitor cell to its first

generation clonal descendants. Sister:sister, two cells produced from the same progenitor during its division. Cousin:cousin, second generation clonal descendants from one progenitor.

5.2.7 Statistics

All statistical analysis was performed using Minitab[103] and IBM SPSS[102]. Throughout this project none of the data has undergone transformation[105-107].

5.2.7.1 Motility statistics

Cell motility is represented as each cell's average speed, calculated as the total distance of the path travelled (in μm) over its lifetime (in hours). The Lilliefors correction of the Kolmogorov-Smirnov test has been used in conjunction with descriptive statistics to determine the distribution of the data[104]. For all experiments all repeats have been statistically compared to ensure no difference either between wells within an experiment or between replicate experiments (Kruskal-Wallis $p>0.05$). The data from all repeats for each group have been grouped together for further analysis[130].

As in chapters 3 and 4 any differences in cell motility between cell fates or generations were tested for to ensure there was no bias between the samples of cells measured and H^2 has been estimated using the slope parameter of an ordinary least squares regression of trait values between related cells for the three different cell:cell relationships[104].

Outliers (identified using scatterplots of the standardised residuals and checked using Mahalanobis and Cooks distances) have been removed from the analysis of H^2 and both H^2 and R^2 values have been given in the results. R^2 values show how much of the variance in the dependent variable (cell speed) is explained by the model[104]. As discussed in the methods, section 2.10.6 outliers represent natural variation in the cell population, however their presence could be explained by bias in the data identified when comparing cell types (generations & fates). Where appropriate such groups have been removed from the analysis, details on whether groups have been removed and why are given in the appropriate results section.

5.2.7.3 Calculating cellular growth rates for common garden and competition assays

For both competition and common garden assays the change in cell number over time was calculated using images of cells in culture at set time intervals. For each position at each time

interval the number of cells in each image was counted. The difference in cell number between time points for each position within the culture was then calculated as a percentage change in cell number over time.

$$\frac{Change\ in\ cell\ number\ (\%)}{Time\ (hours)} = \frac{[(Cell\ number\ at\ Time\ 2\ - Cell\ number\ at\ Time\ 1) \times 100]}{Time\ 2 - Time\ 1}$$

The average percentage change over time was then compared between groups. Calculating the growth rate as a percentage change in time takes account of any differences in cell number at the beginning of the experiments and allows accurate comparison between groups.

5.3 Results

5.3.1 Effect of serum reduction on HeLa, HT1080, MCF7-3 and MDA-2 cells

Cell lines were cultured for 72 hours in 10 % FBS and 0.5 % FBS nutrient media and multiple cell phenotypes measured. Reducing the amount of serum was predicted to increase competition between cells selecting for increased dispersal.

5.3.1.1 Motility of HeLa, HT1080, MCF7-3 and MDA-2 cells over 72 hours culture in 10 % and 0.5 % FBS media

As seen in figure 5.3 there was no significant difference in speed between the different experimental repeats for HeLa, HT1080, MCF7-3 and MDA-2 cell lines in either 10 or 0.5 % media (Kruskal-Wallis test; $p>0.06$). As a result, all data for each repeat has been grouped together for further analysis.

Time-lapse microscopy was used to track cell motility. It was predicted that increased competition would select for increased motility over time. As can be seen in table 5.6 this is true for HT1080 cells which have higher motility in the 0.5 % than the 10 % media (Mann-Whitney U test; $p<0.02$) but not in the other cell lines. MCF7-3 and MDA-2 cells both have higher cell motility in 10 % than 0.5 % FBS media (Mann-Whitney U test; $p<0.001$) and there is no significant difference in motility between the groups for HeLa cells (Mann-Whitney U test; $p=0.358$).

Figure 5.4 shows the median value and distribution of motility in both media for each cell line. All cell lines had a non-normal distribution of cell speeds (Lilliefor corrected Kolmogorov-Smirnov test; $p<0.001$) in both 10 % and 0. 5 % media.

Table 5.6: Median speed and N of HeLa, HT1080, MCF7-3 and MDA-2 cells in 10 and 0.5 % media.

Cell line	**Median Speed [N]**	
	10 %	**0.5 %**
HeLa	11.75 [2467]	11.89 [1858]
HT1080	34.05 [2562]	34.98 [2093]
MCF7-3	9.76 [1322]	8.21 [1977]
MDA-2	24.28 [2452]	16.24 [1224]

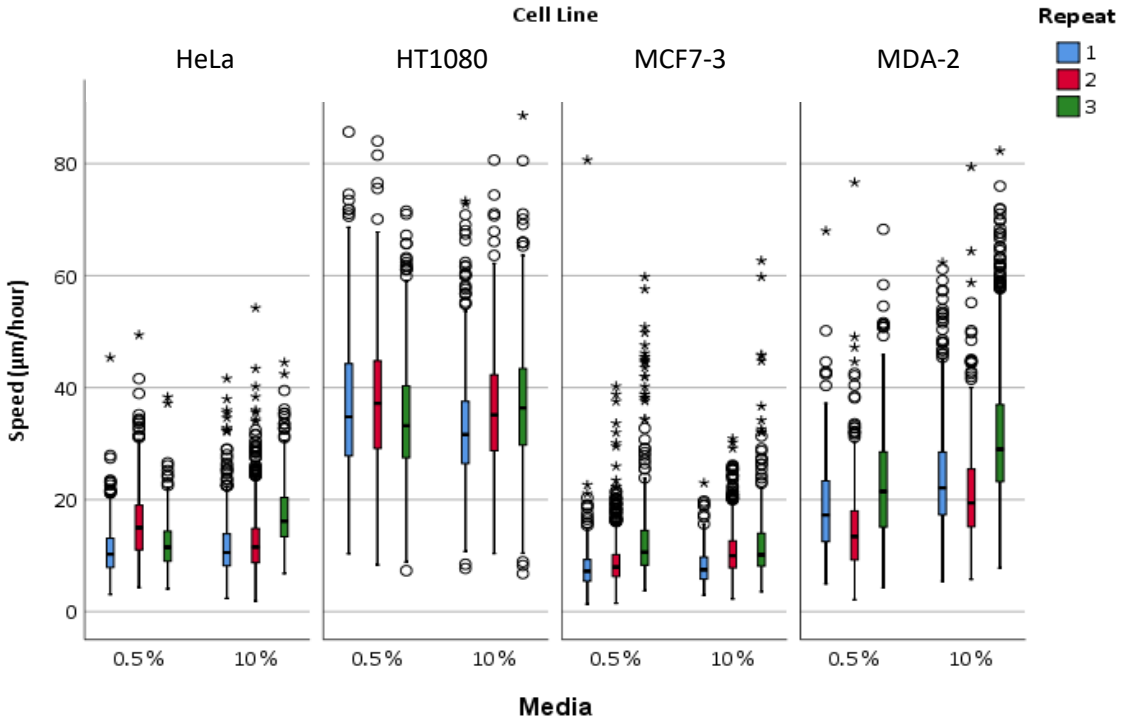


*Figure 5.3: Distribution of HeLa, HT1080, MCF7-3 and MDA-2 cell speeds in 10 and 0.5 % FBS media for each experimental repeat. The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. Data gathered from three independent experiments. N for each cell line, media and repeat is; HeLa 0.5 %; 1 = 823, 2 = 681 and 3 = 354. HeLa 10 %; 1 = 1125, 2 = 990 and 3 = 352. HT1080 0.5 %; 1 = 592, 2 = 587 and 3 = 914. HT1080 10 %; 1 = 947, 2 = 810 and 3 = 805. MCF7-3 0.5 %; 1 = 424, 2 = 1190 and 3 = 364. MCF7-3 10 %; 1 = 162, 2 = 969 and 3 = 191. MDA-2 0.5 %; 1 = 426, 2 = 484 and 3 = 314. MDA-2 10 %; 1 = 668, 2 = 714 and 3 = 1070.*

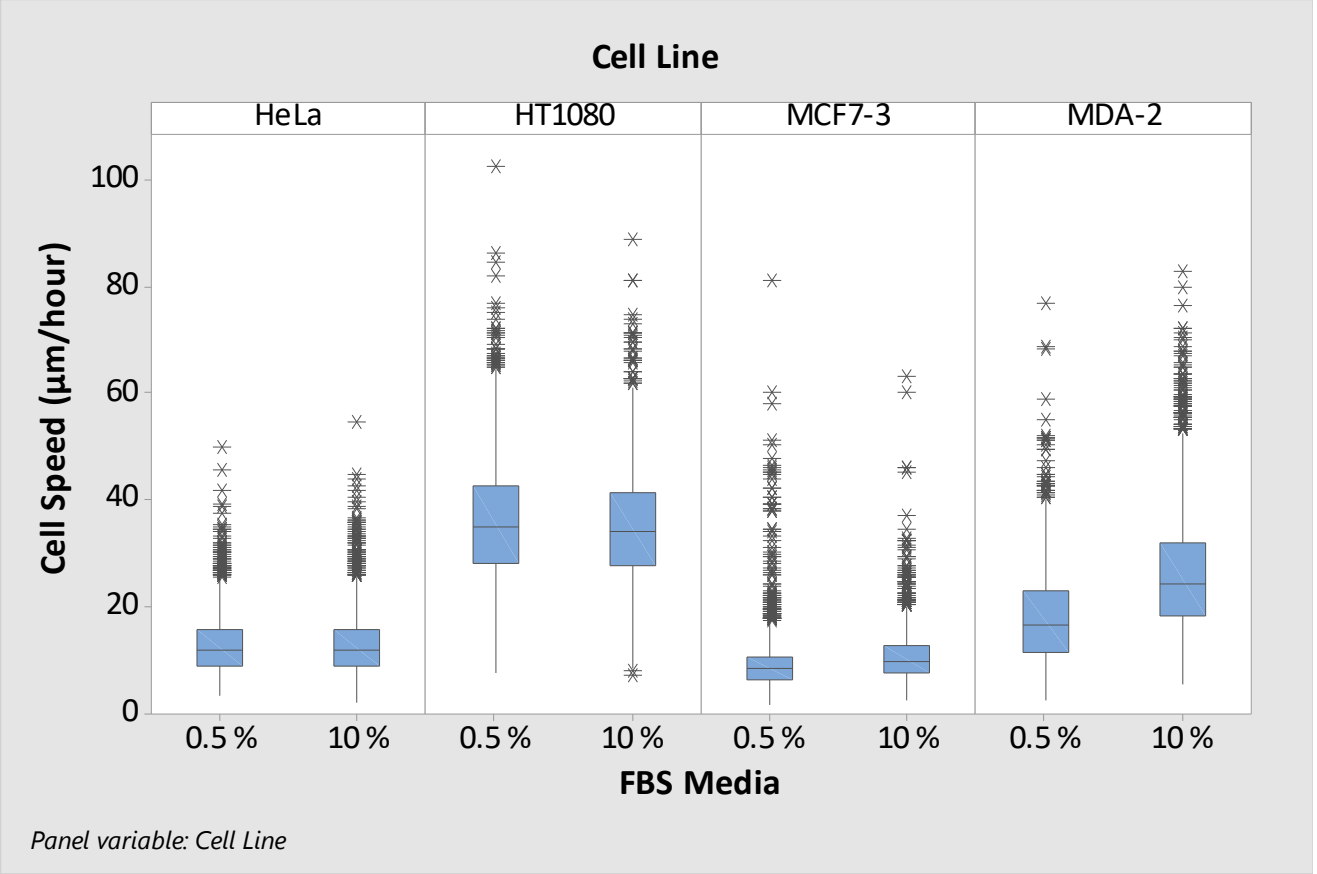


*Figure 5.4: Distribution of cell speeds in 0.5 % and 10 % FBS media for HeLa, HT1080, MCF7-3 and MDA-2 cell lines. The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. Data was gathered from three independent experiments. The N for each cell line is; HeLa 4,325, HT1080 4,655, MCF7-3 3,300 and MDA-2 3,676.*

As shown in table 5.7, generation 1 of MCF7-3 and MDA-2 cells in 0.5 % FBS media had significantly higher motility than generations 2 and 3 (Kruskal-Wallis test; $p<0.001$).

Table 5.7: Median speed (µm/hour) and N for each of the first three generations in HeLa, HT1080, MCF7-3 and MDA-2 cell lines for both 10 and 0.5 % media.

Cell line	**FBS media (%)**	**Median speed [N]**		
		Generation		
		1	**2**	**3**
HeLa	0.5	11.93 [206]	12.39 [367]	11.35 [588]
	10	13.17 [272]	12.61[495]	11.47 [807]
HT1080	0.5	34.84 [390]	35.20 [659]	36.22 [705]
	10	37.94 [236]	37.43 [373]	36.33 [523]
MCF7-3	0.5	9.05 [476]	7.95 [665]	7.85 [600]
	10	9.26 [269]	9.46 [359]	9.58 [376]
MDA-2	0.5	17.87 [639]	15.42 [443]	14.49 [150]
	10	24.83 [272]	25.06 [453]	23.61 [659]

As shown in table 5.8 all cell lines in both media showed significant differences in cell speed between cell fates (Kruskal-Wallis $p<0.001$). For HeLa, MCF7-3 and MDA-2 cells the differences in motility between cell fates were the same for both 10 % and 0.5 % FBS media, cells that died (HeLa) or moved off screen (MCF7-3 and MDA-2) had significantly higher speeds (Mann-Whitney U test, $p<0.001$). However HT1080 cells showed opposing results between the two groups, in 10 % media cells that moved off screen were significantly faster than cells that did not divide however in 0.5 % media those cells that did not divide were significantly faster than those that moved off screen (Mann-Whitney U test, $p<0.001$).

By excluding the first generation of cells and only using relationships between their clonal descendants to calculate the broad-sense heritability, the effect on estimates of H^2 caused by variation between generations and cell fates will be minimised.

Table 5.8: Median speed (µm/hour) and N for each cell fate in HeLa, HT1080, MCF7-3 and MDA-2 cell lines for both 10 and 0.5 % media.

Cell line	**FBS media (%)**	**Median speed [N]**			
		Fate			
		Non-divided	**Divided**	**Moved off screen**	**Died**
HeLa	0.5	11.58 [910]	11.60 [848]	15.19 [130]	12.90 [21]
	10	11.42 [1267]	11.73 [1136]	15.32 [133]	12.69 [11]
HT1080	0.5	36.47 [629]	34.97 [874]	32.65 [69]	33.67 [555]
	10	29.71 [515]	34.76 [1223]	29.36 [15]	35.85 [921]
MCF7-3	0.5	7.831 [1164]	8.638 [774]	9.101 [75]	13.62 [12]
	10	9.62 [754]	9.98 [548]	8.75 [52]	24.29 [16]
MDA-2	0.5	14.94 [837]	18.56 [301]	19.37 [53]	25.13 [51]
	10	22.96 [1208]	25.46 [1131]	23.21 [26]	27.47 [175]

5.3.1.2 Heritability of motility in HeLa, HT1080, MCF7-3 and MDA-2 cells over 72 hours culture in 10 % and 0.5 % FBS media

Cell speed is significantly heritable for all cell:cell relationships in both 0.5 and 10 % FBS media. Table 5.9 shows H^2 values range from 0.796 - 0.295 and R^2 values range from 0.09 - 0.63. Figure 5.5 shows how H^2 has been calculated for the mother:daughter relationship in both 0.5 and 10 % media of the HeLa cell line.

In the HeLa, HT1080 and MCF7-3 cell lines there appears to be a general trend for cells in 0.5 % media to have higher H^2 estimates than those in 10 % media. This is not true in the MDA-2 cell line or for HT1080 sister:sister and MCF7-3 cousin:cousin relationships. For all groups the H^2 estimates for the cousin:cousin relationship whilst significant are lower than those obtained for both mother:daughter and sister:sister. The number of post hoc comparisons meant it was not statistically possible to compare H^2 estimates. R^2 is the amount of variance in cell motility explained by the media group. R^2 values appears to correlate with the results seen for H^2 estimates, groups with higher H^2 estimates also have higher R^2 values. Inclusion of multiple different cell:cell relationships and examination over several generations means the significant results seen here represent stable differences between clonal cell lineages.

Table 5.9: Mean trait value, R^2 and broad-sense heritability of motility in HeLa, HT1080, MCF7-3 and MDA-2 cell lines in 0.5 % or 10 % FBS media for three cell:cell relationships.

Cell line	% FBS media	No of families	Mean speed (µm/hour)	Heritability of cell motility H^2 (R^2)		
				Sister:sister	Mother:daughter	Cousin:cousin
HeLa	10	296	12.55	0.707 (0.50)***	0.706 (0.5)***	0.647 (0.42)***
	0.5	285	12.64	0.796 (0.63)***	0.733 (0.54)***	0.724 (0.53)***
HT1080	10	248	37.25	0.545 (0.3)***	0.558 (0.31)***	0.364 (0.13)***
	0.5	295	37.55	0.593 (0.35)***	0.518 (0.27)***	0.504 (0.25)***
MCF7-3	10	185	9.96	0.479 (0.23)***	0.446 (0.2)***	0.350 (0.12)***
	0.5	163	8.99	0.512 (0.26)***	0.477 (0.23)***	0.295 (0.09)***
MDA-2	10	201	26.05	0.735 (0.54)***	0.770 (0.59)***	0.540 (0.29)***
	0.5	72	16.06	0.619 (0.38)***	0.475 (0.23)***	0.322 (0.10)**

**p<0.01, ***p<0.001

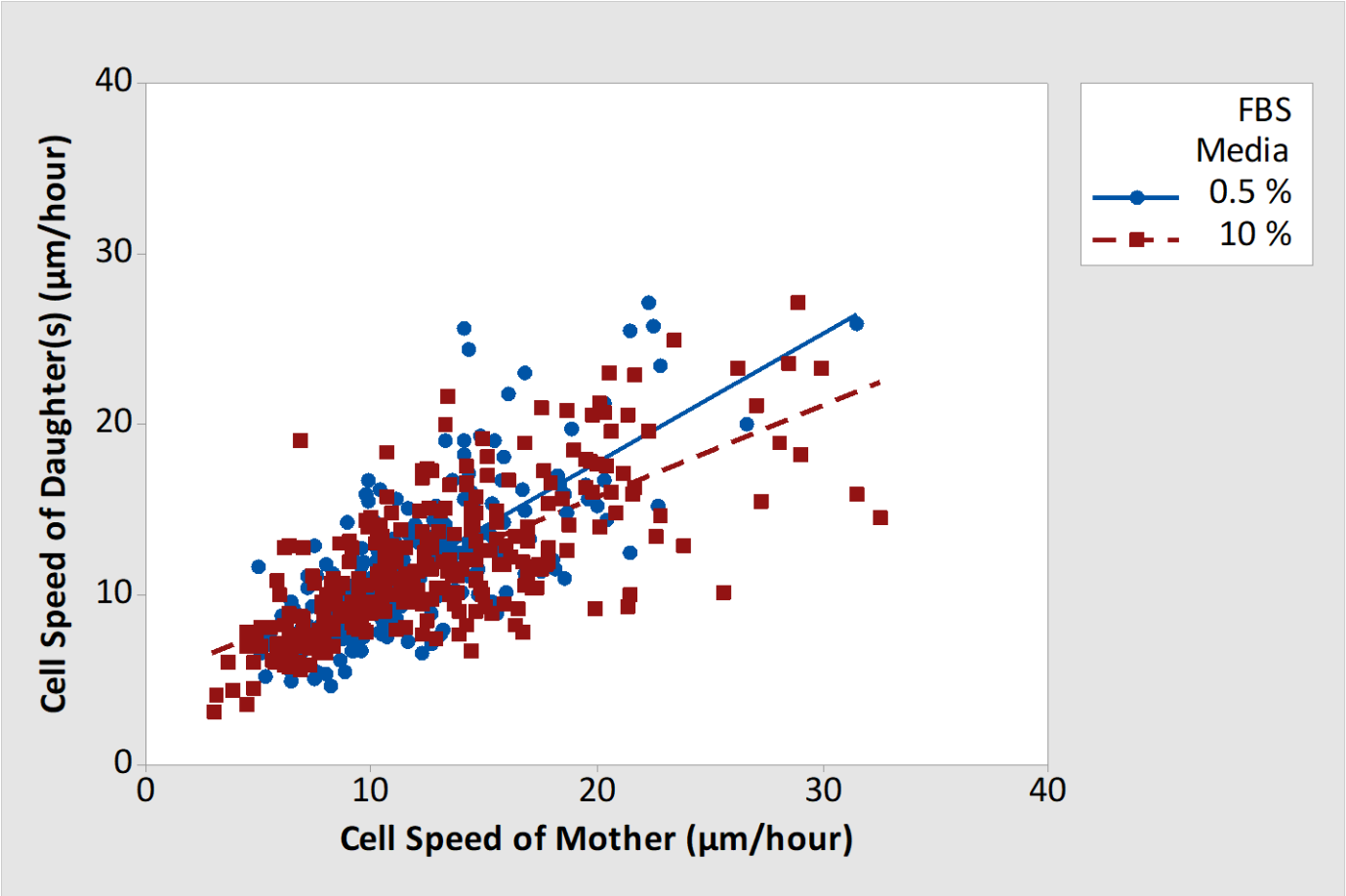


Figure 5.5: Example of the mother:daughter regression for motility of HeLa cells in 0.5 % and 10 % FBS media. Each point represents a mother cell and the mean speed of her daughters, the slope of the regression line is the estimate of broad-sense heritability. As the cells are clonal, the same method is applicable to other cell:cell relationships.

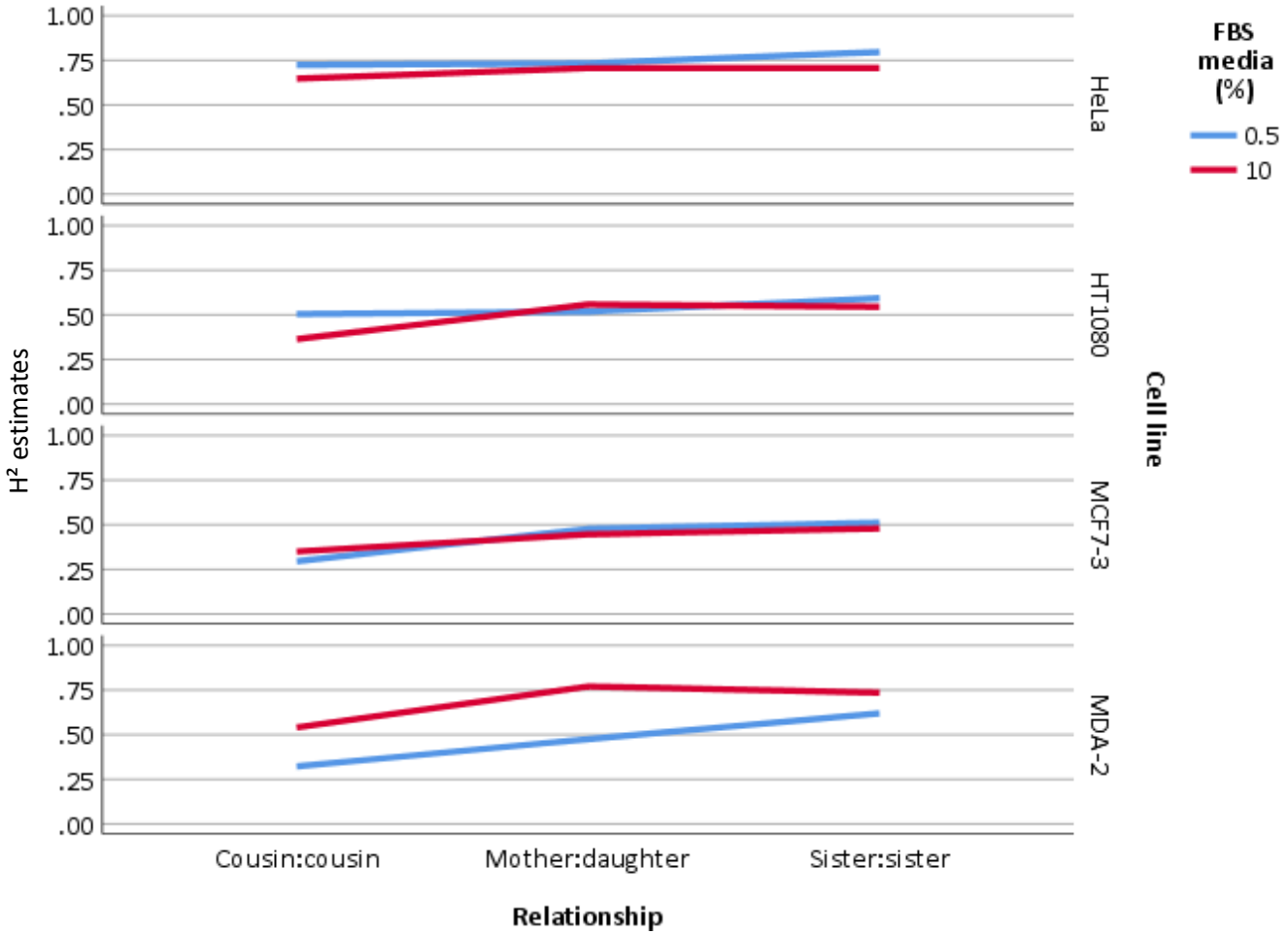


Figure 5.6: Heritability estimates of cell:cell relationships for HeLa, HT1080, MCF7-3, MCF7-2, MDA-1 and MDA-2 cell lines in 0.5 and 10 % media. N for each group is HeLa 10 %= 296, HeLa 0.5 % = 285, HT1080 10 % = 248, HT1080 0.5 % = 295, MCF7-3 10 % = 185, MCF7- 3 0.5 % = 163, MDA-2 10 % = 201, MDA-2 0.5 % = 72.

5.3.1.3 Live cell concentration of HeLa, HT1080, MCF7-3 and MDA-2 cells over 72 hours culture in 10 % and 0.5 % FBS media

MTT assays were used to determine the live cell concentration, a marker for the amount of cell proliferation or death. Reducing the amount of serum was predicted to increase competition for nutrients and decrease proliferation.

It appears that reducing the FBS concentration did not significantly alter the rate of proliferation in MCF7-3 and MDA-2 cells but did in HeLa and HT1080 cells. Figure 5.7 shows the median value and distribution of live cell concentration in both media for each cell line. HeLa and HT1080 cell lines had a significantly higher live cell concentration in 10 % media than 0.5 % media (Mann-Whitney U test, $p<0.001$). There was no difference in live cell concentration between the two media for MCF7-3 or MDA-2 cells (Mann-Whitney U test, Bonferroni correction $p>0.02$). An increase in the rate of cell death in 0.5 % FBS media can be discounted due to the comparison of cell fates (section 5.3.1.1) between groups. All cell lines had a non-normal distribution of cell concentration (Lilliefor corrected Kolmogorov-Smirnov test; $p<0.001$) in both 10 % and 0. 5% media.

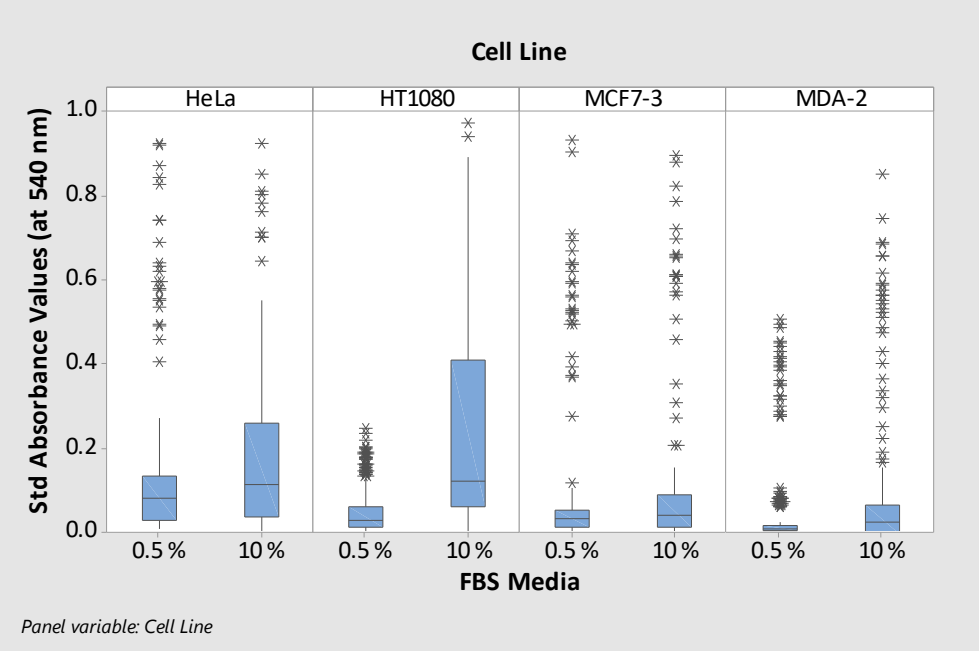


Figure 5.7: Distribution of live cell concentration in 0.5 % and 10 % FBS for HeLa, HT1080, MCF7-3 and MDA-2 cell lines. The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. Data was gathered from ten independent experiments.

5.3.1.4 Protein concentration of HeLa, HT1080, MCF7-3 and MDA-2 cells after 72 hours culture in 10 % and 0.5 % FBS media

Western blots were performed on lysates made after cells had been cultured for 72 hours in either 10 or 0.5 % FBS media. It appears that HT1080, MCF7-3 and MDA-2 cell lines show similar protein levels in both 10 % and 0.5 % FBS media however HeLa cells seem to show lower protein levels in 0.5 % compared to 10 % media. Figure 5.8 shows the difference in the ratio of each protein (relative to the total protein) between media groups for each cell line.

The groups of proteins tested; stress & apoptosis, HDACs, focal adhesions and actin nucleation & polymerisation, were all chosen for their links to motility and their role in cell adaptation to a resource restricted environment. As seen in chapter 4, research into the role of specific epigenetic inhibitors did not yield conclusive results. It was hoped that comparison of a broad range of HDAC proteins would elucidate any alterations in expression occurring at a group level, i.e. if individual HDAC proteins were not significantly increasing or decreasing in expression but overall HDAC protein expression did increase/decrease significantly. Such trends could be missed if only studying single proteins individually. Cells in 0.5 % FBS media face increased competition for resources which would be expected to increase stress and potentially increase the apoptosis rate[266 267]. Only the HeLa cell line showed a significant difference in both the proliferative rate and protein expression between the medias. However, cells in 0.5 % media seemed to have lower expression of stress and apoptosis proteins.

Focal adhesion and actin nucleation and polymerisation proteins are linked to motility[268-270]. It was predicted that expression of these proteins would be increased in cell groups which showed an increase in motility[268-270]. Actin nucleation and polymerisation proteins may also show increased expression correlating to increased proliferative rates[270]. Neither of these correlations was observed in these results. Groups of cells that had increased motility (HT1080 cells in 0.5 %, MCF7-3 and MDA-2 cells in 10 % media) or proliferation (HeLa and HT1080 cells in 10 % media) did not show any alteration in protein expression. The lack of correlation between protein expression and cell phenotypes could be due to the low N (=2) of the western blot experiments.

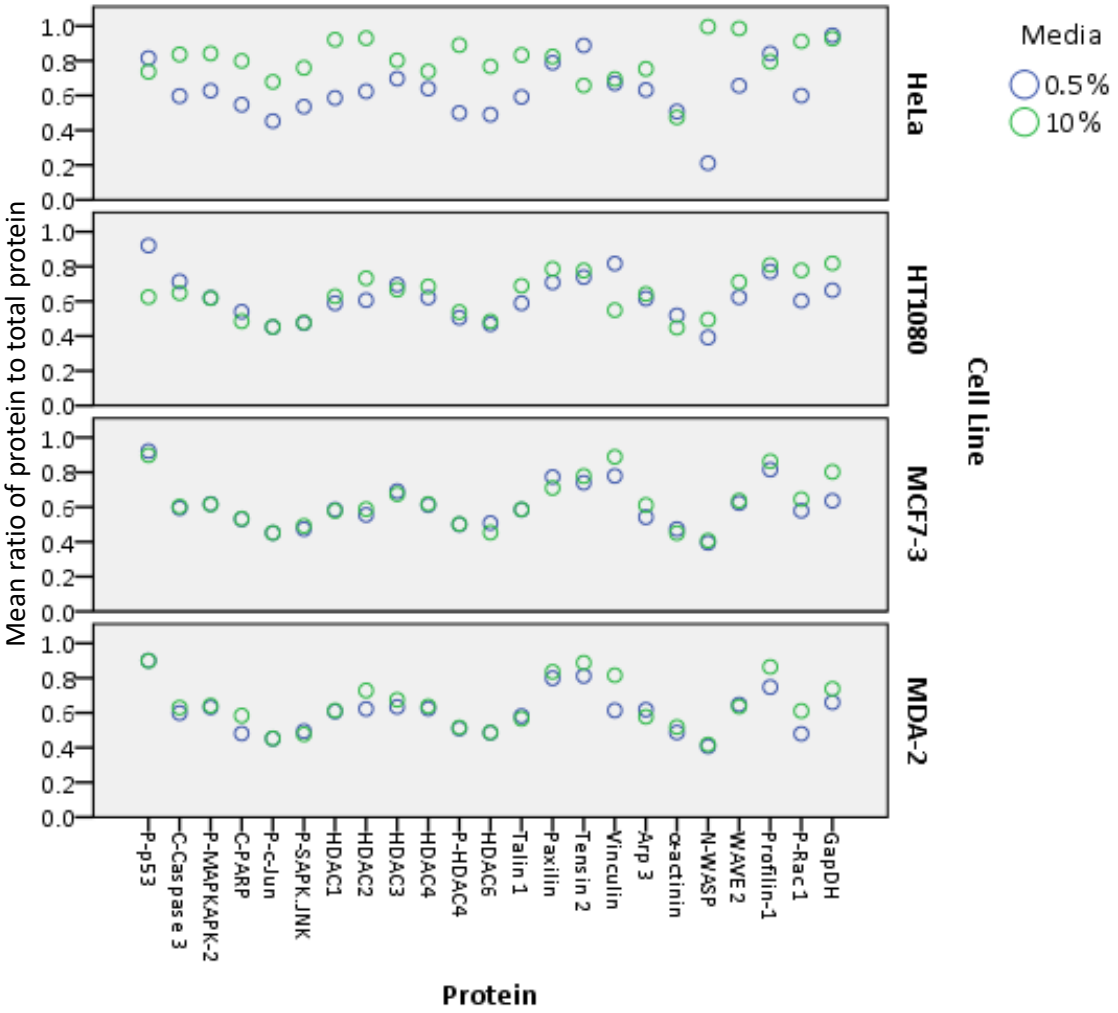


Figure 5.8: Mean ratio of proteins in 0.5 % and 10 % FBS media for HeLa, HT1080, MCF7-3 and MDA-2 cell lines. Data gathered from two independent experiments.

5.3.2 Effect of serum reduction on MCF7-1 cells

MCF7- 1 cells were cultured for 12 weeks in a 5 % and 0.5 % FBS nutrient media and multiple cell phenotypes measured. Evolving cells in a resource restricted environment was predicted to cause adaptation to competition, particularly an increase in population motility over time.

5.3.2.1 Motility of MCF7-1 cells evolved for 12 weeks in 5 % and 0.5 % FBS media

As seen in figure 5.9 there was no significant difference in speed between the different experimental repeats for the ancestral or evolved cell lines (Kruskal-Wallis test; $p>0.05$). As a result, all data for each repeat has been grouped together for further analysis.

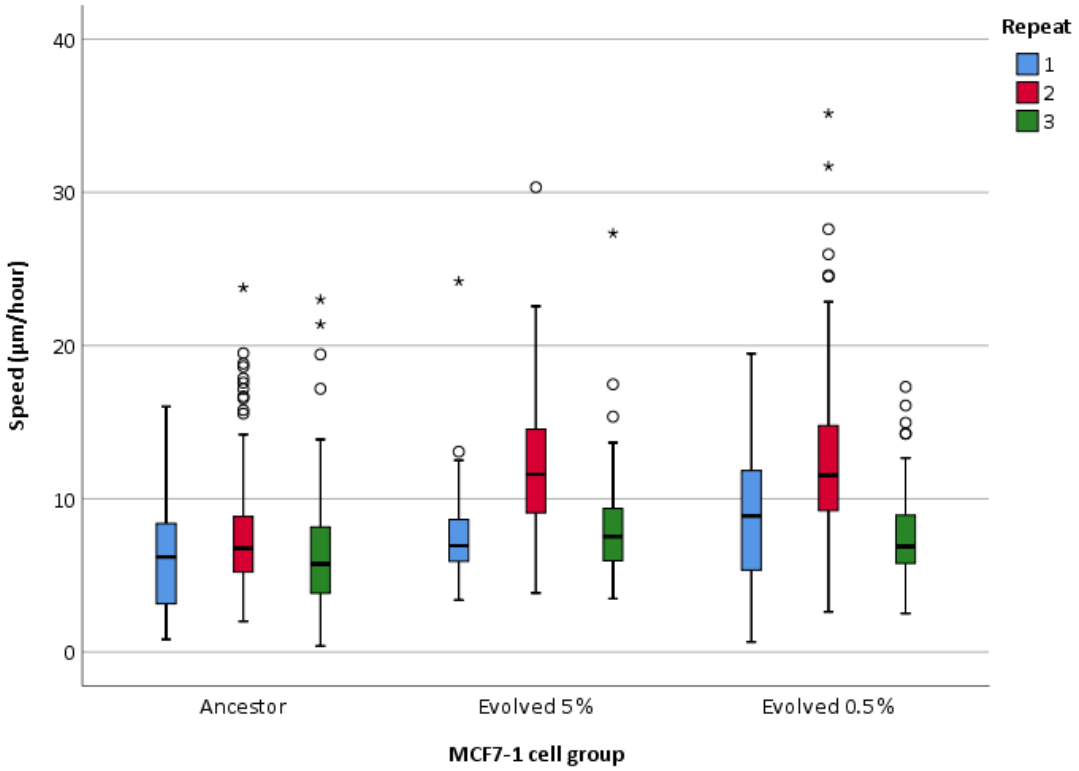


*Figure 5.9: Distribution of MCF7-1 ancestor and evolved cell speeds for each experimental repeat. The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. Data gathered from three independent experiments. N for each cell line and repeat is; Ancestor; 1 = 69, 2 = 192, 3 = 152. Evolved 5 %; 1 = 89, 2 = 84, 3 = 54. Evolved 0.5 %; 1 = 117, 2 = 123, 3 = 137.*

MCF7-1 cells evolved for 12 weeks in a resource restricted environment were predicted to have increased motility. As seen in table 5.10, the results support this hypothesis, there is a significant difference in cell speed between the ancestor population and both groups of evolved cells, with evolved cells having significantly faster motility (Mann-Whitney U test, $p<0.001$). There was no significant difference in speed in the two evolved populations (Mann-Whitney U test, $p=0.543$). The distribution of cell speeds in each group can be seen in figure 5.10, all groups had non-normal distributions of cell speed (Lilliefor corrected Kolmogorov-Smirnov test; $p<0.001$).

Table 5.10: Median speed and N of ancestor and evolved MCF7-1 cells.

MCF7-1 cell group	Median Speed (µm/hour)	N
Ancestor	6.34	413
Evolved 5 %	8.38	228
Evolved 0.5 %	8.91	377

No significant difference was found in cell speed between generations in the Evolved 5 % or the ancestral population (Kruskal-Wallis $p=0.054$). In the Evolved 0.5 % group the third generation was found to have a significantly higher speed than the first and second generations (Mann-Whitney U test, $p<0.001$). As can be seen in table 5.11, the N for generation 3 was very small, H^2 estimates have been calculated using the first generation of cells and their clonal descendants and H^2estimates cannot be calculated for the cousin:cousin relationship. No significant differences were found in cell speed between the different cell fates for any group (Kruskal-Wallis, $p>0.114$). As seen in table 5.12 only two types of cell fate were observed for this experiment. The lack of other cell fates could be due to the smaller time frame (in comparison to previous experiments) that this experiment was conducted over.

Table 5.11: Median speed and N for each of the first three generations of the ancestral and evolved MCF7-1 cells.

MCF7-1 cell group	Median Speed (µm/hour) [N]		
	Generation 1	Generation 2	Generation 3
Ancestor	5.49 [234]	7.09 [176]	8.57 [4]
Evolved 5 %	8.30 [114]	8.57 [154]	10.32 [116]
Evolved 0.5 %	7.39 [89]	8.32 [103]	8.09 [52]

Table 5.12: Median speed and N for each cell fate of the ancestral and evolved MCF7-1 cells.

MCF7-1 cell group	Median Speed (µm/hour) [N]	
	Non-divided	Divided
Ancestor	6.07 [324]	7.98 [90]
Evolved 5 %	8.84 [256]	9.13 [142]
Evolved 0.5 %	8.38 [164]	8.09 [80]

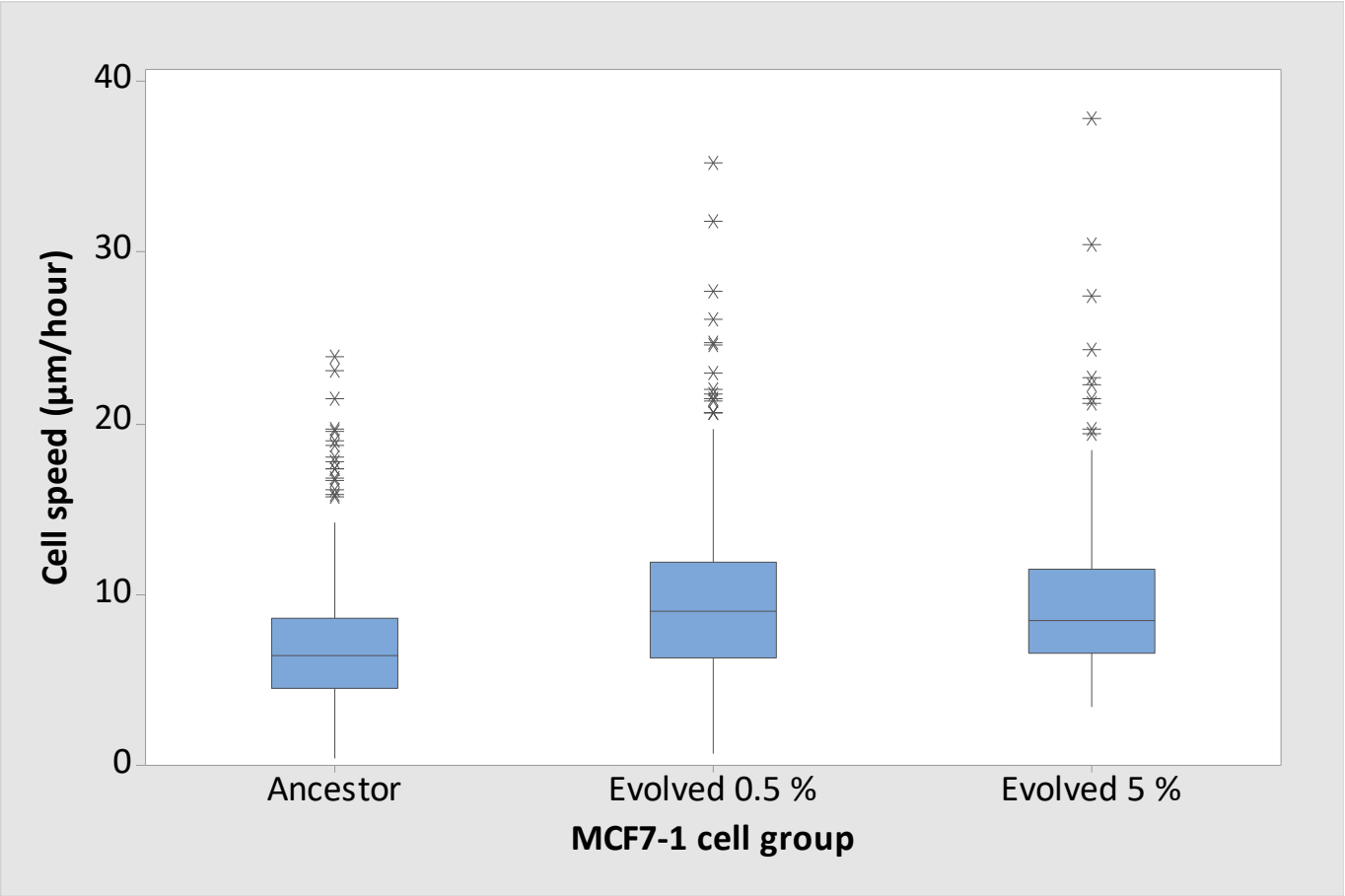


Figure 5.10: Distribution of cell speed for MCF7-1 cells evolved over 12 weeks in 0.5 % and 5 % FBS media and the ancestral population. The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. Data was gathered from three independent experiments. The N for each group is; Ancestor 413, Evolved 0.5 % 377 and Evolved 5 % 228.

5.3.2.2 Distance to point of MCF7-1 cells evolved for 12 weeks in 5 % and 0.5 % FBS media

The distance to point (D2P) of a cell is the distance a single cell moves between time frames. Measuring this over the course of an experiment can indicate how the rate of movement changes over time. As with motility there is a significant difference in mean D2P between the ancestor and both evolved groups with both evolved groups having significantly higher average D2P than the ancestor (Mann-Whitney U test, $p<0.001$). There was no significant difference in D2P of the evolved groups (Mann-Whitney U test, $p=0.347$). Figure 5.11 shows the distribution of mean D2P values for each cell group, the distribution of mean D2P for each cell group was non-normal (Lilliefor corrected Kolmogorov-Smirnov test; $p<0.001$). As seen in figure 5.12, when comparing D2P for each group over time, cells evolved in 0.5 % FBS have fewer non-motile cells than either the ancestor cells or those evolved in 5 % FBS (Kruskal-wallis $p<0.05$).

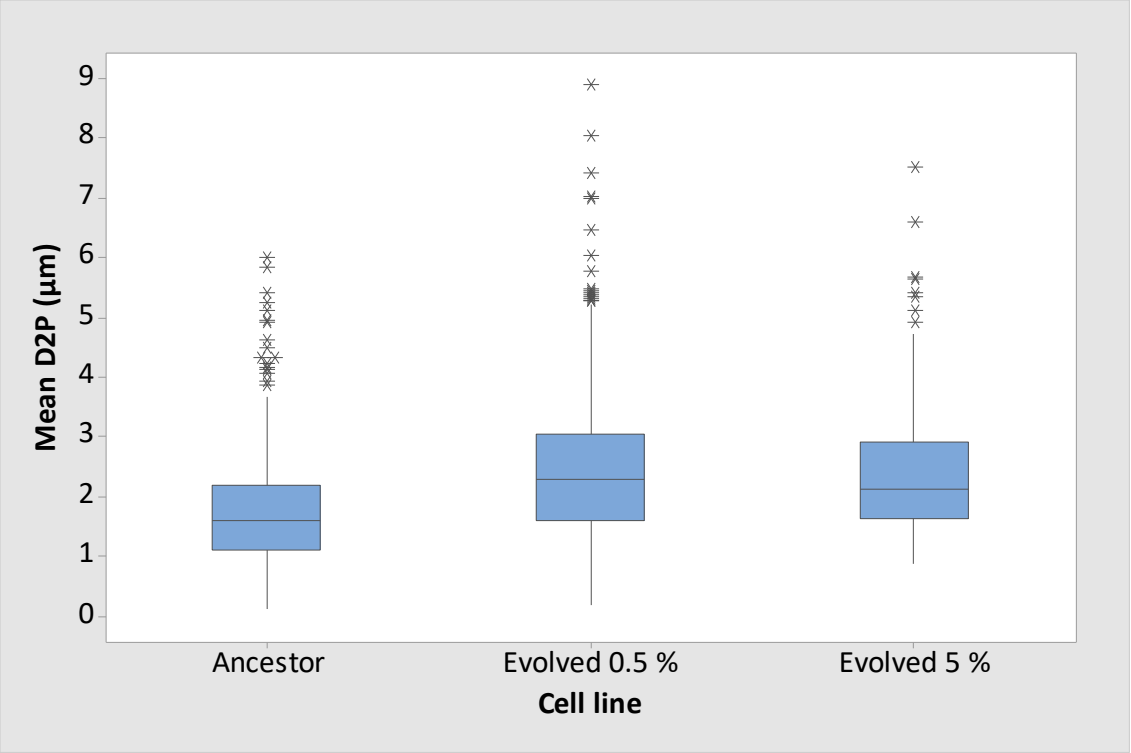


*Figure 5.11: Distribution of distance to point (D2P) measurements for MCF7-1 cells evolved over 12 weeks in 0.5 % and 5 % FBS media and the ancestral population. The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. Data was gathered from three independent experiments. The N for each group is; Ancestor 413, Evolved 0.5 % 377 and Evolved 5 % 228.*

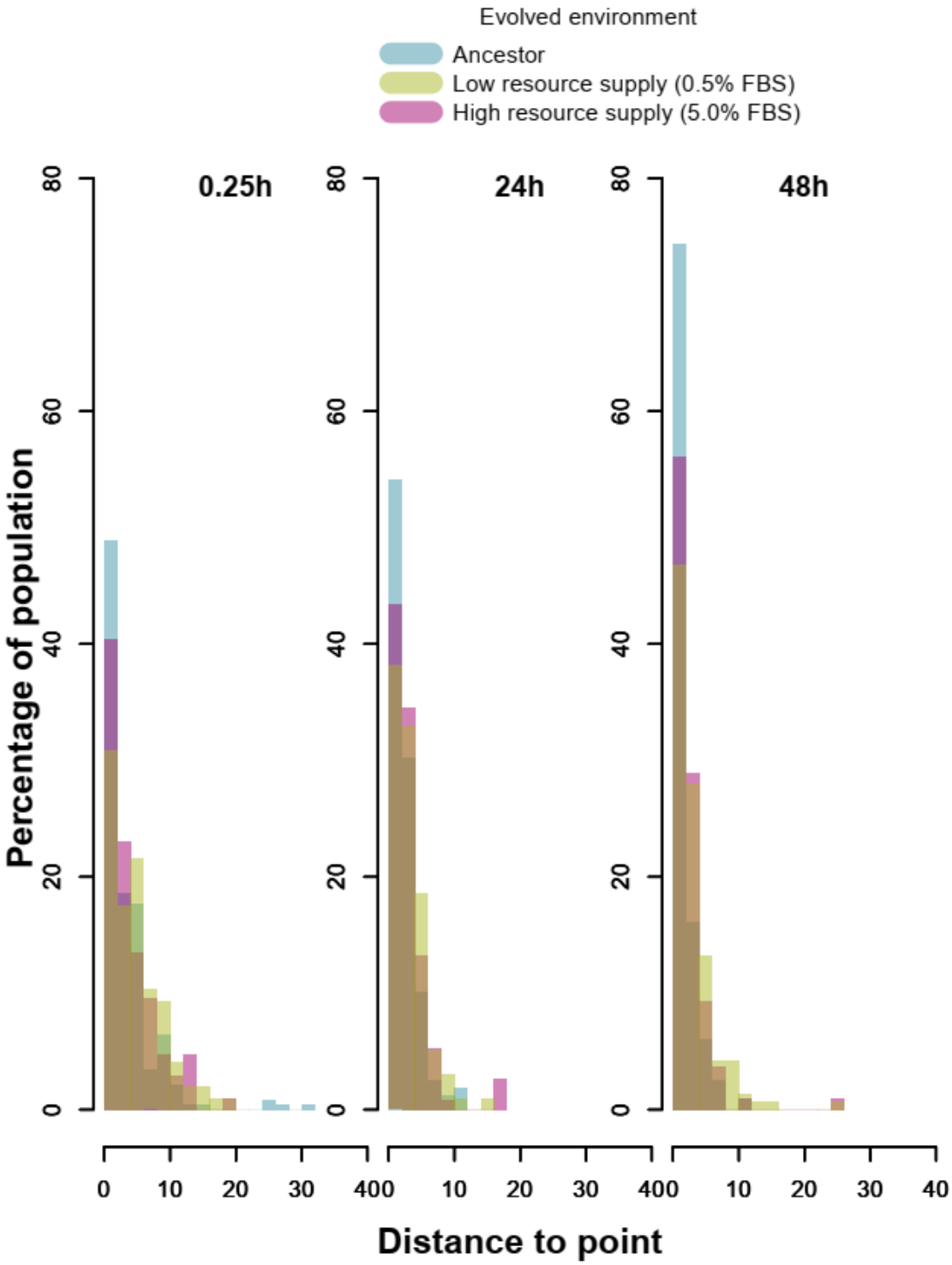


Figure 5.12: Histogram showing distribution of distance to point (D2P) measurements at 0.25, 24 and 48 hours for MCF7-1 cells evolved over 12 weeks in 0.5 % and 5 % FBS media and the ancestral population[38].

5.3.2.3 Heritability of MCF7-1 cells evolved for 12 weeks in 5 % and 0.5 % FBS media

Cell speed is significantly heritable for all cell:cell relationships, table 5.13 shows H^2 values range from 0.816 - 0.214 and R^2 values range from 0.05 - 0.67. H^2 has been estimated as stated in the methods section 5.2.7.3 using simple unweighted regression of cell speeds between related cells. Due to low N, H^2 could not be estimated for the cousin:cousin relationships of any cell group. Figure 5.13 shows how H^2 has been calculated for the mother:daughter relationship in each group of MCF7-1 cells. Statistical comparison of H^2 estimates between groups was not possible due to the number of post hoc comparisons. As shown in figure 5.14 it appears that both the evolved cell groups have greater heritability than the ancestral group. Particularly for the mother:daughter relationships where both the H^2 and R^2 values of the ancestral group is very low.

Table 5.13: Mean trait value, R^2 and broad-sense heritability of motility in MCF7-1 cells evolved for 12 weeks in 5 % and 0.5 % FBS media and the ancestral population for two cell:cell relationships.

Cell group	**No of families**	**Mean speed (µm/hour)**	**Heritability of cell motility H^2 (R^2)**	
			Mother:daughter	**Sister:sister**
Ancestor	87	6.83	0.214 (0.05)*	0.656 (0.43)***
Evolved 5 %	49	9.44	0.647 (0.42)***	0.816 (0.67)***
Evolved 0.5 %	76	9.63	0.650 (0.42)***	0.773 (0.6)***

*p<0.05 **p<0.01, ***p<0.001

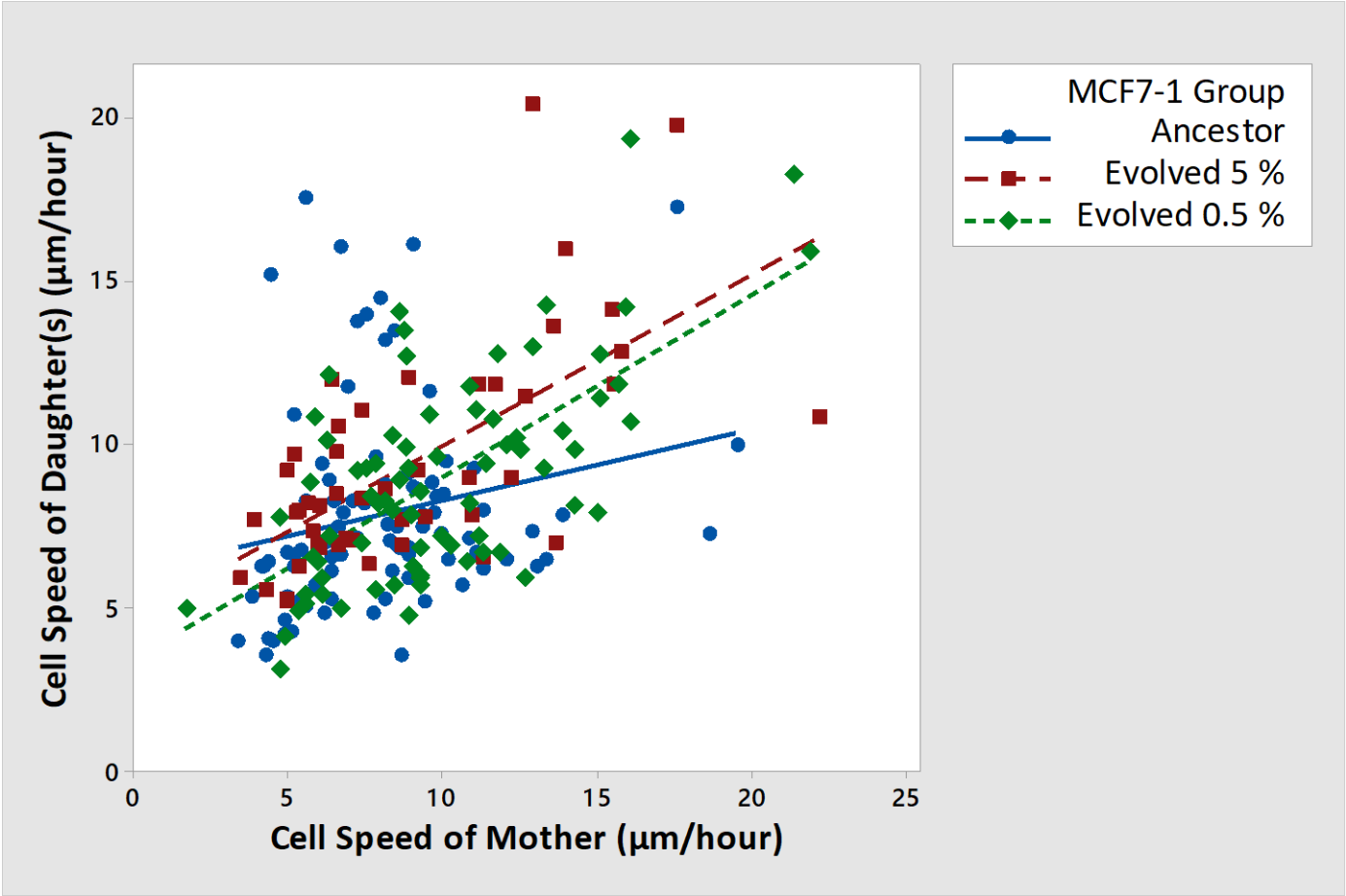


Figure 5.13: Example of the mother:daughter regression for motility in MCF7-1 cells evolved over 12 weeks in 0.5 % and 5 % FBS media and the ancestral population. Each point represents a mother cell and the mean speed of her daughters, the slope of the regression line is the estimate of broad-sense heritability. As the cells are clonal, the same method is applicable to other cell:cell relationships.

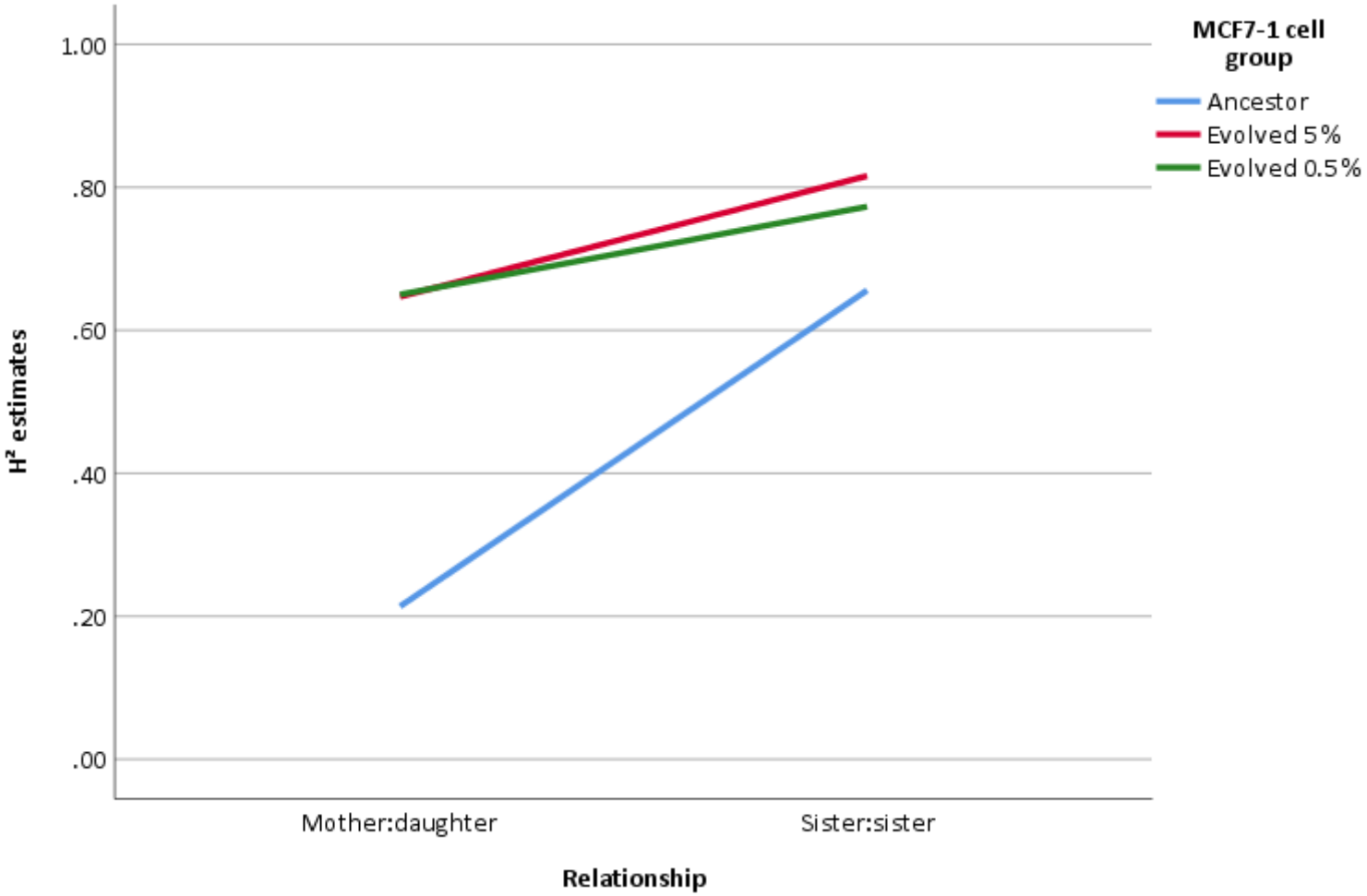


Figure 5.14: Heritability estimates of cell:cell relationships for ancestral and evolved MCF7-1 cells. N for each group is Ancestor = 87, Evolved 5 % = 49, Evolved 0.5 % = 79.

5.3.2.4 Common garden assays comparing growth of MCF7-1 cells in 5 % and 0.5 % FBS media

The rate of proliferation for each cell group was calculated as the percentage change in cell number over time (section 5.2.7.3) and measured in both the evolved media; 5 and 0.5 % FBS. When comparing growth at 12 hours to 72 hours, only the Evolved 5 % cells showed a significant difference with higher growth in the 5 % media (Mann-Whitney U test, $p<0.001$).

In both the 5 % and 0.5 % media Ancestor cells showed the lowest growth rates (Mann-Whitney U test, $p<0.001$). In the 0.5 % media there was no significant difference between the two evolved lines (Mann-Whitney U test, $p=0.152$) however in the 5 % media Evolved 5 % cells showed the highest growth rate (Mann-Whitney U test, $p<0.001$).

The percentage change in cell number for each cell group in both media can be seen in figure 5.15 data had a mix of normal and non-normal distributions (Lilliefor corrected Kolmogorov-Smirnov test).

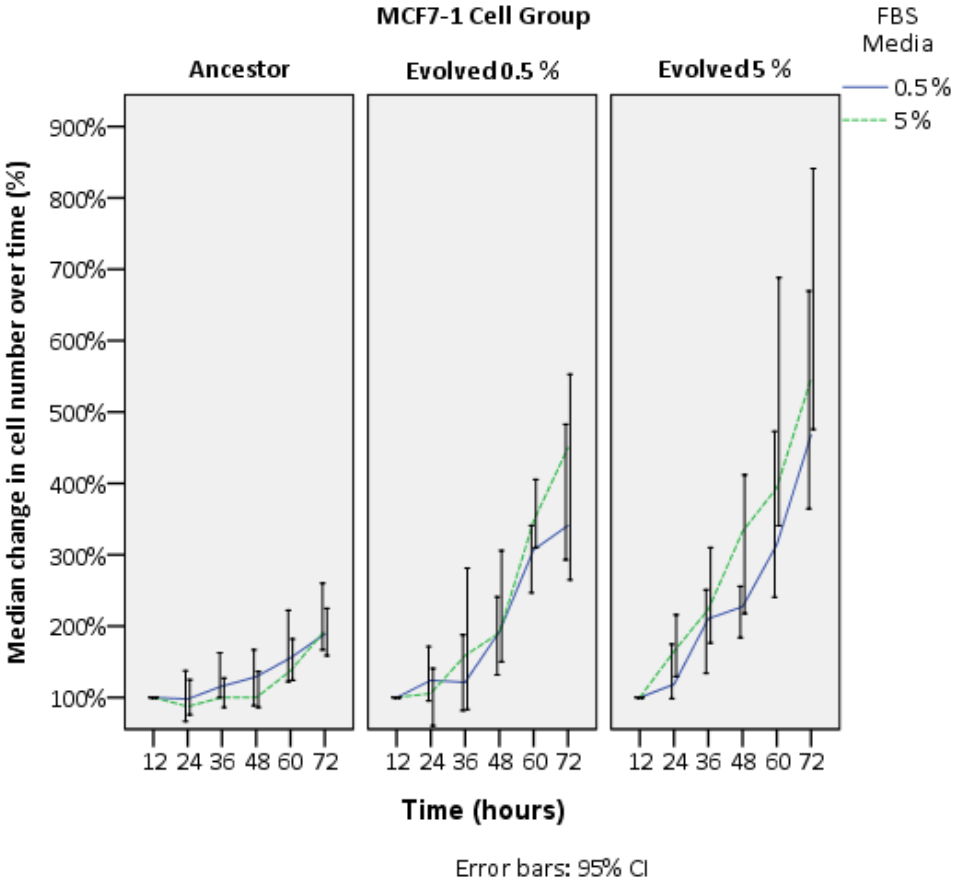


Figure 5.15: The median percentage change in cell number over time for MCF7-1 cells evolved over 12 weeks in 0.5 % and 5 % FBS media and the ancestral population when cultured in 5 % and 0.5 % FBS media. N = 3.

5.3.2.5 Spheroid spread of MCF7-1 cells evolved for 12 weeks in 5 % and 0.5 % FBS media

Figure 5.16 is an example of how MCF7-1 cells evolved in 5 % media for 12 weeks cultured as spheroids spread over time in 5 % and 0.5 % media. Spheroid spread was calculated as the percentage increase in area over 72 hours (section 5.2.4.3) and compared for each group.

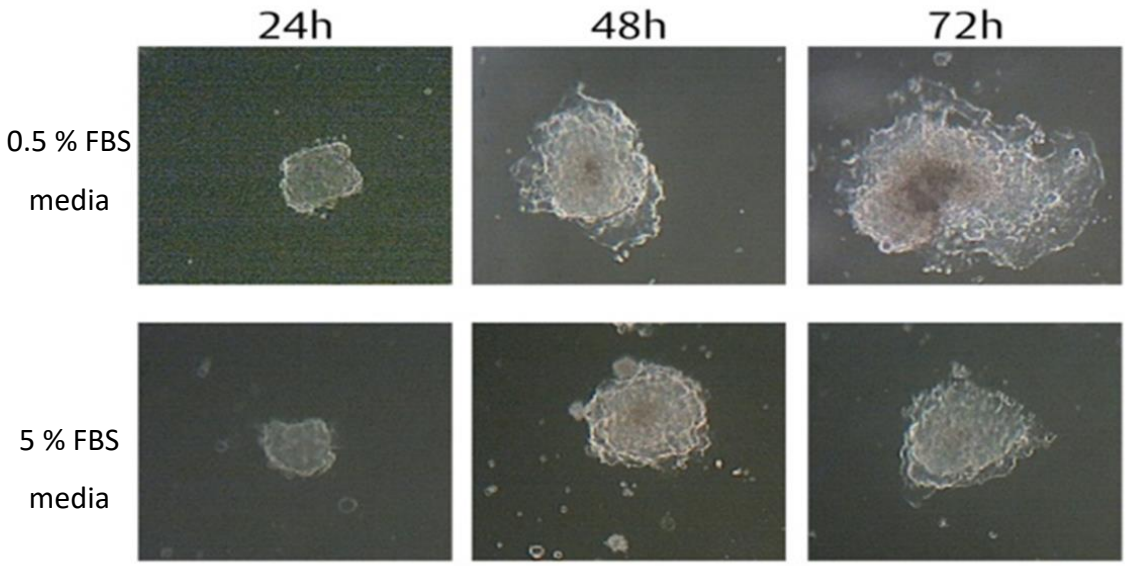
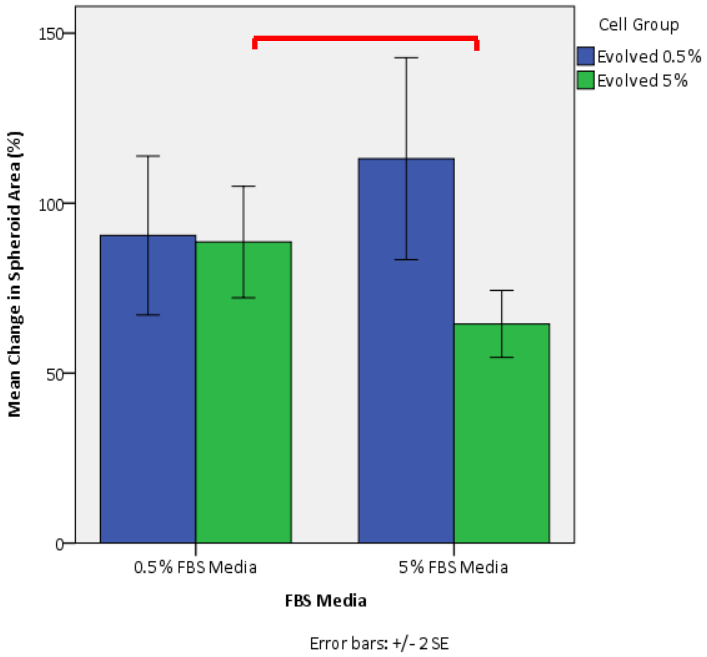


Figure 5.16: Example image of MCF7-1 cells evolved in 5 % media for 12 weeks cultured as spheroids then left to spread in 5 and 0.5 % media.

Figure 5.17 shows the mean average change in spheroid area for each group, significant differences between groups has been indicated with a red line. Evolved 5 % cells had a greater change in area when cultured in 0.5 % FBS media than in their adapted 5 % FBS media (independent T-test, p=0.002). This is despite them showing a faster growth rate in the 5 % media (as seen in the common garden assay section 5.3.2.4) indicating that migration has increased in the 0.5 % media.

Evolved 0.5 % cells did not show a significant difference in the change in spheroid area between the two media (independent T-test, p=0.249). Comparison of spheroid spread between the both evolved cell groups in their adapted environments showed no significant difference (independent T-test, p=0.057). Independent T-tests were used to compare the change in spheroid area between cell groups and between the media for each cell group. Parametric tests could be used as all groups had a normal distribution for the change in area (Lilliefor corrected Kolmogorov-Smirnov test; p>0.101) and were not heteroscedastic (Levene's test p>0.223).



<u>Figure 5.17:</u> Percentage change in spheroid area after 72 hours for MCF7-1 cells evolved over 12 weeks in 0.5 % and 5 % FBS media in both their adapted and the alternative media. The height of the bar is the mean average of each group and error bars extend for two standard errors either side. Significant differences between groups are indicated by a red line. Data gathered from three independent repeats.

5.3.2.6 Testing fluorescent tags for competition assays

MCF7-1 ancestor cells were labelled with fluorescent probes to test their suitability for use in competition assays, the fluorescent signal had to be maintained between cell divisions for a period of 72 hours. For all experiments cell number has been calculated as a percentage change over time to allow comparisons between groups (section 5.2.7.3). Three fluorescent tags were tested, Qtracker was unsuccessful at fluorescently labelling cells. CellLight and CellTracker did fluorescently label cells but inhibited cell proliferation so are not suitable for use in competition assays.

5.3.2.6.1 CellLight

As seen in figure 5.18 cells labelled with CellLight show a significant decrease in cell number over time (Kruskal-Wallis, $p<0.001$) the control group shows a significant increase (Kruskal-Wallis, $p<0.001$). This suggests that either CellLight is inhibiting cell proliferation and/or increasing apoptosis or that photobleaching is occurring and the fluorescent marker is degrading over time leading to an inaccurate cell count. This make GFP CellLight an unsuitable fluorescent tag for use in competition assays. The percentage change in cell number over time was non-normally distributed (Lilliefor corrected Kolmogorov-Smirnov test; $p<0.001$).

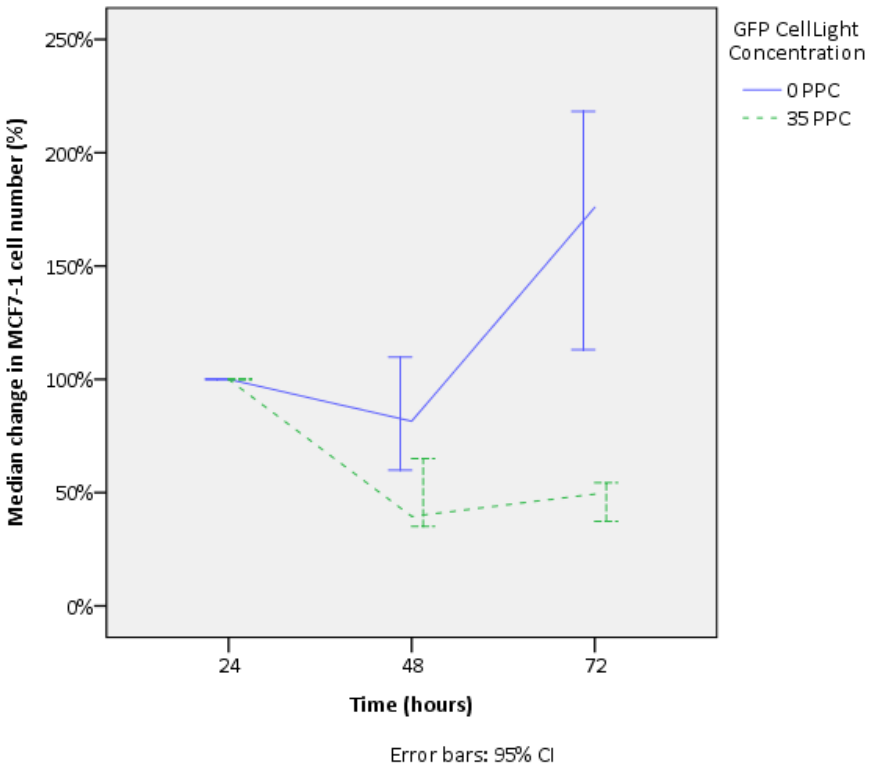


Figure 5.18: The median percentage change in cell number over time of MCF7-1 cells with or without GFP nuclear CellLight. Data gathered from three independent experiments.

5.3.2.6.2 CellTracker

As seen in figure 5.19, there was no significant difference in cell growth between 10 µM red CellTracker and the control group (Mann-Whitney U test, bonferroni correction, p=0.03) indicating that red CellTracker does not inhibit proliferation and so may be suitable for use in competition assays. Caution must be taken when interpreting results as the percentage change in cell number over time was non-normally distributed for all groups (Lilliefor corrected Kolmogorov-Smirnov test; p<0.001). Even if red CellTracker is suitable for use in competition assays green CellTracker will also be needed to distinguish between cell groups.

There is a significant difference at all time points between the control group and cells labelled with 10 µM green CellTracker (Kruskal-Wallis, p=0.276). As mentioned in section 5.3.2.6.1 the decrease in cell number over time indicates that either green CellTracker is inhibiting cell proliferation and/or increasing apoptosis or photobleaching of the fluorescent label is occurring leading to an inaccurate cell count.

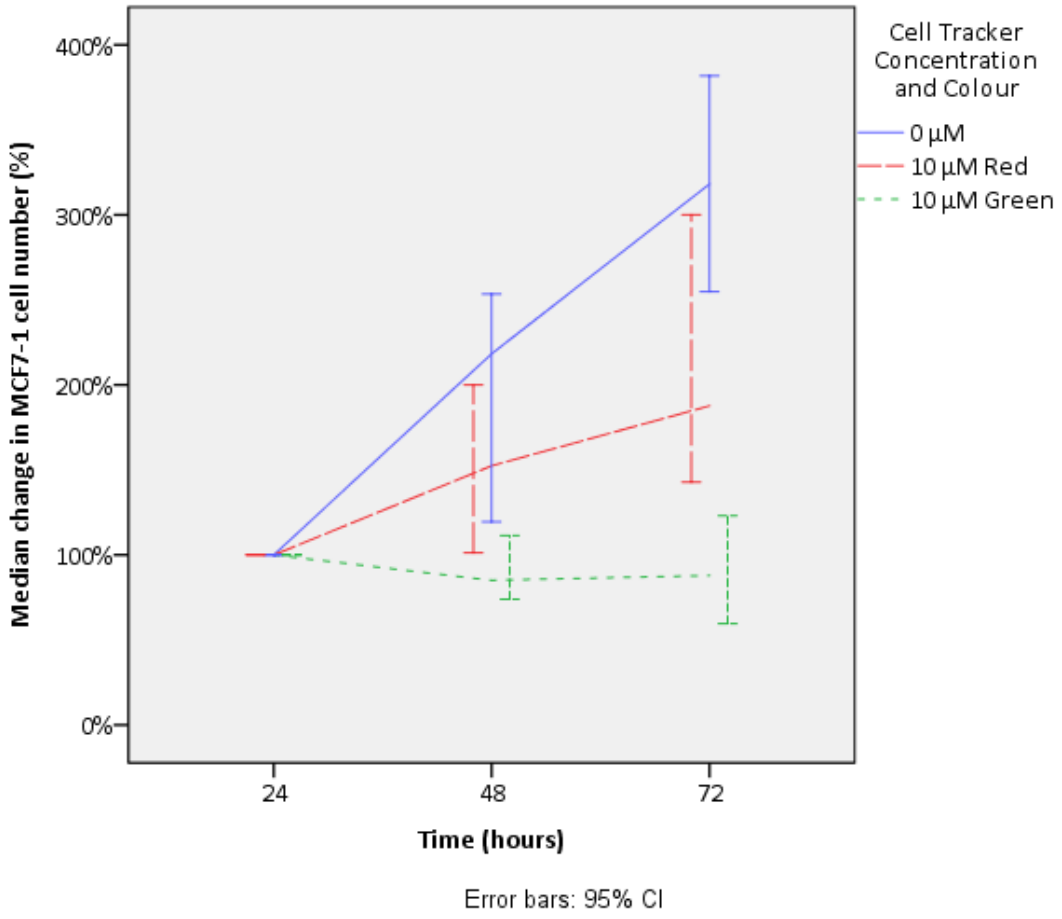


Figure 5.19: The median percentage change in cell number over time of MCF7-1 cells for CellTracker concentrations (µM). Data gathered from three independent experiments.

To test whether the decrease in cell number is due to inaccurate fluorescent labelling with green CellTracker, cell counts were obtained both with a fluorescent filter (only fluorescent cells will be counted) and without (the total number of cells present in the field of view). To investigate whether the decrease in cell number is due to green CellTracker inhibiting proliferation a range of concentrations (0, 2.5, 3.25, 3.75, 5 and 6.5 µM) of the green CellTracker were investigated to see if any concentration could stably label cells without altering proliferation.

Only cells labelled with 5 µM green CellTracker showed a significant difference between the counts, with higher growth observed without a filter (Mann-Whitney U test, $p=0.009$). This suggests that the fluorescent label is maintained over the 72hour period and is not degrading over time leading to an inaccurate cell count.

As seen in figure 5.20, only cells in 0 µM showed a significant increase in the percentage change in cell number over time (Kruskal-Wallis, $p<0.001$) no other concentration showed a significant difference (Kruskal-Wallis, $p>0.037$). This suggests that green CellTracker is inhibiting cell proliferation/or increasing apoptosis, either of which make it unsuitable for use in competition assays. The percentage change in cell number over time was non-normally distributed for all groups (Lilliefor corrected Kolmogorov-Smirnov test; $p<0.001$).

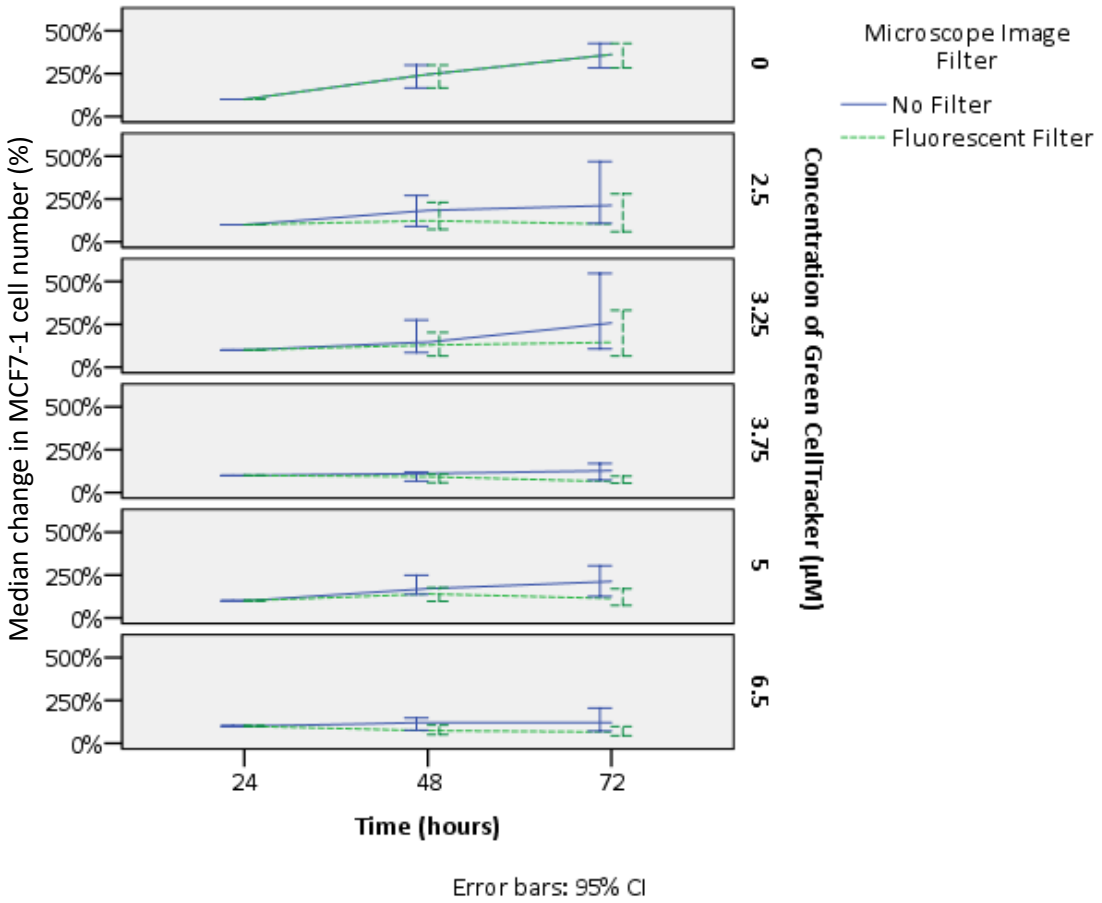


Figure 5.20: The median percentage change in cell number over time of MCF7-1 cells for green CellTracker concentrations (µM). Cell counts were obtained both with (only fluorescent cells counted) and without (all cells present in the field of view counted) fluorescent filters. Data gathered from three independent experiments.

5.3.3 Effect of resource restriction on MCF7-2 cells

MCF7- 2 cells were cultured for 6 months in 5 % and 0.5 % FBS media, 0 mM and 1 mM Glucose media and at pH 6.5 and pH 7. Multiple cell phenotypes were then measured and compared to test for adaptations in the cell populations. Due to mycoplasma infection it was not possible to include all cell groups in each experiment. Unless gained prior to a clear contamination test, results could not be included or used for analysis.

5.3.3.1 Motility of MCF7-2 cells evolved for 6 months in reduced glucose and pH environments

Time-lapse microscopy measured motility of MCF7-2 cells evolved for 6 months in acidic (pH 7), low glucose (1 mM and 0 mM Glucose) and their respective control media (MEM and DMEM Control).

The results in table 5.14 suggest that restricting glucose selects for increased motility but removing it completely does not as only cells in 1 mM Glucose had significantly higher speed than the DMEM control cells (Mann-Whitney U tests, $p<0.001$). Culturing cells in an acidic environment appears to negatively impact cell motility, with cells cultured at pH 7 being significantly slower than MEM control cells (Mann-Whitney U tests, $p=0.001$).

The distribution of cell speeds in each group can be seen in figure 5.22, all groups had non-normal distributions of cell speed (Lilliefor corrected Kolmogorov-Smirnov test; $p<0.001$). Due to small N it was not possible to conduct statistical analysis on the generations or fate of cells, suggesting that resource restriction over long periods of time may negatively impact cell proliferation.

Table 5.14: Median speed and N of control and evolved groups of MCF7-2 cells.

MCF7-2 cell group	**Median speed (µm/hour)**	**N**
MEM Control	12.19	1520
pH 7	7.42	375
DMEM Control	8.05	347
1 mM Glucose	10.59	354
0 mM Glucose	8.24	301

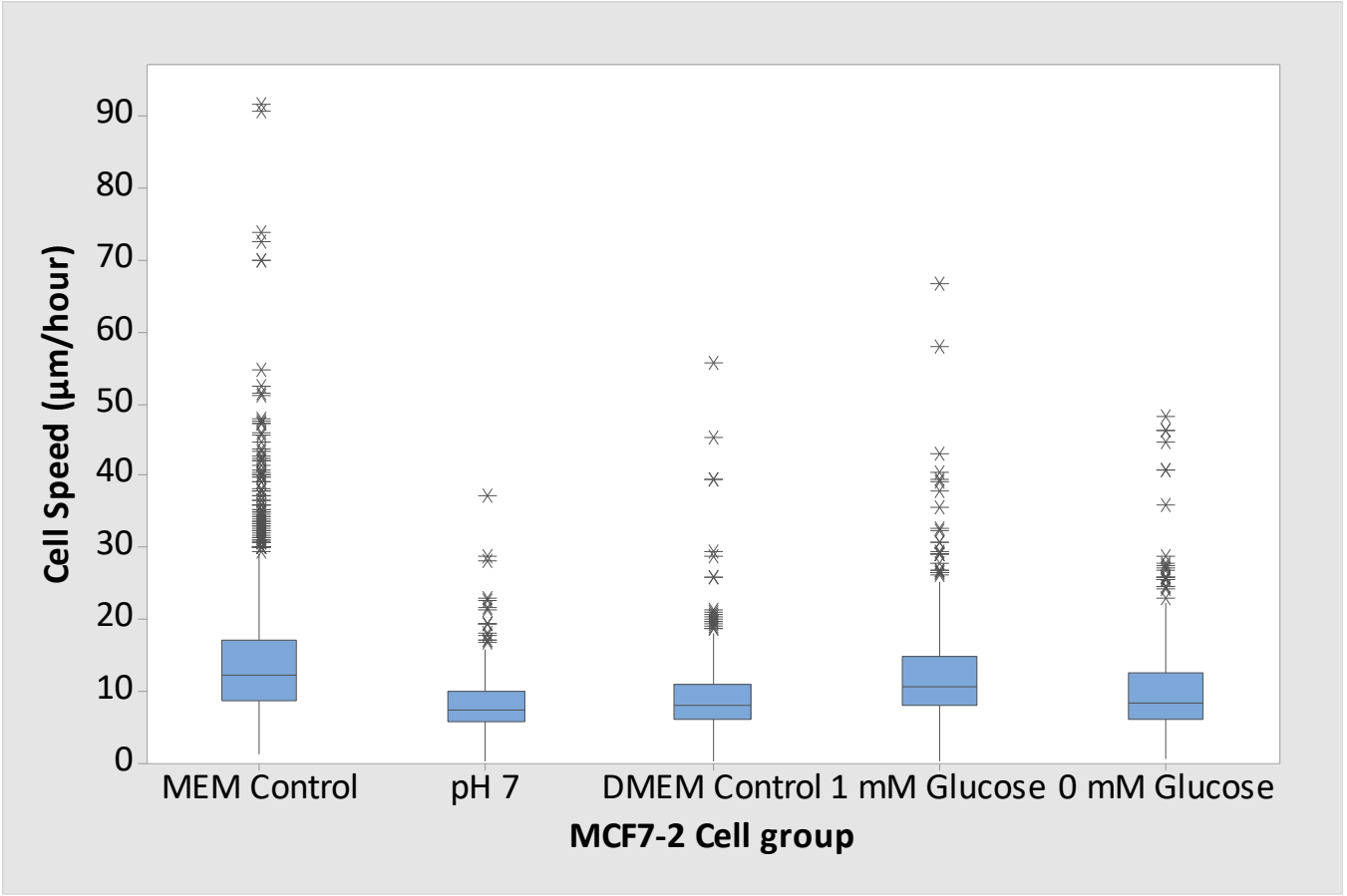


*Figure 5.21: Distribution of cell speed for MCF7-2 cells evolved over 6 months in low glucose and low pH media and the ancestral populations. The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. Data was gathered from two independent experiments. The N for each group is; 0 mM Glucose 301, 1 mM Glucose 354, DMEM Control 347, MEM Control 1,520 and pH 7 375.*

5.3.3.2 Live cell concentration of MCF7-2 cells cultured in resource restricted media

MTT assays were used to determine the live cell concentration, a marker for the amount of cell proliferation or death. Increasing resource restriction was predicted to increase competition and decrease proliferation. This hypothesis is not supported by the results.

As seen in figure 5.22 and table 5.15, over 72 hours there is a significant difference in live cell concentration between groups (Kruskal-Wallis $p<0.001$) with cells in 5 % FBS showing less proliferation than the MEM control and pH 6.5 groups which were all lower than cells in pH 7. The DMEM control group had a significantly greater live cell concentration than either 1 mM or 0 mM. All groups had a non-normal distribution of cell concentration (Lilliefor corrected Kolmogorov-Smirnov test; $p<0.01$).

Table 5.15: Median standardised absorbance values and N of control and evolved groups of MCF7-2 cells.

MCF7-2 cell group	**Standardised Absorbance values (OD at 540 nm)**	**N**
MEM Control	0.059	192
pH 7	0.078	209
pH 6.5	0.054	196
5 % FBS	0.033	72
DMEM Control	0.193	208
1 mM Glucose	0.107	144
0 mM Glucose	0.117	216

If resource restriction was causing a reduction in proliferation it would be expected that all of the evolved groups would have lower live cell concentrations than their respective controls. Cells evolved in pH 7 media have a higher average live cell concentration than their control. As seen in figure 5.23 there is a negative correlation (-0.023) between average motility and proliferation. These results could suggest a trade-off between increased motility and reduced growth, cells in 1 mM and 0 mM Glucose showed less proliferation and greater motility than the DMEM control and cells at pH 7 showed more proliferation and lower motility than the MEM control.

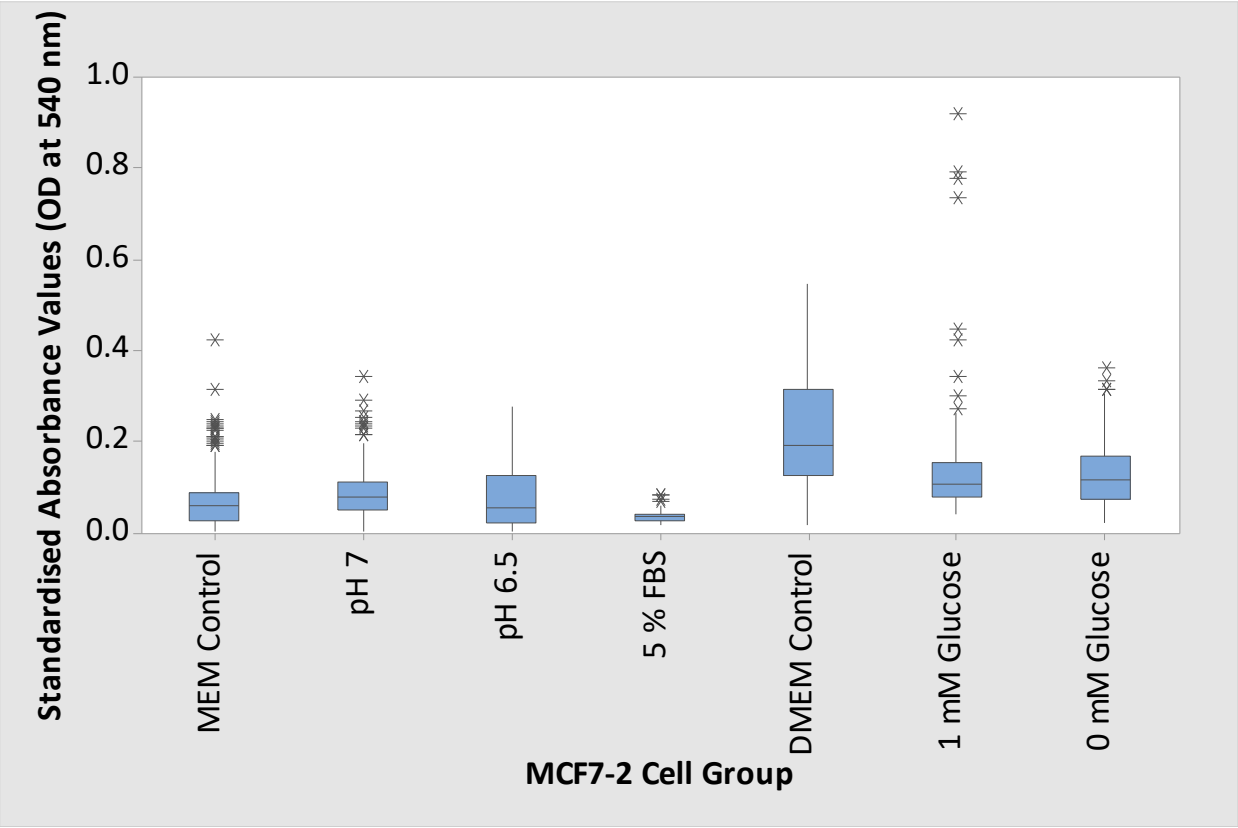


*Figure 5.22: Distribution of live cell concentration for MCF7-2 cells evolved over 6 months in low glucose, low pH and low FBS media and the ancestral population. The median is the line within the box and the box represents the interquartile range. The whiskers extend to the top and bottom 25 % of data values (excluding outliers). Outliers are represented by * and are data values extending at least 1.5 times the interquartile range from the box. Data was gathered from three independent experiments for 0 mM Glucose, pH 7 and DMEM Control cells, two independent experiments for 1 mM Glucose and pH 6.5 cells and only one experiment for 5 % FBS cells.*

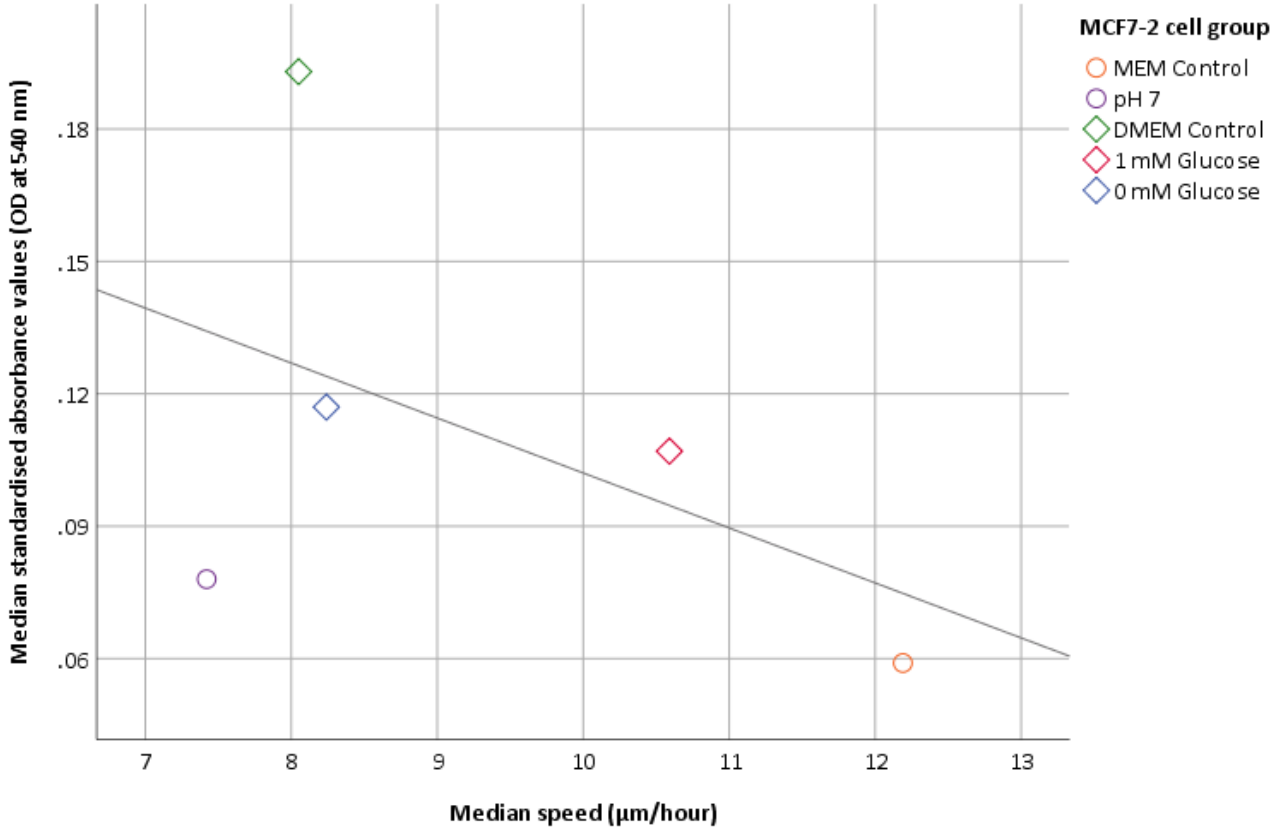


Figure 5.23: Relationship between the average proliferation (standardised absorbance values at OD 540 nm) and motility (µm/hour) for five groups of MCF7-2 cells. The average is the median of each group and data has been gathered from three independent repeats for the proliferative rate and two independent repeats for the motility.

5.3.3.3 Wound healing rate of MCF7-2 cells cultured in resource restricted media

Wound healing assays measure collective migration and were calculated as the percentage reduction in wound size over time. It was predicted that decreasing resources would increase motility however this was not seen in the results. The only significant difference between groups was that cells in 5 % FBS had a significantly faster mean wound healing rate than those in pH 7 (ANOVA Tukey p = 0.03) or the MEM control (ANOVA Tukey p = 0.02).

Figure 5.24 shows the mean percentage change in wound area for all groups, data was normally distributed for all groups (Lilliefor corrected Kolmogorov-Smirnov test; p>0.201). These results are different to those obtained when individual cell motility was measured and demonstrate the importance of measuring different aspects of metastasis.

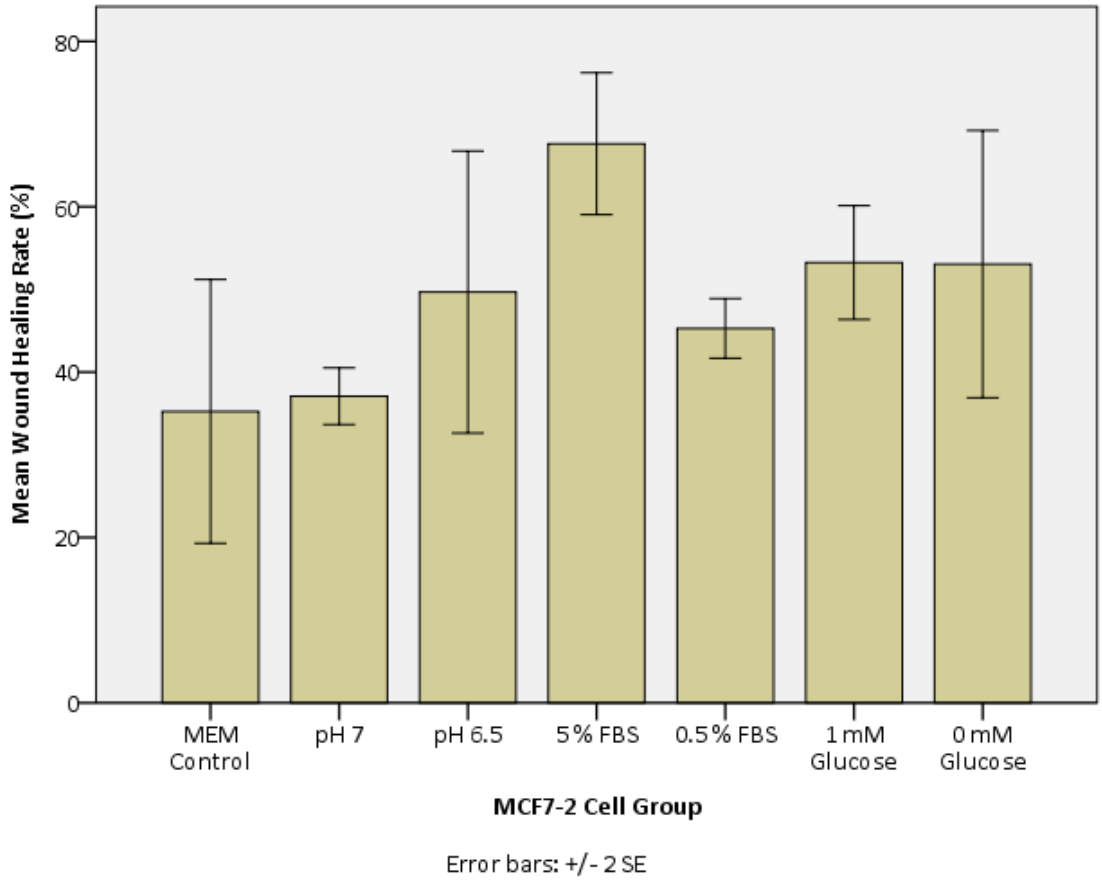


Figure 5.24: Wound healing rate for MCF7-2 cells evolved over 6 months in low glucose, low pH and low FBS media and the ancestral population. The height of the bar is the mean average of each group and error bars extend for two standard errors either side. Data gathered from three independent repeats.

5.4 Discussion

In this chapter experimental evolution techniques have been developed to enable the application of dispersal theory to cancer cells *in vitro*. Selection experiments were used to culture cancer cells in vitro in resource restricted environments for varying periods of time and their response to selection measured. Results suggest that selection for 12 weeks in a resource restricted environment does select for increased motility[38].

5.4.1 Adaptation of experimental evolution techniques to *in vitro* cancer cell culture

Techniques had to be adapted for use in cancer cell culture[36]. Whilst cancer cells do bear many similarities to microorganisms making them well suited for experimental evolution, they are also far more complex having evolved as part of a multicellular organism before undergoing their own individual mutations[62].

5.4.1.1 Creating a serum free media

Cancer cells are fastidious in their growth. Each cell line has specific media requirements and 'lab adapted' strains are those that have adapted to their culture conditions. Whilst many media supplements are well defined serum is a very variable media supplement. Serum is the liquid component of clotted blood it contains many substances whose concentrations and effects are unknown[64 243]. It is collected from individual animals sent to slaughter so cannot be standardised from batch to batch or made specific to a certain cell line[230 258 271]. This variation makes standardisation of protocols difficult both within and between laboratories[64]. As different batches as well as different cell lines have been tested throughout this project this could explain some of the variation in results.

Attempts were made to create a serum free media suitable for cancer cell culture. If successful, this serum replacement media would have allowed standardisation across experiments[64]. Cell culture media has been specifically adapted to maintain cell culture *in vitro*. Changing the conditions by removing supplements or altering concentrations resulted in cell death. Only a small selection of serum-free media formulations exist and these are confined to specific cell lines. Furthermore, these formulations are proprietary, they cannot be reproduced and the exact composition remains unknown[243]. Serum free medium must be cell type specific, the lack of success in this project suggests either a lack of essential components or incorrect concentrations. Cell survival and proliferation could not be maintained in the

serum free media, even when using a commercially available serum replacement supplement. The time and cost of developing a serum free media meant that in the scope of this project it was more viable and effective to use standard basal medium with serum supplementation[64].

5.4.1.2 Creating a heterogeneous environment

Biomolecular gradients play a vital role in cell growth, migration and differentiation[247] and have been linked to cancer metastasis[248]. It was hoped that culture of cells in a biomolecular gradient would demonstrate the selective pressures of resource restriction on motility.

Cell migration *in vitro* is a relatively slow process (approximately 10 µm/hour). This creates problems in setting up and maintaining a resource gradient over the length of time needed, whilst also maintaining a large enough population of cells to be statistically relevant. Attempts were made to culture cells in a resource gradient, section 5.2.3.1, these attempts were unsuccessful due to the low numbers of cells able to be maintained in the cultures or the low motility of the cell populations resulting in low cell numbers (table 5.2).

Trial of these methods and equipment demonstrated that the transwell assay (table 5.2) would be the most suitable technique for selection of motile cells in *in vitro* culture. Whilst in this project the transwell assay proved unsuccessful, this could be due to the small pore size of the inserts, 0.4 µm is much smaller than the average cancer cell used in this project (average size 40-100 µm). The cell types used could also have played a role in the lack of success, smaller more motile cells may migrate in high enough numbers to make re-culture of cells a possibility. It could also be that the time period monitored (between 24 hours and 14 days) was not long enough for selection to have occurred. Using a transwell insert with larger pore sizes as well as culture of cells for an extended period of time in resource restricted environments before selection using a transwell assay may increase experiment success.

Cells in culture are not in an entirely homogenous environment (Figure 4.10, section 4.3.3) the effects of kin competition may still act as a selective pressure to increase cell motility even in the absence of a predictable controlled resource gradient[14 41].

5.4.1.3 Testing of fluorescent probes for use in competition assays

Competition assays involve the co-culturing of two populations creating competition for resources. They measure the relative fitness of the populations to each other, usually by the number of survivors or each population's contribution to future generations[14 235]. There may

be trade-offs between cell phenotypes such as motility and growth. Competition assays provide a comprehensive measure of fitness as they measure how a combination of phenotypes enable cell survival[14 51 235].

Before competition assays could be attempted the use of fluorescent probes and their effect on cell phenotypes had to be tested. All of the fluorescent tags tried in this project (section 5.2.4.4) are unsuitable for use in this project. CellLight and CellTracker significantly reduced cell growth (figures 5.18 and 5.19) and QTracker was unsuccessful at labelling cells. Whilst red CellTracker did not show a significant reduction in cell growth, it did appear to have an effect over the 72 hour period (figure 5.19). This makes them unsuitable for use in this project as labelling cells for competition assays would add another level of selection to the experiment and the reduction in cell growth may invalidate results.

The fluorescent tags chosen in this project all acted via different mechanisms and targeted various locations so would not have been expected to affect cell characteristics in the same manner[84]. CellLight uses endocytosis of BacMam technology (a baculovirus coupled with a mammalian promoter) to induce nuclear expression of fluorescent molecules[217 272 273]. Transfection of a cell, forcing it to express a foreign protein may reduce its fitness through reduction of resources needed for essential cell functions.

A range of green CellTracker concentrations were trialled (figure 5.20) to determine if a lower concentration would allow its use in competition assays. No concentration was found that did not affect cell growth. As seen in figure 5.20 the reduction in growth when cells were labelled with CellTracker was not due to a lack of precision in fluorescent labelling or to a reduction in the fluorescent signal making it unlikely that the results seen were due to issues with fluorescent signalling. CellTracker is a cytoplasmic dye which should passively diffuse through the cell membrane and undergo a hydrolysis reaction resulting in glutathione conjugation[217]. A study by Lulevich et al (2010) found a significant effect on the cell characteristics of rigidity and adhesion when MDA-MB-468 and MLC-SV40 cells were fluorescently labelled with CMFDA (green CellTracker) [274]. Glutathione distribution is heterogeneous and whilst abundant (in the millimolar range) the CellTrackers reaction with intracellular proteins may be depleting glutathione and affecting cell function[275].

These results suggest that fluorescently labelling cells reduces their fitness meaning *in vitro* co-culture of cancer cell lines for use in competition assays is not possible. Using competition assays as a measure of cell fitness has disadvantages, cell behaviour is comprised of integrated

phenotypes and phenotypic variability could arise from either developmental plasticity or phenotypic flexibility[15]. By measuring a discrete phenotypic trait (in this case proliferation in co-culture) the population 'fitness' could be over or under estimated. Cell lines with a faster proliferative rate would be expected to outcompete those with a lower proliferative rate irrespective of their environment. An alternative to competition assays would be quantitative analysis of multiple cell traits either in an individual cell or in population and comparison between cell lines. By using multiple different techniques and comparison of many different phenotypes this may give a better representation of fitness and allow comparison between groups without modifying populations.

5.4.2 Selection of cells in serum restricted media

Dispersal theory predicts that culture of cells in a resource restricted environment would increase competition and hence select for increased motility[14 36 39]. Serum contains growth and adhesion factors which promote cell proliferation and cell attachment, it is also a source of proteins, minerals, lipids and hormones[64]. FBS is an essential component of cell culture providing many nutrients essential to cell survival[64]. Reducing the serum concentration was predicted to increase competition between cells.

5.4.2.1 Selection over 72 hours in HeLa, HT1080, MCF7-3 and MDA-2 cells

Four cell lines, HeLa, HT1080, MDA-2 and MCF7-3 were cultured in a high (10 % FBS) and low (0.5 % FBS) nutrient media for 72 hours. For MCF7-3 and MDA-2 cells decreasing the serum concentration over a 72-hour period did not increase motility, cells in 10 % FBS displayed higher motility than those in 0.5 % FBS. However, there was no difference in the growth of these cell lines between the 10 and 0.5 % FBS media. This could suggest that MCF7-3 and MDA-2 cells exposed to a reduction in serum maintained their proliferation and reduced their motility to do so. Both HeLa and HT1080 cells showed reduced proliferation in the 0.5 % FBS media, HeLa cells showed no difference in motility between the media and HT1080 cells showed increased motility in the 0.5 % FBS media. This could suggest that these cell lines reduced their proliferation in order to maintain cell motility. These results indicate there might be a trade-off between proliferation and motility.

HeLa, HT1080 and MCF7-3 cells tended to show greater H^2 estimates for motility in 0.5 % FBS than 10 % FBS, MDA-2 cells had greater H^2 estimates in 10 % FBS than 0.5 % FBS (table 5.9). The magnitude of a phenotypes response to selection depends both on the heritability and

strength of selection[276]. Adaptation is the process of 'fit' genotypes (in this case cells with higher motility) rising in frequency, the mutation rate and the ancestral range in motility will affect how quickly this occurs. In a population with a low mutation rate and a narrow range in motility acquisition of new phenotypes may take longer, expansion of the clones already in the population will occur meaning the shift in population average is less marked[113]. HT1080 cells had the greatest variance and range in cell motility (seen in figure 5.4), this could explain why only this cell line is showing a significant increase in cell motility. HT1080 cells also had the fastest average division time, 17.1 hours. This faster division rate along with a greater range in cell motility could mean selection is only producing a measurable effect on the HT1080 cell line[113].

Studies on phytoplankton have shown that organisms exposed to resource limitation over a short period of time show plastic alterations in molecular phenotype including protein expression[277]. Western blot analysis on the protein expression of each cell line in 10 % and 0.5 % FBS media shows that apart from HeLa cells there does not seem to be a difference in protein expression between the two media (figure 5.8). Over 72 hours selection may not have had chance to act at a population level; this is partly due to the phenotypic plasticity cancer cells display when adapting to their environment [15 29 276]. Other studies have shown over short time periods (24-48 hours) cancer cells are capable of producing plastic responses to a reduction in nutrient levels. These results may simply show a cell populations immediate response to the lack of nutrients and not selection for a particular trait.

5.4.2.2 Selection over 12 weeks in MCF7-1 cells

MCF7-1 cells were cultured in serum reduced media (5 % and 0.5 %) for 12 weeks. This longer time frame allowed selection to have occurred and for populations to have adapted to their environments. The results presented in section 5.3.2 are part of a larger project by this lab, see appendix A. The data presented in section 5.3.2 has been further expanded on by other members of this lab leading to an increase in N and inclusion and testing of more cell groups for some experiments.

Both evolved cell lines had significantly higher motility in comparison to the ancestor (figure 5.10). In the data gathered as part of this project there was no statistically significant difference between the two evolved groups when comparing their migration in the spheroid spreading assays (figure 5.17). However further work done by this lab, see appendix A, where growth rate was accounted for in spheroid spread, did show a significant difference between

groups with Evolved 0.5 % cells having faster spheroid spread than both Evolved 5 % and Ancestor cells[38]. These results suggest that dispersal theory can be applied to cancer cells and that resource competition selects for increased motility.

Spheroid spreading assays involve the collective migration of cells and have been used to study spreading behaviours of cancers[251-253]. This type of migration is different to measuring individual cell motility and is a more accurate representation of an *in vivo* environment[249]. Measuring individual cell motility gives a more precise measure of population motility and may be more accurate as it is possible to gather data on each individual cell (generation, fate and relatedness to other cells). Using multiple different techniques and applying them to experimental evolution methods enables a more accurate measure of cell fitness and more robust testing of hypotheses. Resource restriction appears to increase both individual cell motility and collective migration.

Both evolved cell groups also had higher H^2 than the ancestor (table 5.13) which indicates that the proportion of motility due to genetic components has increased or that the environmental variation has decreased. Reducing the serum concentration may well have lowered the amount of environmental variation but as shown in figures 4.10 and 4.11 even at very low concentrations solutions can still be heterogeneous at a cellular level. The R^2 values for ancestor mother:daughter cells was very low, indicating that the regression model explained the data poorly which could explain why the H^2 estimates were much lower in this group than in the evolved cells. The increase in H^2 estimates could also occur due to genetic drift or if the changes in motility of the evolved lines was a plastic response modulated by epigenetic mutations[38 47].

When the distance to point was measured (section 5.3.2.2), the distance each cell moves between time frames[38]. It was found that Evolved 0.5 % cells had fewer non-motile cells at any given time point than both Evolved 5 % and Ancestor cells[38], seen in figure 5.12. Figure 5.10 suggests that reducing serum concentration is selecting for increased motility, figure 5.12 suggests that this increase in motility is due to an increase in the rate of movement. Cells cultured in 0.5 % FBS are consistently motile as resource competition is always high however in 5 % media cells may increase their motility in response to increased competition, such as an increase in the number of surrounding cells. Further work is currently ongoing in this area to investigate motility in response to surrounding cells.

Common garden assays showed Evolved 5 % cells had a significant difference between media for both their growth rate and spheroid migration. With higher growth in the 5 % media but greater migration in the 0.5 % media. This indicates an initial plastic response of increased motility, at the cost of decreased proliferation when placed under increased resource restriction. Experiments on E. coli have found a trade-off between growth and motility with bacteria grown in a cyclical environment showing a plastic alteration of increased motility at the cost of reduced growth rates when resource availability becomes heterogeneous[278].

Evolved 5 % cells also showed a higher growth rate than Evolved 0.5 % cells but only in the 5 % media. It may be that in 0.5 % media there is a maximum growth rate based on the concentration of essential nutrients. If evolved 0.5 % cells have reduced their proliferation due to resource restriction it may take time for adaptation to occur and the proliferation rate to increase[113].

5.4.3 Selection of MCF7-2 cells over 6 months in resource restricted media

Lab adapted cell lines acquired from commercial sources come with recommended culture media. The culture media provides the necessary nutrients, growth factors and hormones for cell growth as well as helping regulate pH and osmotic pressure[64 80 81]. Attempts were made to culture MCF7-2 cells for a longer time period than previously achieved and in a greater range of nutrient deprived environments. Media was altered so that cells were cultured in reduced glucose and FBS and at an acidic pH. It was hypothesised that reducing the nutrient concentrations in culture will increase the competition between cells, this increased competition due to resource limitation would select for an increase in motility as motile cells migrate away from competition and towards areas with higher nutrient concentrations[38].

Glucose is an energy source so a reduction in glucose would be expected to increase competition, dispersal theory predicts this should select for increased motility. Media contains 5.5 mM of glucose; this is loosely based on the approximate concentration of glucose in the arteries and capillaries[279 280]. Cells were cultured for 6 months at reduced glucose concentrations of 1 and 0 mM.

Most cell lines, and all those used in this project, grow well at pH 7.4. Media controls and buffers the pH of cells in culture by a balance between dissolved carbon dioxide (CO_2) and bicarbonate ions (HCO_3^-)[281]. Whilst pH itself is not a resource, hypoxic environments that can induce acidic pH have been shown to increase metastatic progression and capability[244] and

molecular processes are greatly influenced by environmental pH. Culture pH was lowered using lactic acid to pH 7 and pH 6.5.

As previously discussed, serum is an essential component of cell culture providing many nutrients essential to cell survival[64]. Reducing the serum concentration to 5 and 0.5 % has been shown to select for increased motility but with a potential trade off in reduced proliferation (section 5.4.2.2). Selection of MCF7-2 cells over a longer time period under the same serum restricted conditions was conducted to see whether similar adaptations occurred and to allow comparison between the glucose restricted and acidic pH cultures.

5.4.3.1 Selection over 6 months in glucose restricted media

Figure 5.21 shows cells evolved in 1 mM Glucose have increased motility in comparison to both their control and cells in 0 mM Glucose. No significant increase in motility was measured for cells cultured in 0 mM Glucose. When comparing the rates of collective migration and proliferation no significant differences were seen between the two groups. The similarities between groups in their collective migration may be caused by concentration gradients created when cells are cultured as spheroids. This increase in resource heterogeneity for other nutrients may be causing cells in 0 mM Glucose to display a plastic increase in motility, cells increase their motility to migrate to areas with higher resource concentrations, when they achieve this their motility may revert to previous levels. Plasticly increasing motility may aid in balancing the trade-off seen between motility and proliferation, figure 5.23.

In cancer cells glucose is metabolised via glycolysis which has been linked to attenuation of apoptosis[242 282 283]. It may be that removing glucose altogether not only completely removes an energy source but also increases apoptosis, if a strong selective pressure is applied on a small range of phenotypes it may lead to population extinction. This would occur irrespective of whether natural selection was occurring[113]. Whilst population extinction has not occurred it may be that cells in 0 mM Glucose may have to expend more energy on maintaining their proliferation.

5.4.3.2 Selection over 6 months at an acidic pH

Extracellular pH is known to affect cell behaviour *in vivo*[284] and an acidic pH increases the metastatic potential of cancer cells[245]. MCF7-2 cells maintained in culture at pH 7 showed reduced motility compared to the control (figure 5.21) however when the rate of collective

migration was compared there was no difference between cells cultured at an acidic pH and the control (figure 5.24). The growth rate of cells cultured at pH 7 is significantly higher than both the control and cells at pH 6.5, figure 5.22. The increased proliferation of cells in pH 7 could explain why there is no difference in the collective migration of pH 7 cells and the control. Wound closure could be accelerated due to higher proliferation not similar motility. These results suggest that pH 7 does not select for increased motility.

As discussed in section 5.4.2.2 there may be a trade-off between motility and proliferation. Cells cultured at pH 7 may display significantly lower motility as they have significantly greater proliferation. Song et al (2006) found that upon culture in acidic pH initial cell growth was reduced but this increased once cells had adapted to the media[285]. They hypothesised that this adaptation may not directly contribute to increased metastatic potential but increase chance of survival once a cell extravasates and reaches a foreign tissue[285]. These results suggest that an acidic environment is not a selective pressure for increased motility however other studies have shown culture at an acidic pH increases cell invasiveness[245 286 287]. Metastasis is a complex multistage process involving multiple cell phenotypes, in this project migration has only been measured *in vitro* in a 2D environment. Measurements of cell migration in 3 D invasive assays conducted on hydrogel layers (table 5.2) or *in vivo* studies may provide a more comprehensive analysis of how acidic pH effects metastasis[245].

5.4.3.3 Selection over 6 months in serum restricted media

Figure 5.22 shows that cells in 5 % FBS had reduced proliferation compared to all other groups. It was not possible to measure the growth rate of cells in 0.5 % FBS however observation of cells in culture noted that they seemed to have a reduced growth rate. As seen in figure 5.24 there is a significant difference between groups in the collective migration of cells, the rate of wound healing is significantly higher in the 5 % group than the control and cells in pH 7. When the differences in proliferation between cells in 5 % FBS and the MEM control are taken into account this could suggest that cell migration is increased in the serum restricted groups. Cells in 5 % FBS have a non-significant higher rate of wound healing than cells in 0.5 % FBS. This is different to the results seen when selection occurred over a 12 week period (section 5.4.2.2) and could be due to the lower N in these experiments. These results support the hypothesis that decreasing FBS concentration increases competition and is a selective pressure for motility.

Chapter 6 – General discussion and conclusions

Studying the evolution of cancer progression is of critical importance to clinical treatment of the disease[10]. Understanding the fundamental principles behind cancer evolution will allow us to study a tumour microenvironment and predict how tumours may adapt or when disease progression may occur[288]. For example, drug resistance in tumours is one of the key examples of how understanding tumour evolution can impact on patient treatment[289]. Tumour evolution depends on heritable differences between cells in traits affecting cell survival or proliferation[129]. Whilst evolutionary processes are acknowledged to play a significant role in the initiation and progression of cancer[9 29 114], the underlying assumption that a variety of cell traits linked to disease progression are heritable within the cell population has never been directly tested. In the course of this project cell motility was shown to be a heritable trait, the first-time heritability has been estimated for a cancer cell trait and one of very few estimates of H^2 in cells of multicellular organisms[142].

This research group has also obtained heritability estimates on the generation times of four cell lines, HeLa, HT1080, MCF7-3 and MDA-2, see appendix B. These H^2 estimates, whilst still significant, were considerably lower than those measured for cell motility. Differences in heritability estimates can have a variety of causes though in general lower H^2 estimates would be expected for a trait closely linked to cell fitness, such as generation time[124], as genetic variation is likely to be reduced by natural selection[124]. These results provide evidence of an important but previously untested assumption that there is heritable variation in cancer cell populations for cell traits[114 115]. Similar methods could be used to estimate evolvability parameters in a range of traits, such as ECM degradation or survival in the circulatory system[114 115]. Gaining heritability estimates in a variety of cancer cell types in a range of traits under different environmental conditions is important as heritability is specific to a population in an environment and so in order to keep testing evolutionary theories in cancer cell lines better understanding is needed of the fundamental principles.

Using an experimental evolution approach has allowed testing of the dispersal theory hypothesis; that when competition for resources is high, selection will favour motile cells[29 38]. The results in chapter 5 indicate that dispersal theory could apply to cancer cells *in vitro* and that serum restriction selects for cell motility (see appendix A). These results may help explain why cancer metastasis repeatedly occurs in solid tumour progression[19 25]. Despite the energy demands of motility and the uncertain advantage to cell fitness, cells able to disperse increase

their own chances of survival as they may migrate to an area with less competition or higher resource concentration[29 38 39]. Kin selection (where selection on individuals favours traits that increase the fitness of close relatives), may also act to maintain a motile phenotype within the population as dispersing cells reduce competition at the original site increasing survival of their relatives [39]. Cancer cells are clonal so a dispersing cell's genome should be preserved within the original population, even if that dispersing cell is unsuccessful, the increased chance of survival for the remaining clonal population increases and the genome is maintained and increased throughout the population[39]. H^2 estimates of cell motility increased for MCF7-1 cell lines cultured under serum restriction for 12 weeks, this implies there was an increase in the genetic contribution to motility, indicating that selection favoured motile genotypes[47]. Obtaining H^2 estimates for other cell traits – such as proliferation, and comparison of these within and between selections lines could be used to investigate potential trade-offs i.e. between cell growth and motility (section 5.4.2).

The results seen in chapter 5 suggest that resource competition is selecting for an increase in the motile phenotype however H^2 estimates for the evolved cell lines also increases. As discussed, this could occur due to an increase in the genetic contribution to motility or a decrease in environmental variation. The complexity of the human genome[61] as well as the genomic instability of cancer cell lines[72 76 139 290] and their sensitivity to environmental conditions[17 163 215 291] means caution must be taken when inferring selection. As discussed in chapter 4, phenotypic plasticity in response to the environment can be modulated via epigenetic changes[154 173 174] and maintained across multiple cell generations[168 169]. Further research is needed to clarify how cancer cells are adapting to their environment and which mechanisms of action they are using to do so. If cells increase their motility as a plastic response to environmental selective pressures, then focus should be on quantifying and controlling the tumour microenvironment[38 292]. If motility and hence metastasis arise as a result of selection on genetic and epigenetic mutations, then preventative treatments might focus on minimising tumour cell population and mutation rates[38 293].

6.1 Implications on studying the evolution of multicellularity

Experimental evolution techniques studying bacteria have already been used to elucidate many of the mechanisms leading to the evolution of multicellularity[294 295]. Use of bacteria as a simpler model system has allowed studies of the core fundamental evolutionary principles such as which selective forces lead to multicellularity, the genetic underpinnings and their

evolutionary history and why multicellularity has evolved multiple times[294 295]. These same techniques could be repeated on cancer cell lines to understand cancer disease progression[11].

Heritability has been measured for proliferation and marker gene expression in Chinese hamster ovaries but otherwise there is almost no data on the extent of heritable phenotypic variation between clonal somatic cells of multicellular animals[142]. Heritability data is available for a few unicellular eukaryotes; growth traits in *Saccharomyces cerevisiae*[143] and six clock traits in *Neurospora crassa*[296]. For both studies geographically disparate strains were crossed to maximise genetic variation but the broad-sense heritability estimates obtained were similar to those gained in this project; 0.77 and 0.42-0.87, indicating that cancer cell lines undergo rapid phenotypic diversification. Measuring the extent of heritable phenotypic variation from cell populations of multicellular organisms has relevance in examining the evolution of multicellularity[11]. The evolution of multicellularity was one of the most significant events in the history of life, it has occurred independently many times across all domains of life, but its initial evolution remains poorly understood[294 295]. Multicellular organisms use cooperation between cells with phenotypic diversification to achieve a higher level of functionality[13 295]. Herron et al (2018) demonstrated that the heritability of a collective trait – such as organism size, depends on the heritability of the individual cells with that organism and that the heritability of the collective trait is usually higher than that of the individual cell trait heritability[297]. There are many mechanisms to ensure and promote cooperation between component cells however as can be seen in the evolution of cancer these controls are not always successful[295]. Whilst cancer cells represent a breakdown in multicellularity, tumour formation involves multicellular cooperation[13]. Obtaining heritability estimates of primary non-cancerous cell lines and comparing these to H^2 estimates of cancer cells may help to answer fundamental questions about selection on individual cells within multicellular organisms.

Multicellular development from a single cell minimizes conflict from mutant cell lineages as they can rarely succeed beyond the lifespan of the multicellular individual[11]. Whilst rare, some cancers have evolved to become contagious between individuals of the same species; Canine transmissible venereal tumour is a sexually transmitted cancer in dogs[298], devil facial tumour disease is a transmissible cancer in Tasmanian devils[299] and transmissible cancers have also been found in some marine bivalves[300]. Heritability estimates are just as important as the strength of selection when predicting evolutionary outcomes[297]. Transmissible cancers represent the 'next step' in cancer progression as tumours can survive the lifespan of the

multicellular organism they derive from. Expanding the research in this project to obtain H^2 estimates for traits in transmissible cancer cells and comparing these to non-transmissible cancer cells may provide insight into disease progression and how the heritability of individual cell traits affects tumour formation.

6.2 The experimental approach

One criticism of *in vitro* experimental evolution is that the laboratory conditions used to study evolution cannot mirror the complexity of a natural environment (in this case *in vivo* tissue) [37]. Whilst an *in vitro* environment will always be less complex than *in vivo*, this simplicity allows environmental control and application of the selective pressures of interest. Environmental complexity can be increased using 3 D cell culture methods such as culture on decellularised tissue. Decellularised tissues contain ECM components and bound signalling molecules meaning *in vitro* culture on such substrates bears more similarities to an *in vivo* environment[301 302]. Further quantitative analysis on *in vivo* tumour cells would be needed to enable the evolutionary dynamics of cancer to be understood and used for clinical applications such as personalised medicine or new therapeutic targets.

Another caveat of *in vitro* cell culture is the frequent bottlenecking of cell populations. In order to maintain cells in culture they must be passaged. This involves removal of all cells from their culture vessel and a random subset of the population being extracted and re-plated onto a new vessel. Whilst attempts are made to minimise the effects of this process by transferring 25 % of the population this may still create variation in the population's genome. Bottlenecking also occurs when experiments are performed, it is not possible to include the entire population of cells in an experiment so again, a random subset of the population is chosen. Only these cells will be measured and used to gather results. Bottlenecking will alter the effects genetic drift and selection have on the cell population[14]. This process is not a naturally occurring one and is another difference between an *in vitro* and *in vivo* environment.

This project does not aim to mimic an *in vivo* tumour microenvironment but to capture the principles of evolution and elucidate the selective pressures favouring motility. Experimental evolution can provide support for a given theory but cannot indicate its relative importance in explaining patterns in nature where many selective forces may be acting simultaneously[37]. Once the simple evolutionary parameters are understood they can be studied and tested *in vivo*[37]. The novel techniques applied here are a proof of concept which can be further adapted for future work into the evolution of cancer metastasis.

Cell lines are used extensively throughout cancer research to study biochemical pathways and processes[293]. The main advantages of *in vitro* cancer models is the level of control on environmental conditions and the reproducibility[7 293]. Almost all continuous cancer cell line cultures are derived from malignant cancers and all the cells lines used in this project were derived from metastatic tumours[68 69 73 77 293]. Cells that have already undergone metastasis may react differently to cells obtained from primary tumours. Cells which have already successfully disseminated have already undergone selection for, and acquired, a metastatic phenotype. One way of ensuring relevance to a wide array of cancer types is to test a broad range of cell lines[64]. This project has shown applicability in results across three distinct cancer types; breast, cervical and fibrosarcoma, but this research could be extended to include cell lines originating from primary tumours as well a range of other cancer types. Acquiring additional cell lines and including a broader range of cancer types in experimental work would be suitable when trying to ascertain mechanistic actions or providing proof of concept. However, long-term robustness in results as well as personalisation to individual patients could be better achieved through use of patient derived cell lines[303].

6.3 Measuring metastasis *in vitro*

The *in vivo* tumour microenvironment is highly complex[17] and metastasis is difficult to measure *in vitro* as it comprises multiple phenotypes[18]. This is further complicated as fitness is measured as the reproductive success of an individual and cannot be measured at a single time point[14 113]. The main limitations of two-dimensional *in vitro* cultures are the phenotypic and genotypic selection that occurs during adaptation to *in vitro* conditions as well as isolation of the cells from an *in vivo* microenvironment[293]. The main advantage of studying cancer cells *in vitro* is that it is a much simpler model system that not only enables precise control of the environment but also allows composite phenotypes, such as metastasis to be broken down into their constituent traits. Using experimental evolution techniques in this simpler model system allows specific selective pressures to be applied and measurement of its effect on the trait of interest[37]. In solid tumours cell motility contributes to cancer metastasis as cells must be able to migrate away from the primary tumour in order to establish secondary, metastatic tumours[21 24]. Measuring a discrete phenotype, such as motility, in an *in vitro* environment allows testing of individual evolutionary hypotheses and direct observation of selection.

Measurement of multiple cell traits and comparing their responses to selection will give clearer insight into how cells respond to their environment. Spheroid cell culture could provide

a more accurate measure of cell migration as the increased complexity of the surrounding environment means a cells biological activity is closer to that seen *in vivo*[249 304]. Including data obtained from spheroid spreading assays will help show whether results gained in a 2 D culture environment bear any relevance to an *in vivo* environment[304].

6.4 Genetics and phenotype

The reduction in costs of sequencing technology makes it possible to gain gene expression data for single cancer cells[305]. Large scale genomic studies (such as The Cancer Genome Atlas) have attempted to catalogue biologically significant mutations within different cancer types[306 307]. Genetic profiling of tumour biopsies has provided a personalised prognostic tool to assist treatment decisions[308], such as patients with HER2 positive breast tumours receiving trastuzumab treatment[309]. Whilst the molecular genetic variation within tumours is known to be substantial, little is known about the molecular basis of phenotypic variation[310 311]. Understanding the molecular basis of phenotype variation has implications in tumour progression and treatment. Khan et al (2018) demonstrated how knowledge of the molecular phenotype of a patient's tumour, anti-EGFR monoclonal antibody resistance conferred via RAS pathway mutations, could be combined with cancer evolutionary dynamics and serial genomic profiling to not only detect whether treatment would be successful but also predict the expected time until treatment failure[289]. Study of the molecular basis of phenotype will allow better understanding of disease progression and combining this data with evolutionary dynamics will allow powerful predictive models on disease treatment and outcome.

Understanding of how genotype and phenotype are connected will provide information on the gene and phenotype under selection[312]. Whilst the cancer genome can be analysed to reveal its temporal history there are problems when inferring causality from a single time point after selection has occurred[57]. Comparisons of gene expression data between cells from the same cell line but with varying trait values, or between evolved and control lines in selection experiments, could reveal possible targets of selection and help link genotype and phenotype together.

6.5 Future directions

Recent papers have shown that using an experimental evolution approach to study cancer is an emerging field[313-315]. Jong et al (2019) argue that experimental evolution of cancer cell lines is a model system to investigate the impact of phenotypic heterogeneity and phenotypic plasticity in response to therapy *in vitro* and is comparable to the phenotypic landscape in a

patient with advanced disease[314]. They exposed adherent and non-adherent cells from a non-small cell lung cancer cell line, NCI-H2122, to an acidic and alkaline pH to generate selection lines that were genetically related but phenotypically distinct and suggest such a system could be used to test the effect of therapies on multiple phenotypes and to design therapies for different phenotypes[314]. This conclusion is supported by Acar et al (2019) who also state that *in vitro* experimental evolution could be used as a model system to study the evolutionary dynamics of cancer with implications for therapy development[315].

Acar et al (2019) used experimental evolution to investigate the effects evolutionary herding has on the evolution of drug resistance[315]. Evolutionary herding involves controlling the tumour cell population to delay or prevent resistance by exploitation of evolutionary trade-offs between resistance and growth[315]. Using HCC827 non-small cell lung cancer cells they demonstrated that the evolutionary dynamics of drug resistance are deterministic and predictable and that in a highly heterogeneous tumour a more effective therapeutic response might be disease management and control aided by modelling of evolutionary dynamics to predict therapeutic response[315].

Lai et al (2019) exposed HCT116 cells, a colorectal cancer cell line to 8 rounds of selection via spheroid forming assays after which a subpopulation became malignant and expressed stem markers, Nanog, Oct4 and Lgr5. Using this experimental evolution approach they identified a potential prognostic biomarker for colorectal cancer, DPEP1[313]. This study has provided proof of concept in the use of experimental evolution in the identification of therapeutic targets.

These recently published papers, along with our own work (appendices A and B) provide a strong foundation for the field of cancer experimental evolution. The techniques developed in this project and the results gained have already (see appendix A) and will continue to contribute to ongoing research in this field.

References

1. Hanahan, D. and Weinberg, R. A. The Hallmarks of Cancer. *Cell* **100,** 57-70 (2000)
2. Cancer Research UK. *Worldwide Cancer Statistics* (CRUK) 2018
3. Coleman, W. B. Chapter 4 – Neoplasia in *Molecular Pathology: The Molecular Basis of Human Disease*. Pages 71-97 Second Edition (Academic Press) 2018
4. Talmadge, J. E. and Fidler, I. J. The biology of cancer metastasis: historical perspective. *Cancer Res.* **70,** 5649-69 (2010)
5. Curtius, K., Wright, N. A. and Graham, T. A. An evolutionary perspective on field cancerization. *Nature Reviews Cancer* **18,** 19-32 (2018)
6. Bernards, R. and Weinberg, R. A. A progression puzzle. *Nature* **418,** 823 (2002)
7. Polyak, K. Breast cancer: origins and evolution. *J Clin Invest.* **117,** 3155-3163 (2007)
8. Marusyk, A. and Polyak, K. Tumor heterogeneity: Causes and consequences. *Biochimica et Biophysica Acta* **1805,** 105-117 (2010)
9. Gerlinger, M., et al. Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. *The New England Journal of Medicine* **366,** 883-892 (2012)
10. Casás-Selves, M. and DeGregori, J. How cancer shapes evolution, and how evolution shapes cancer. *Evolution (N Y)* **4,** 624-634 (2011)
11. Grosberg, R. K. and Strathmann, R. R. The evolution of multicellularity: a minor major transition?. *Annu Rev Ecol Evol Syst.* **38,** 621-654 (2007)
12. Trigos, A. S., Pearson, R. B., Papenfuss, A. T. and Goode, D. L. How the evolution of multicellularity set the stage for cancer. *BJC.* **118,** 145-152 (2018)
13. Aktipis, C. A., et al. Cancer across the tree of life: cooperation and cheating in multicellularity. *Philosophical Transactions Royal Society B* **370,** 20140219 (2015)
14. Futuyma, D. J. *Evolutionary Biology.* 3rd Edition (Sinauer Associates Inc) 1998
15. Forsman, A. Rethinking phenotypic plasticity and its consequences for individuals, populations and species. *Heredity* **115,** 276-284 (2015)
16. Gupta, P. B., et al. Stochastic State Transitions Give Rise to Phenotypic Equilibrium in Populations of Cancer Cells. *Cell* **146,** 633-644 (2011)
17. Bindra, R. S. and Glazer, P. M. Genetic instability and the tumor microenvironment: towards the concept of microenvironment-induced mutagenesis. *Mutation Research* **569,** 75-85 (2005)
18. Klein, C. A. The Metastasis Cascade. *Science* **321,** 1785-1787 (2008)

19. Chen, J., Sprouffske, K., Huang, Q. and Maley, C. C. Solving the Puzzle of Metastasis: The Evolution of Cell Migration in Neoplasms. *Plos One* **6,** e17933 (2011)
20. Urquidi, V. Contrasting Expression of Thrombospondin-1 and Osteopontin Correlates with Absence or Presence of Metastatic Phenotype in an Isogenic Model of Spontaneous Human Breast Cancer Metastasis. *Clinical Cancer Research* **8,** 61-74 (2002)
21. Valastyan, S. and Weinberg, R. A. Tumor Metastasis: Molecular Insights and Evolving Paradigms. *Cell* **147,** 275-292 (2011)
22. Friedly, P. Prespecification and plasticity: shifting mechanisms of cell migration. *Current Opinion in Cell Biology* **16,** 14-23 (2004)
23. Olson, M. F. and Sahai, E. The actin cytoskeleton in cancer cell motility. *Clinical and Experimental Metastasis* **26,** 273 (2009)
24. Palmer, T. D., Ashby, W. J., Lewis, J. D. and Zijlstra, A. Targeting tumor cell motility to prevent metastasis. *Advanced Drug Delivery Reviews* **63,** 568-581 (2011)
25. Hanahan, D. and Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **144,** 646-674 (2011)
26. Tarin, D., et al. Mechanisms of Human Tumor Metastasis Studied in Patients with Peritoneovenous Shunts. *Cancer Res.* **44,** 3584-3592 (1984)
27. Butler, L. M., et al. Suberoylanilide Hydroxamic Acid, an Inhibitor of Histone Deacetylase, Suppresses the Growth of Prostate Cancer Cells in Vitro and in Vivo. *Cancer Res.* **60,** 5165-5170 (2000)
28. Fidler, I. J. Metastasis: Quantitative Analysis of Distribution and Fate of Tumor Emboli Labeled with 1-5-Iodo-2'-deoxyuridine. *J Natl Cancer Inst*. **45,** 773-782 (1970)
29. Merlo, L. M., Pepper, J. W., Reid, B. J. and Maley, C. C. Cancer as an evolutionary and ecological process. *Nature reviews Cancer* **6,** 924-935 (2006)
30. Rainey, P. and Travisano, M. Adaptive radiation in a heterogeneous environment. *Nature* **394,** 69– 72 (1998)
31. Pérez-Ramos, I. M., Matías, L., Gómez-Aparicio, L. and Godoy, Ó. Functional traits and phenotypic plasticity modulate species coexistence across contrasting climatic conditions. *Nature Communications* **10,** 2555 (2019)
32. Siezen, R. J., et al. Phenotypic and genomic diversity of Lactobacillus plantarum strains isolated from various environmental niches. *Environmental Microbiology* **12,** 758-773 (2010)
33. Ramaswamy, S., Ross, K. N., Lander, E. S. and Golub, T. R. A molecular signature of metastasis in primary solid tumors. *Nature Genetics* **33,** 49-54 (2003)

34. Hu, M., et al. Distinct epigenetic changes in the stromal cells of breast cancers. *Nature Genetics* **37,** 899-905 (2005)
35. Yachida, S., et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* **467,** 1114-1119 (2010)
36. Taylor, T. B., Johnson, L. J., Jackson, R. W., Brochurst, M. A. and Dash, P. R. First steps in experimental cancer evolution. *Evol Appl.* **6,** 535-548 (2013)
37. Buckling, A., Maclean, C. R., Brockhurst, M. A. and Colegrave, N. The Beagle in a bottle. *Nature* **12,** 824-829 (2009)
38. Taylor, T. B., Wass, A. V., Johnson, L. J. and Dash, P. Resource competition promotes tumour expansion in experimentally evolved cancer. *BMC Evolutionary Biology* **17,** 268-277 (2017)
39. Bowler, D. E. and Benton, T. G. Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological Reviews* **80,** 205-225 (2005)
40. Taylor, T. B. and Buckling, A. Competition and Dispersal in *Pseudomonas aeruginosa. The American Naturalist* **176,** 83-89 (2010)
41. Hamilton, W. D. and May, R. M. Dispersal in stable habitats. *Nature* **269,** 578-581 (1977)
42. Lowe, W. H. What drives long-distance dispersal? A test of theoretical predictions. *Ecology* **90,** 1456-1462 (2009)
43. Spaan, R. S., Epps, C. W., Ezenwa, V. O. and Jolles, A. E. Why did the buffalo cross the park? Resource shortages, but not infections, drive dispersal in female African buffalo (*Syncerus caffer*). *Ecology and Evolution* **9,** 5651-5663 (2019)
44. Aktipis, C. A., Maley, C. C. and Pepper, J. W. Dispersal evolution in neoplasms: the role of disregulated metabolism in the evolution of cell motility. *Cancer Prev Res*. **5,** 266-275 (2012)
45. Amend, S.R., Gatenby, R.A., Pienta, K.J. and Brown, J.S. Cancer foraging ecology: diet choice, patch use, and habitat selection of cancer cells. *Current Pathobiology Reports* **6,** 209-218 (2018).
46. Wray, N. R. & Visscher, P. M. Estimating trait heritability. *Nature Education* **1,** 29 (2008)
47. Fletcher, H. & Hickey, I. *Genetics Fourth Edition*. Ch 8. (Garland Science) 2013.
48. Kawecki, T. J., et al. Experimental Evolution. *Trends in Ecology & Evolution* **27,** 547-560 (2012)
49. Graham, T. A. and Sottoriva, A. Measuring cancer evolution from the genome. *The Journal of Pathology* **241,** 183-191 (2016)

50. Woo, Y. H. and Li, W. H. DNA replication timing and selection shape the landscape of nucleotide variation in cancer genomes. *Nature Communications* **3,** 1004 (2012)
51. Lenski, R. E., Rose, M. R. Simpson, S. C. and Tadler, S. C. Long-Term Experimental Evolution in *Escherichia coli*. I. Adaptation and Divergence. *The American Naturalist* **138,** 1315-1341 (1991)
52. Lenski, R. E. and Burnham, T. C. Experimental evolution of bacteria across 60,000 generations, and what it might mean for economics and human decision-making. *Biochem.* **20,** 107-124 (2018)
53. Novak, M., Pfeiffer, T., Lenski, R. E., Sauer, U. and Bonhoeffer, S. Experimental Tests for an Evolutionary Trade-Off between Growth Rate and Yield in *E. coli. The American Naturalist* **168,** 242– 251 (2006)
54. Barrick, J. E., et al. Genome evolution and adaptation in a long-term experiment with *Escherichia coli. Nature* **461,** 1243– 1247 (2009)
55. Barrett, R. D. H., et al. Rapid evolution of cold tolerance in stickleback. *Proc Biol Sci.* **278,** 233-238 (2011)
56. Turner, P. E. and Chao, L. Prisoner's dilemma in an RNA virus. *Nature* **398,** 441– 443 (1999)
57. Swallow, J. G., Carter, P. A. and Garland Jr, T. Artificial Selection for Increased Wheel-Running Behaviour in House Mice. *Behaviour Genetics* **28,** 227-237 (1998)
58. Rose, M. R. Artificial selection on a fitness-component in *Drosophila Melanogaster. Evolution* **38,** 516-526 (1984)
59. Elena, S. F. and Lenski, R. E. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nature Review Genetics* **4,** 457-469 (2003)
60. Arrigoni, C., Bersini, S., Gilardi, M. and Moretti, M. In Vitro Co-Culture Models of Breast Cancer Metastatic Progression towards Bone. *International Journal of Molecular Sciences* **17,** 1405 (2016)
61. Cooper, G. M. *The Cell: A Molecular Approach.* 2nd edition (Sinauer Associates) 2000
62. Ajani, J. A. Challenges imposed by the complexity of cancer genome. *The Lancet Oncology* **14,** Pe291-e292 (2013)
63. Opasic, L., Zhou, D., Werner, B., Dingli, D. and Traulsen, A. How many samples are needed to infer truly clonal mutations from heterogenous tumours? *BMC Cancer.* **19,** 403 (2019)
64. Freshney, R. I. *Culture Of Animal Cells: A Manual Of Basic Technique And Specialized Applications.* Sixth Edition (Wiley) 2010
65. Bodnar, A. G., et al. Extension of Life-Span by Introduction of Telomerase into Normal Human Cells. *Science* **279,** 349-352 (1998)

66. Cancer Research UK. *Breast Cancer Statistics*. (Breast Cancer Research UK) 2015
67. Carrel, A. and Burrows, M. T. Cultivation of Tissues in vitro and its Technique. *JEM.* **13,** 387-396 (1911)
68. Cailleau, R., Young, R., Olive, M. and Reeves, W. J. Breast tumor cell lines from pleural effusions. *J Natl Cancer Inst.* **53,** 661-674 (1974)
69. Soule, H. D., Vasquez, J., Long, S., Albert, S. and Brennan, M. A human cell line from a pleural effusion derived from a breast carcinoma. *J Natl Cancer Inst.* **51,** 1409-1416 (1973)
70. Holliday, D. L. and Speirs, V. Choosing the right cell line for breast cancer. *Breast Cancer Research* **13,** 215 (2011)
71. Lee, A. V., Oesterreich, S. and Davidson, N. E. MCF-7 Cells – Changing the Course of Breast Cancer research and Care for 45 Years. *JNCI*. **107,** djv073 (2015)
72. Osborne, C. K., Hobbs, K. and Trent, J. M. Biological differences among MCF-7 human breast cancer cell lines from different laboratories. *Breast Cancer Res Treat.* **9,** 111-121 (1987)
73. Landry, J. J. M., et al. The Genomic and Transcriptomic Landscape of a HeLa Cell Line. *G3: Genes, Genomes, Genetics* **3,** 1213-1224 (2013)
74. Gey, G. O., Coffman, W. D. and Kubicek, M. T. Tissue culture studies of the proliferative capacity of cervical carcinoma and normal epithelium. *Cancer Res.* **12,** 264-265 (1952)
75. Bulzomi, P. et al. The pro-apoptotic effect of quercetin in cancer cell lines requires ERβ – dependent signals. *Journal of Cellular Physiology* **227,** 1891-1898 (2012)
76. Stephens, P. J., et al. Massive Genomic Rearrangement Acquired in a Single Catastrophic Event during Cancer Development. *Cell* **144,** 27-40 (2011)
77. Rasheed, S., Nelson-Rees, W. A., Toth, E. A., Arnstein, P. & Gardner, M. B. Characterization of a newly derived human sarcoma cell line (HT-1080). *Cancer* **33,** 1027-1033 (1974)
78. Scherer, W. F., Syverton, J. T. and Gey, G. O. Studies on the Propagation in vitro of Poliomyelitis Viruses. *J Exp Med.* **97,** 695-710 (1953)
79. Prat, A., et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Research* **12,** R68 (2010)
80. Eagle, H. Nutrition Needs of Mammalian Cells in Tissue Culture. *Science* **122,** 501-504 (1955)
81. Eagle, H. The Specific Amino Acid Requirements of a Mammalian Cell (Strain L) in Tissue Culture. *J Biol Chem.* **214,** 839-852 (1955)

82. Denizot, F., and Lang, R. Rapid colorimetric assay for cell growth and survival: Modifications to the tetrazolium dye procedure giving improved sensitivity and reliability. *Journal of Immunological Methods* **89,** 271-277 (1986)
83. Berthois, Y., Katzenellenbogen, J. A., and Katzenellenbogen, B. S. Phenol red in tissue culture media is a weak estrogen: Implications concerning the study of estrogen-responsive cells in culture. *PNAS*. **83**, 2496-2500 (1986)
84. Rous, P. and Jones, F. S., A Method for Obtaining Suspensions of Living Cells from the Fixed Tissues and for the Plating Out of Individual Cells. *J Exp Med*. **23,** 549-555 (1916)
85. Riss, T. L., et al. *Cell Viability Assays* (Bethesda) 2016
86. Liu, Y., Peterson, D. A., Kimura, H. and Schubert, D. Mechanism of Cellular 3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide (MTT) Reduction. *Journal of Neurochemistry* **69,** 581-593 (1997)
87. Stephens, D. J. and Allan, V. J. Light Microscopy Techniques for Live Cell Imaging. *Science* **300,** 82-86 (2003)
88. Vybrant® MTT Cell Proliferation Assay Kit (V-13154) Molecular Probes (2002)
89. Twentyman, P. R. and Luscombe, M. A study of some variables in a tetrazolium dye (MTT)
90. Yokoyama, W. M., Thompson, M. L. and Ehrhardt, R. O. Cryopreservation and Thawing of Cells. *Curr Protoc Immunol* **99,** A.3G.1-A.3G.5 (2012)
91. Brand, K.G. and Syverton, J. T. Results of Species-Specific Hemagglutination Tests on "Transformed," Nontransformed, and Primary Cell Cultures. *JNCI.* **28,** 147-157 (1962)
92. Kniss, D. A. and Summerfield, T. L. Discovery of HeLa Cell Contamination in HES Cells. *Reprod Sci.* **21,** 1015-1019 (2014)
93. Kuhlmann, I. The prophylactic use of antibiotics in cell culture. *Cytotechnology* **19,** 95-105 (1995)
94. Drexler, H. G. and Uphoff, C. C. Mycoplasma contamination of cell cultures: Incidence, sources, effects, detection, elimination, prevention. *Cytotechnology* **39,** 75-90 (2002)
95. Marx, V. Cell-line authentication demystified. *Nature Methods* **11,** 483-488 (2014)
96. Carter, M., and Shieh, J. *Chapter 5 – Microscopy* in *Guide to Research Techniques in Neuroscience* Second Edition (Academic Press 2015)
97. Schindelin, J., et al. Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9,** 676-682 (2012)
98. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9,** 671-675 (2012).

99. Meijering, E., Dzyubachyk, O. and Smal, I. Methods for Cell and Particle Tracking. *Methods in Enzymology* **504,** 183-200 (2012)
100. Collins, T. J. ImageJ for microscopy. *BioTechniques* **43,** S25-S30 (2007)
101. Meijering, E. MTrackJ: An ImageJ Plugin for Motion Tracking and Analysis https://imagescience.org/meijering/software/mtrackj/
102. IBM Corp. Released 2016. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp.
103. Minitab 18 Statistical Software (2017) Minitab Inc. www.minitab.com
104. Pallant, J. *SPSS Survival Manual 4th Edition Ch. 3* (Open University Press) 2010
105. Packard, G. C. Unanticipated consequences of logarithm transformation in bivariate allometry. *Journal of Comparative Physiology B.* **181,** 841-849 (2011)
106. Kruuk, L. E., Slate, J. and Wilson, A. J. New answers for old questions: the evolutionary quantitative genetics of wild animal populations. *Annu Rev Ecol Evol* S. **39,** 525-548 (2008)
107. Feng, C. et al. Log-transformation and its implications for data analysis. *Shanghai Archives of psychiatry* **26,** 105-109 (2014)
108. Gillet, J. P., Varma, S. and Gottesman, M. M. The Clinical Relevance of Cancer Cell Lines*. J Natl Cancer Inst.* **105,** 452-458 (2013)
109. Zhang, J., Cunningham, J. J., Brown, J. S. and Gatenby, R. A. Integrating evolutionary dynamics into treatment of metastatic castrate-resistant prostate cancer. *Nat Commun.* **8,** 1816 (2017)
110. Darwin, C. *The Origin of Species; and, The Voyage of the Beagle/Charles Darwin; with an introduction by Richard Dawkins.* (Everyman's Library) 2003
111. Fisher, R. A. *The Genetical Theory of Natural Selection* (Dover Publications, Inc.) 1930
112. Greaves, M. and Maley, C. C. Clonal evolution in cancer. *Nature* **481,** 306-313 (2012)
113. Barton, N. H., Briggs, D. E. G., Eisen, J. A., Goldstein, D. B. and Patel, N. H. *Evolution* (Cold Spring Harbor Laboratory Press) 2007
114. Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194,** 23-28 (1976).
115. Aktipis, C. and Nesse, R. M. Evolutionary foundations for cancer biology. *Evol Appl.* **6,** 144-159 (2013)
116. Hatzikirou, H., Basanta, D., Simon, M., Schaller, K. and Deutsch, A. 'Go or Grow': the key to the emergence of invasion in tumour progression? *Math Med Biol.* **29,** 49-65 (2012)
117. Lee, H., et al. Evolution of Tumour Invasiveness: The Adaptive Tumour Microenvironment Landscape Model. *Cancer Res*. **71,** 6327 – 6337 (2011)

118. Wright, S. The Relative Importance Of Heredity And Environment In Determining The Piebald Pattern Of Guinea-Pigs. *Proc Natl Acad Sci.* **6,** 320-332 (1920)

119. Fisher, R. A. The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Trans Roy Soc Edin.* **52,** 399 – 433 (1918)

120. Guthrie, S. and Pidduck, H. G. Heritability of elbow osteochondrosis within a closed population of dogs. *Journal of Small Animal Practice* **31,** 93-96 (1990)

121. Kosova, G., Abney, M. and Ober, C. Heritability of reproductive fitness traits in a human population. *PNAS.* **107,** 1772-1778 (2010)

122. Letcher, B. H., Coombs, J. A. and Nislow, K. H. Maintenance of phenotypic variation: repeatability, heritability and size-dependent processes in a wild brook trout population. *Evolutionary Applications* **4,** 602-615 (2011)

123. Loeske, E. B., et al. Heritability of fitness in a wild mammal population. *PNAS*. **97,** 698-703 (2000)

124. Merilä, J. and Sheldon, B. C. Lifetime Reproductive Success and Heritability in Nature. *The American Naturalist* **155,** 301-310 (2000)

125. Scwaegerle, K. E. and Levin, D. A. Quantitative Genetics of Fitness Traits in a Wild Population of Phlox. *Evolution* **45,** 169-177 (1991)

126. Falconer, D. S. and Mackay T. F. C. *Introduction to Quantitative Genetics* (Longman, Harlow) 1996

127. Hill, W. G., Goddard, M. E. and Visscher, P. M. Data and Theory Point to Mainly Additive Genetic Variance for Complex Traits. *PLoS Genetics* **4,** e1000008 (2008)

128. Byers, D. L. Components of phenotypic variance. *Nature Education* **1,** 161 (2008)

129. Visscher, P. M., Hill, W. G. and Wray, N. R. Heritability in the genomics era – concepts and misconceptions. *Nature Reviews Genetics* **9,** 255-256 (2008)

130. Griffiths, A. J. F., Miller, J. H., Suzuki, D. T., Lewontin R. C. and Gelbart, W. M. *An Introduction to Genetic Analysis.* 7th Edition Heritability of a trait. (W. H. Freeman) 2000

131. Spencer, S. L., Gaudet, S., Albeck, J. G., Burke, J. M. and Sorger, P. K. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **459,** 428-432 (2009)

132. Navin, N., et al. Tumour evolution inferred by single-cell sequencing. *Nature* **472,** 90-95 (2011)

133. Nik-Zainal, S., et al. The Life History of 21 Breast Cancers. *Cell* **149,** 994 – 1007 (2012)

134. Shah, S. P., et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486,** 395-359 (2012)

135. Schlappack, O. K., Zimmermann, A. and Hill, R. P. Glucose starvation and acidosis: effect on experimental metastatic potential, DNA content and MTX resistance of murine tumour cells. *Br J Cancer.* **64,** 663-670 (1991)

136. Bhadriraju, K., et al. Large-scale time-lapse microscopy of Oct4 expression in human embryonic stem cell colonies *Stem Cell Research.* **17,** 122-129 (2016)

137. Nikon. *Imaging Software NIS Elements.* (Nikon Corporation) 2017

138. Boddy, A. M., Huang, W. and Aktipis, A. Life History Trade-Offs in Tumours. *Current Pathobiology Reports* **6,** 201-207 (2018)

139. Torsvik, A., et al. U-251 revisited: genetic drift and phenotypic consequences of long-term cultures of glioblastoma cells. *Cancer Med.* **3,** 812-824 (2014)

140. Jackson, A. L. and Lawrence, A. L. The Mutation Rate and Cancer. *Genetics* **148**, 1483-1490 (1993)

141. Kaden, D., Gadi, I. K., Bardwell, L., Gelman, R. and Sager, R. Spontaneous Mutation Rates of Tumorigenic and Nontumorigenic Chinese Hamster Embryo Fibroblast Cell Lines. *Cancer Research* **49,** 3374-3379 (1989)

142. Duesberg, P., Stindl, R. and Hehlmann, R. Explaining the high mutation rates of cancer cells to drug and multidrug resistance by chromosome reassortments that are catalyzed by aneuploidy. *PNAS* **97,** 14295-14300 (2000)

143. Adey, A., et al. The haplotype-resolved genome and epigenome of the aneuploid HeLa cancer cell line. *Nature* **500,** 207-211 (2013)

144. Dinalankara, W. and Bravo, H. C. Gene Expression Signatures Based on Variability can Robustly Predict Tumor Progression and Prognosis. *Cancer Informatics* **14,** 71-81 (2015)

145. Andor, N., et al. Pan-cancer analysis of the extent and consequences of intratumor heterogeneity. *Nature Medicine* **22,** 105-113 (2016)

146. Davies, S. L., et al. Functional heterogeneity and heritability in CHO cell populations. *Biotechnology and Bioengineering* **110,** 260-274 (2013)

147. Bloom, J. S., Ehrenreich, I. M., Loo, W., Võ Lite, T. and Kruglyak, L. Finding the sources of missing heritability in a yeast cross. *Nature* **494,** 234-237 (2013)

148. Wu, A., et al. Cell motility and drug gradients in the emergence of resistance to chemotherapy. *PNAS.* **40,** 16103–16108 (2013).

149. Díaz Arenas, C. and Cooper, T. F. Mechanism and selection of evolvability: experimental evidence. *FEMS Microbiology Reviews.* **37,** 572-582 (2013)

150. Colegrave, N. and Collins, S. Experimental evolution: experimental evolution and evolvability. *Heredity* **100,** 464-470 (2008)

151. Shaver, A. C., et al. Fitness evolution and the rise of mutator alleles in experimental Escherichia coli populations. *Genetics* **162,** 557-566 (2002)

152. Araten, D. J., et al. A Quantitative Measurement of the Human Somatic Mutation Rate. *Cancer Res.* **65,** 8111-8117 (2005)

153. Nachman, M. W. and Crowell, S. L. Estimate of the Mutation Rate per Nucleotide in Humans. *Genetics* **156,** 297-304 (2000)

154. Werner, B. and Sottoriva, A. Variation of mutational burden in healthy human tissues suggests non-random strand segregation and allows measuring somatic mutation rates. *PLoS Computational Biology* **14,** e1006233 (2018)

155. Berger, S. L, Kouzarides, T., Shiekhattar, R. and Shilatifard, A. An operational definition of epigenetics. *Genes Dev.* **23,** 781-783 (2009)

156. Trerotola, M., Relli, V., Simeone, P. and Alberti, S. Epigenetic inheritance and the missing heritability. *Human Genomics* **9,** 17-29 (2015)

157. Soo-You, J. and Jones, P. A. Cancer Genetics and Epigenetics: Two Sides of the Same Coin? *Cancer Cell* **22,** 9-20 (2012)

158. Dawson, M. A. and Kouzarides, T. Cancer Epigenetics: From Mechanism to Therapy. *Cell* **150,** 12-27 (2012)

159. Junttila, M. R. and De Sauvage, F. J. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature* **501,** 346-354 (2013)

160. Lewis, C. E. and Pollard, J. W. Distinct role of macrophages in different tumor microenvironments. *Cancer Res.* **66,** 605-612 (2006)

161. Rosato, R. R., Almenara, J. A. and Grant, S. The Histone Deacetylase Inhibitor MS-275 Promotes Differentiation or Apoptosis in Human Leukemia Cells through a Process Regulated by Generation of Reactive Oxygen Species and Induction of $p21^{CIP1/WAF1\ 1}$. *Cancer Res.* **63,** 3637-3645 (2003)

162. Kruidenier, L., et al. A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. *Nature* **488,** 404-408 (2012)

163. Burrell, R. A., McGranahan, N., Bartek, J. and Swanton, C. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* **501,** 338-345 (2013)

164. Graaf, A., et al. Rate, spectrum, and evolutionary dynamics of spontaneous epimutations. *PNAS.* **112,** 6676-6681 (2015)

165. Waddington, C. H. The epigenotype. *Endeavour* **1,** 18–20 (1942).

166. Burggren, W. Epigenetic Inheritance and Its Role in Evolutionary Biology: Re-Evaluation and New Perspectives. *Biology (Basel).* **5,** 24 (2016)

167. Johnson, L. J. and Tricker, P. J. Epigenomic plasticity within populations: its evolutionaery significance and potential. *Heredity* **105,** 113-121 (2010)

168. Jablonka, E. and Raz, G. Transgenerational Epigenetic Inheritance: Prevalence, Mechanisms, and Implications for the Study of Heredity and Evolution. *The Quarterly Review of Biology* **84,** 131-176 (2009)

169. Lind, M. I. and Spagopoulou, F. Evolutionary consequences of epigenetic inheritance. *Heredity* **121,** 205-209 (2018)

170. Cubas, P., Vincent, C. and Coen, E. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* **401,** 157-161 (1999)

171. Payne, S., McCarthy, S., Johnson, T., North, E. and Blum, P. Nonmutational mechanism of inheritance in the Archaeon *Sulfolobus solfataricus. PNAS.* **115,** 12271-12276 (2018)

172. Zhang, Y., Latzel, V., Fischer, M. and Bossdorf, O. Understanding the evolutionary potential of epigenetic variation: a comparison of heritable phenotypic variation in epiRILs, RILs, and natural ecotypes of *Arabidopsis thaliana. Heredity* **121,** 257-265 (2018)

173. Geng, Y., Gao, L. and Yang, J. Epigenetic Flexibility Underlying Phenotypic Plasticity In: Lüttge U., Beyschlag W., Francis D., Cushman J. (eds) Progress in Botany. *Progress in Botany* vol 74. (Springer, Berlin, Heidelberg) 2013

174. Duncan, E. J., Gluckman, P. D. and Dearden, P. K. Epigenetics, plasticity, and evolution: How do we link epigenetic change to phenotype? *JEZ-B Molecular and Developmental Biology* **322,** 208-220 (2014)

175. Allis, C. D. and Jenuwein, T. The molecular hallmarks of epigenetic control. *Nature Reviews Genetics* **17,** 487-500 (2016)

176. Maley, C. C., et al. Classifying the evolutionary and ecological features of neoplasms. *Nature Reviews Cancer* **17,** 605-619 (2017)

177. Yang, H., Yan, B., Liao, D., Huang, S. and Qiu, Y. Acetylation of HDAC1 and degradation of SIRT1 form a positive feedback loop to regulate p53 acetylation during heat-shock stress. *Cell Death and Disease* **6,** e1747 (2015)

178. Deng, S., et al. Hypoxia-induced LncRNA-BX111 promotes metastasis and progression of pancreatic cancer through regulating ZEB1 transcription. *Oncogene* **37,** 5811-5828 (2018)

179. Garcia, B. A., Shabanowitz, J. and Hunt, D. F. Characterization of histones and their post-translational modifications by mass spectrometry. *Current Opinion in Chemical Biology* **11,** 66-73 (2007)

180. Lewin, B. *Genes IX* (Jones and Bartlett Publishers) 2008

181. Di Cerbo, V. and Schneider, R. Cancers with wrong HATs: the impact of acetylation. *Briefings in Functional Genomics* **12,** 231-243 (2013)

182. Gangisetty, O., Cabrera, M. A. and Murugan, S. Impact of epigenetics in aging and age related neurodegenerative diseases. *Frontiers in Bioscience* **23,** 1445-1464 (2018)

183. Adams, R. L. P. and Lindsay, H. What is hemimethylated DNA? *FEBS.* **320,** 243-245 (1993)

184. Tatamiya, T., Saito, A., Sugawara, T. and Nakanishi, O. Isozyme-selective activity of the HDAC inhibitor MS-275. *American Association for Cancer Research* **64,** 2451 (2004)

185. McCabe, M. T., Brandes, J. C. and Vertino, P. M. Cancer DNA Methylation: Molecular Mechanisms and Clinical Implications. *Clin Cancer Res.* **15,** 3927-3937 (2009)

186. Klutstein, M., Nejman, D., Greenfield, R. and Cedar, H. DNA Methylation in Cancer and Aging. *Cancer Research* **76,** 3446-3450 (2016)

187. Steensel, B. and Belmont, A. S. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* **169,** 780-791 (2017)

188. Gkountela, S., et al. Circulating Tumor Cell Clustering Shapes DNA Methylation to Enable Metastasis Seeding. *Cell* **176,** 98-112 (2019)

189. Chimonidou, M., et al. DNA Methylation of Tumor Suppressor and Metastasis Suppressor Genes in Circulating Tumor Cells. *Clinical Chemistry* **57,** 1169-1177 (2011)

190. Kulis, M. and Esteller, M. 2- DNA Methylation and Cancer. *Advances in Genetics* **70,** 27-56 (2010)

191. Das, P. M. and Singal, R. DNA Methylation and Cancer. *Journal of Clinical Oncology* **22,** 4632-4642 (2004)

192. Pilka, E. S., James, T. and Lisztwan, H. J. Structural definitions of Jumonji family demethylase selectivity. *Drug Discovery Today* **20,** 743-749 (2015)

193. Pederson, M. T. and Helin, K. Histone demethylases in development and disease. *Trends in Cell Biology* **20,** 662-671 (2010)

194. Hoffmann, I., et al. The role of histone demethylases in cancer therapy. *Molecular Oncology* **6,** 683-703 (2012)

195. Park, S. Y., Park, J. W. and Chun, Y. S. Jumonji histone demethylases as emerging therapeutic targets. *Pharmacological Research* **105,** 146-151 (2016)

196. Sadoul, K., Boyault, C., Pabion, M. and Khochbin, S. Regulation of protein turnover by acetyltransferases and deacetylases. *Biochimie* **90,** 306-312 (2008)

197. Heinemann, B., et al. Inhibition of demethylases by GSK-J1/J4. *Nature* **514,** E1-E2 (2014)

198. Wang, J. D., et al. The histone demethylase UTX enables RB-dependent cell fate control. *Genes & Dev.* **24,** 327-332 (2010)

199. Siegel. P. M. and Massagué, J. Cytostatic and apoptotic actions of TGF-β in homeostasis and cancer. *Nature Reviews Cancer* **3,** 807-820 (2003)

200. Macias, M. J., Martin-Malpartida, P. and Massagué, J. Structural determinants of Smad function in TGF-β signalling. *Trends in Biochemical Sciences* **40,** 296-308 (2015)

201. Bang, J., et al. Cell Proliferation and Motility Are Inhibited by G1 Phase Arrest in 15-kDa Selenoprotein-Deficient Chang Liver Cells. *Mol Cells.* **38,** 457-465 (2015)

202. Kagawa, Y., et al. Cell Cycle-Dependent Rho GTPase Activity Dynamically Regulates Cancer Cell Motility and Invasion In Vivo. *PLoS One* **8,** e83629 (2013)

203. Na, J., et al. Histone H3K27 Demethylase JMJD3 inCooperation with NF-κB RegulatesKeratinocyte Wound Healing. *Journal of Investigative Dermatology* **136,** 847-858 (2016)

204. Knapinska, A. M., Estrada, C. A. and Fields, G. B. The Roles of Matrix Metalloproteinases in Pancreatic Cancer. *Progress in Molecular Biology and Translational Science* **148,** 339-354 (2017)

205. Reunanen, N. and Kähäri, V. Matrix Metalloproteinases in Cancer Cell Invasion.*Madam Curie Biosciences Database* (Landes Bioscience) 2013

206. De Ruijter, A. J. M., Van Gennip, A. H., Caron, H. N., Kemp, S. and Van Kuilenburg, A. B. P. Histone deacetylases (HDACs) : characterization of the classical HDAC family. *Biochem J.* **370,** 737-749 (2003)

207. Kim, H. J. and Bae, S. C. Histone deacetylase inhibitors: molecular mechanisms of action and clinical trials as anti-cancer drugs. *Am J Transl Res.* **3,** 166-179 (2011)

208. Martin, G. S. Cell signaling and cancer *Cancer. Cell* **4,** 167-174 (2003)

209. Saito, A., et al. A synthetic inhibitor of histone deacetylase, MS-27-275, with marked in vivo antitumor activity against human tumors. *PNAS.* **13,** 4592-4597 (1999)

210. Dokmanovic, M., Clarke, C. and Marks, P. A. Histone Deacetylase Inhibitors: Overview and Perspectives. *Molecular Cancer Research* **5,** 981-989 (2007)

211. Miller, T. A., Witter, D. J. and Belvedere, S. Histone Deacetylase Inhibitors. *J Med Chem.* **46,** 5097-5116 (2003)

212. Ghoshdastider, U., Popp, D., Burtnick, L. D. and Robinson, R. C. The expanding superfamily of gelsolin homology domain proteins. *Cytoskeleton* **70,** 775-795 (2013)

213. Gimona, M., Buccione, R., Courtneidge, S. A. and Linder, S. Assembly and biological role of podosomes and invadopodia. *Current Opinion in Cell Biology* **20,** 235-241 (2008)

214. Gururaj, A., Rayala, S. K. and Kumar, R. p21-activated kinase signaling in breast cancer. *Breast Cancer Research* **7,** 5-12 (2004)

215. Ferrao, P. T., Behren, A., Anderson, R. L. & Thompson, E. W. Editorial: Cellular and Phenotypic Plasticity in Cancer. *Front Oncol.* **5,** 171 (2015)

216. Hedley, D. W. and Chow, S. Evaluation of methods for measuring cellular glutathione content using flow cytometry. *Cytometry* **15,** 349-358 (1994)

217. Invitrogen. *Molecular Probes Handbook.* (Thermo Fisher Scientific) 2010

218. Riddle, M. R., Boesmans, W., Caballero, O., Kazwiny, Y. and Tabin, C. J. Morphogenesis and motility of the *Astyanax mexicanus* gastrointestinal tract. *Developmental Biology* **441,** 285-296 (2018)

219. Wang, X., et al. Engineering anastomosis between living capillary networks and endothelial cell-lined microfluidic channels. *Lab Chip.* **16,** 282-290 (2016)

220. Bertulli, C., et al. Image-Assisted Microvessel-on-a-Chip Platform for Studying Cancer Cell Transendothelial Migration Dynamics. *Sci Rep.* **8,** 12480 (2018)

221. Heinze, T., Liebert, T., Heublein, B. and Hornig, S. *Polysaccharides II. Advances in Polymer Science.* Function Polymers Based on Dextran. (Springer) 2006

222. Wang, L., et al. A small molecule modulates Jumonji histone demethylase activity and selectively inhibits cancer growth. *Nature Communications* **4,** 2035 (2013)

223. Carthew, R. W. and Sontheimer, E. J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **136,** 642-655 (2010)

224. Chen, G. and Deng, X. Cell Synchronization by Double Thymidine Block. *Bio-protocol* **8,** e2994 (2018)

225. Banfalvi, G. Chapter 1: Cell Cycle Synchronization.*Methods in Molecular Biology* (Springer Science) 2011

226. Sawicka, A. and Seiser, C. Sensing core histone phosphorylation — A matter of perfect timing. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1839,** 711-718 (2014)

227. Rossetto, D., Avvakumov, N. and Côté, J. Histone phosphorylation. *Epigenetics* **7,** 1098-1108 (2012)

228. Bartlett, J. M. S. ad Stirling, D. A Short History of the Polymerase Chain Reaction. *Methods in Molecular Biology* **226,** 3- 6 (2003)

229. Johnson, D. S., Mortazavi, A., Myers, R. M. and Wold, B. Genome-Wide Mapping of in Vivo Protein-DNA Interactions. *Science* **316,** 1497-1502 (2007)

230. Brunner, D., et al. Serum-free Cell Culture: The Serum-free Media Interactive Online Database. *ALTEX.* **27,** 53-62 (2010)

231. Cressie, N. *Statistics for Spatial Data* (Wiley Interscience) 1993

232. Xavier, J. B., Martinez-Garcia, E. and Foster, K. R. Social evolution of spatial patterns in bacterial biofilms: when conflict drives disorder. *Am Nat.* **174,** 1-12 (2009)

233. Bonte, D., et al. Costs of Dispersal. *Biological Reviews* **87,** 290-312 (2012)

234. Ljungfeldt, L. E. R., Espedal, P. G., Nilsen, F., Skern-Mauritzen, M. and Glover, K. A. A common-garden experiment to quantify evolutionary processes in copepods: the case of emamectin benzoate resistance in the parasitic sea louse *Lepeophtheirus salmonis. BMC Evolutionary Biology* **14,** 108 (2014)

235. Wiser, M. J. and Lenski, R. E. A Comparison of Methods to Measure Fitness in Escherichia coli. *PLos One* **10,** e0126210

236. Rabinovitch, P. S., June, C. H. and Kavanagh, T. J. Introduction to Functional Cell Assays. *Annals of the New York Academy of Sciences* **677,** 252-264 (1993)

237. Chen, Y. Scratch Wound Healing Assay. *Bio-Protocol* **2,** e100 (2012)

238. Lampugnani, M. G. Cell migration into a wounded area *in vitro. Methods in Molecular Biology* **96,** 177-182 (1999)

239. Haycock, J.W. 3D Cell Culture: A Review of Current Approaches and Techniques. *Methods in Molecular Biology* **695,** 1-15 (2011)

240. Hirschhaeusera, F., et al. Multicellular tumor spheroids: An underestimated tool is catching up again. *Journal of Biotechnology* **148,** 3-15 (2010)

241. Hamilton, G. Multicellular spheroids as an in vitro tumor model. *Cancer Letters* **131,** 29-34 (1998)

242. Liberti, M. V. and Locasale, J. W. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem Sci.* **41,** 211-218 (2016)

243. Van der Valk, J., et al. Optimization of chemically defined cell culture media – Replacing fetal bovine serum in mammalian in vitro methods. *Toxicology in Vitro* **24,** 10053-1063 (2010)

244. Young, S. D., Marshall, R. S. and Hill, R. P. Hypoxia induces DNA overreplication and enhances metastatic potential of murine tumor cells. *PNAS.* **85,** 9533-9537 (1988)

245. Rofstad, E. K. Mathiesen, B., Kindem, K. and Galappathi, K. Acidic Extracellular pH Promotes Experimental Metastasis of Human Melanoma Cells in Athymic Nude Mice. *AACR.* **66,** 6699-6707 (2006)

246. Vamvakidou, A. P., et al. Heterogeneous Breast Tumoroids: An In Vitro Assay for Investigating Cellular Heterogeneity and Drug Delivery. *Journal of Biomolecular Screening* **12,** 13-20 (2007)

247. Somaweera, H., Ibraguimov, A. and Pappas, D. A review of chemical gradient systems for cell analysis. *Analytica Chimica Acta* **907,** 7-17 (2016)

248. Keenan, T. M. and Folch, A. Biomolecular gradients in cell culture systems. *Lab on a Chip* **8,** 34-57 (2008)

249. Fennema, E., Rivron, N., Rouwkema, J., Van Blitterswijk, C. and De Boer, J. Spheroid Culture as a tool for creating 3D complex tissues. *Trends in Biotechnology* **31,** 108-115 (2013)

250. Friedl, P., Hegerfeldt, Y. and Tusch, M. Collective cell migration in morphogenesis and cancer. *Int J Dev Biol.* **48,** 441-449 (2004)

251. Howes, A. L., et al. The phosphatidylinositol 3-kinase inhibitor, PX-866, is a potent inhibitor of cancer cell motility and growth in three-dimensional cultures. *Mol Cancer Ther.* **6,** 2505-2514 (2007)

252. Fujiwara, S., et al. Silencing hypoxia-inducible factor-1α inhibits cell migration and invasion under hypoxic environment in malignant gliomas. *Int J Oncol.* **30,** 793–802 (2007)

253. Indovina, P., Rainaldi, G. and Santini, M. T. Hypoxia increases adhesion and spreading of MG-63 three-dimensional tumor spheroids. *Anticancer Res.* **28,** 1013–22 (2008)

254. Spina, D. PDE4 inhibitors: current status. *British Journal of Pharmacology* **155,** 308-315 (2008)

255. Carpenter, G. and Cohen, S. Epidermal Growth Factor. *The Journal of Biological Chemistry* **265,** 7709-7712 (1990)

256. Burgess, W. H. and Maciag, T. The Heparin-binding (Fibroblast) growth factor family of proteins. *Annual Review of Biochemistry* **58,** 575-606 (1989)

257. Dexamethasone. *The American Society of Health-System Pharmacists* 15.01.2020

258. Gospodarowicz, D. and Moran, J. S. Growth factors in mammalian cell cultures. *Annual Review in Biochemistry* **45,** 31-558 (1976)

259. Mahmood, T. and Yang, P-C. Western Blot: Technique, Theory, and Trouble Shooting. *North American Journal of Medical Sciences* **4,** 429-434 (2012)

260. BIORAD Quick Start™ Bradford Protein Assay. Instruction Manual (Bio-Rad Laboratories Inc.)2000

261. Laemmli, U. K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* **227,** 680-685 (1970)

262. Bio-Rad. *Mini-Protean® Precast Gels*. Instruction Manual and Application Guide. (Bio-Rad Laboratories) 2011

263. Bio-Rad. *Trans-Blot Turbo Transfer System*. Instruction Manual. (Bio-Rad) 2019

264. Soundy, P. and Harvey, B. *Western Blotting as a Diagnostic Method*. **In:** Walker, J. M. and Rapley, R. (eds) *Medical Biomethods Handbook*. (Humana Press) 2015

265. Alamer, M. and Darbre, P. D. Effects of exposure to six chemical ultraviolet filters commonly used in personal care products on motility of MCF-7 and MDA-MB-231 human breast cancer cells in vitro. *Journal of Applied Toxicology* **38,** 148-159 (2018)

266. Amoyel, M. and Bach, E. A. Cell competition: how to eliminate your neighbours. *Development* **141,** 988-1000 (2014)

267. Fulda, S., Gorman, A. M. and Samali, A. Cellular Stress Responses: Cell Survival and Cell Death. *International Journal of Cell Biology* **2010,** 214074 (2009)

268. Rikitake, Y. and Takai, Y. Chapter 3 - Directional Cell Migration: Regulation by Small G Proteins, Nectin-like Molecule-5, and Afadin in *International Review of Cell and Molecular Biology*. Pages 97-143 Volume 287 (Elsevier) 2011

269. Hirata, H., Sokabe, M. and Lim, C. T. Chapter 6 - Molecular Mechanisms Underlying the Force-Dependent Regulation of Actin-to-ECM Linkage at the Focal Adhesions in *Progress in Molecular Biology and Translational Science.* Pages 135-154 Volume 126 (Elsevier) 2014

270. Bearer, E. L. Role of Actin Polymerization in Cell Locomotion: Molecules and Models. *American Journal of Respiratory Cell and Molecular Biology* **8,** 582-591 (1993)

271. Gstraunthaler, G. Alternatives to the Use of Fetal Bovine Serum: Serum-free Cell Culture. *ALTEX.* **20,** 275-281 (2003)

272. Boyce, F. M. and Bucher, N. L. R. Baculovirus-mediated gene transfer into mammalian cells. *PNAS*. **93,** 2348-2352 (1996)

273. Kost, T. A., Condreay, J. P. and Jarvis, D. L. Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nat Biotechnol.* **23,** 567-575 (2005)

274. Lulevich, V., Shih, Y., Lo, S. H. and Liu, G. Cell tracing dyes significantly change single cell mechanics. *J Phys Chem B.* **113,** 6511-6519 (2009)

275. Montero, D., Tachibana, C., Winther, J. R. and Appenzeller-Herzog, C. Intracellular glutathione pools are heterogeneously concentrated. *Redox Biology* **1,** 508-513 (2013)

276. Linnen, C. R. and Hoekstra, H. E. Measuring Natural Selection on Genotypes and Phenotypes in the Wild. *Cold Spring Harb Symp Quant Biol.* **74,** 155-168 (2010)

277. Tamminen, M., et al. Proteome evolution under non-substitutableresource limitation. *Nature Communications* **9,** 4650 (2018)

278. Yi, X. and Dean, A. M. Phenotypic plasticity as an adaptation to a functional trade-off. *eLife* **5,** e19307 (2016)

279. Cengiz, E. and Tamborlane, W. V. A Tale of Two Compartments: Interstitial Versus Blood Glucose Monitoring. *Diabetes Technology and Therapeutics* **11,** S1-11 (2009)

280. American Diabetes Association. Screening for Type 2 Diabetes. *Diabetes Care* **23,** (2000)

281. Dulbecco, R. and Freeman, G. Plaque production by the polyoma virus. *Virology* **8,** 396-397 (1959)

282. Danial, N. N., BAD and glucokinase reside in a mitochondrial complex that integrates glycolysis and apoptosis. *Letters to Nature* **424,** 952-956 (2003)

283. Warburg, O. On the Origin of Cancer Cells. *Science* **123,** 309-314 (1956)

284. White, K. A., Grillo-Hill, B. K. and Barber, D. L. Cancer cell behaviors mediated by dysregulated pH dynamics at a glance. *Journal of Cell Science* **130,** 663-669 (2017)

285. Song, C. W., Griffin, R. and Park, H. J. Influence of Tumor pH on Therapeutic Response *Cancer Drug Discovery and Development: Cancer Drug Resistance* (Humana Press Inc) 2006

286. Kato, Y., et al. Acidic extracellular microenvironment and cancer. *Cancer Cell International* **13,** (2013)

287. Riemann, A., et al. Extracellular Acidosis Modulates the Expression of Epithelial-Mesenchymal Transition (EMT) Markers and Adhesion of Epithelial and Tumor Cells. *Neoplasia* **21,** 450-458 (2019)

288. Khan, K. H., et al. Longitudinal Liquid Biopsy and Mathematical Modeling of Clonal Evolution Forecast Time to Treatment Failure in the PROSPECT-C Phase II Colorectal Cancer Clinical Trial. *Cancer Discovery* **8,** 1-16 (2018)

289. Holohan, C., Van Schaeybroeck, S., Longley, D. B. and Johnston, P. G. Cancer drug resistance: an evolving paradigm. *Nature Reviews Cancer* **13,** 714-726 (2013)

290. Liu, Y., et al. Multi-omic measurements of heterogeneity in HeLa cells across laboratories. *Nat Biotech.* **37,** 314-322 (2019)

291. Kim, I. S., et al. Microenvironment-derived factors driving metastatic plasticity in melanoma. *Nature Communications* **8,** 14343 (2017)

292. Blakwill, F. R., Capasso, M. and Hagemann, T. The tumor microenvironment at a glance. *Journal of Cell Science* **125,** 5591-5596 (2013)

293. Cekanova, M. and Rathore, K. Animal models and therapeutic molecular targets of cancer: utility and limitations. *Drug Design, Development and Therapy* **8,** 1911-1922 (2014)

294. Lyons, N. A. and Kolter, R. On the evolution of bacterial multicellularity. *Current Opinion in Microbiology* **24,** 21-28 (2015)

295. William, C., Ratcliff, R., Denison, F., Borrello, M. and Travisano, M. Experimental evolution of multicellularity. *PNAS.* **109,** 1595-1600 (2012)

296. Kim, T. S., Logsdon, B. A., Park, S., mezey, J. G. and Lee, K. Quantitative Trait Loci for the Circadian Clock in *Neurospora crassa*. *Genetics* **177,** 2335-2347 (2007)

297. Herron, M. D., Zamani-Dahaj, S. A. And Ratcliff, W. C. Trait heritability in major transitions. *BMC Biology* **16,** 145 (2018)

298. Murgia, C., Pritchard, J. K., Kim, S. Y., Fassati, A. and Weiss, R. A. Clonal Origin and Evolution of a Transmissible Cancer. *Cell* **126,** 477-487 (2006)

299. Loh, R., et al. The Pathology of Devil Facial Tumor Disease (DFTD) in Tasmanian Devils (*Sarcophilus harrisii*). *Veterinary Pathology* **43,** 890-895 (2006)

300. Metzger, M. J., et al. Widespread transmission of independent cancer lineages within multiple bivalve species. **534,** 705-709 (2016)

301. Tian, X., et al. Organ-specific metastases obtained by culturing colorectal cancer cells on tissue-specific decellularized scaffolds. *Nature Biomedical Engineering* **2,** 443-452 (2018)

302. Grey, J. F. E., et al. The use of decellularised animal tissue to study disseminating cancer cells. *Journal of Cell Science* **132,** (2019)

303. Welte, Y., Davies, C., Schäfer, R. and Regenbrecht, C. R. Patient derived cell culture and isolation of $CD133^{+}$ putative cancer stem cells from melanoma. *Journal of Visualised Experiments* **13,** e50200 (2013)

304. Pampaloni, F., Reynaud, E. G. and Stelzer, E. H. K. The third dimension bridges the gap between cell culture and live tissue. *Nature Reviews Molecular Cell Biology* **8,** 839-845 (2007)

305. Lawson, D. A., Kessenbrock, K., Davis, R. T., Pervolarakis, N. and Werb, Z. Tumour heterogeneity and metastasis at single-cell resolution.. *Nat Cell Biol* **20,** 1349 (2018)

306. Cowin, P. A., Anglesio, M., Etemadmoghadam, D. and Bowtell, D. D. L. Profiling the Cancer Genome. *Annual Review of Genomics and Human Genetics* **11,** 133-159 (2010)

307. Shaw, R. M., et al. The Cancer Genome Atlas Pan-Cancer analysis project. *Nature Genetics* **45,** 1113-1120 (2013)

308. Jeibouei, S., et al. Personalized medicine in breast cancer: pharmacogenomics approaches. *Pharmacogenomics Personalised Medicine* **12,** 59-73 (2019)

309. Gajria, D. and Chandarlapaty, S. HER2-amplified breast cancer: mechanisms of trastuzumab resistance and novel targeted therapies. *Expert Review of Anticancer Therapy* **11,** 263-275 (2011)

310. Swanton, C. Intratumor heterogeneity: evolution through space and time. *Cancer Res.* **72,** 4875-4882 (2012).

311. Burrell, R. A. & Swanton, C. Re-evaluating clonal dominance in cancer evolution. *Trends in Cancer* **2,** 263-276 (2016).

312. Meyer, M., et al. Single cell-derived clonal analysis of human glioblastoma links functional and genomic heterogeneity. *Proc Natl Acad Sci USA*. **112,** 851-856 (2015)

313. Lai, P. L., et al. Selection of a Malignant Subpopulation from a Colorectal Cancer Cell Line. *bioRxiv* 578823 (2019)

314. Jong, E. D., Chan, I. C. W. and Nedelcu, A. M. A Model-System to Address the Impact of Phenotypic Heterogeneity and Plasticity on the Development of Cancer Therapies. *Frontiers in Oncology.* **9,** 842 (2019)

315. Acar, A., et al. Exploiting evolutionary herding to control drug resistance in cancer. *bioRxiv* 566950 (2019)

Appendix A

Resource competition promotes tumour expansion in experimentally evolved cancer

Taylor et al. BMC Evolutionary Biology (2017) 17:268

DOI 10.1186/s12862-017-1117-6

BMC Evolutionary Biology

RESEARCH ARTICLE Open Access

Resource competition promotes tumour expansion in experimentally evolved cancer

Tiffany B. Taylor[1,2*], Anastasia V. Wass[1], Louise J. Johnson[1] and Phil Dash[1]

Abstract

Background: Tumour progression involves a series of phenotypic changes to cancer cells, each of which presents therapeutic targets. Here, using techniques adapted from microbial experimental evolution, we investigate the evolution of tumour spreading - a precursor for metastasis and tissue invasion - in environments with varied resource supply. Evolutionary theory predicts that competition for resources within a population will select for individuals to move away from a natal site (i.e. disperse), facilitating the colonisation of unexploited resources and decreasing competition between kin.

Results: After approximately 100 generations in environments with low resource supply, we find that MCF7 breast cancer spheroids (small in vitro tumours) show increased spreading. Conversely, spreading slows compared to the ancestor where resource supply is high. Common garden experiments confirm that the evolutionary responses differ between selection lines; with lines evolved under low resource supply showing phenotypic plasticity in spheroid spreading rate. These differences in spreading behaviour between selection lines are heritable (stable across multiple generations), and show that the divergently evolved lines differ in their response to resource supply.

Conclusions: We observe dispersal-like behaviour and an increased sensitivity to resource availability in our selection lines, which may be a response to selection, or alternatively may be due to epigenetic changes, provoked by prolonged resource limitation, that have persisted across many cell generations. Different clinical strategies may be needed depending on whether or not tumour progression is due to natural selection. This study highlights the effectiveness of experimental evolution approaches in cancer cell populations and demonstrates how simple model systems might enable us to observe and measure key selective drivers of clinically important traits.

Keywords: Experimental evolution, Dispersal, Metastasis, Resource competition, Microenvironment, Plasticity, Epigenetic

Background

Solid tumours are largely curable if they are treated before they spread. However, once cancer cells become metastatic and move beyond the location of the primary tumour, mortality rates increase drastically [1]. Metastatic and invasive tumours – those that spread beyond the primary location – show increased spreading to adjacent tissues, which is caused by increased cell motility [2]. As such, targeting premetastatic traits might be a novel approach to prevent the evolution of cancerous traits that would facilitate spreading and invasive behaviours [3]. Indeed there is some evidence to suggest that oxygenation of a tumour inhibits metastasis [4]. However, cancer cells do not need to evolve motility systems de novo, but can co-opt existing mechanisms enabling rapid changes in phenotype [5]. Motility is a normal cellular behaviour for many human cell types, either constitutively, or under particular conditions such as development and tissue repair. Therefore, to understand the processes underlying changes in the behaviour of cancerous cells we must first understand the drivers of change.

Solid tumours, if left untreated, will often progress to metastatic tumours [6]. This is puzzling from an evolutionary perspective. Unlike other hallmarks of cancer such as apoptosis resistance, evasion of growth suppression, or replicative immortality [7], metastasis is not immediately concerned with cell survival or reproduction and appears to have no inherent selective value within a tumour [8]. Nor does motility ensure cell fitness outside the original tumour: of the estimated 106–107 cells that emigrate daily

* Correspondence: t.b.taylor@bath.ac.uk
1 School of Biological Sciences, University of Reading, Whiteknights, Reading RG6 6AH, UK
2 Milner Centre for Evolution and Department of Biology and Biochemistry, University of Bath, Claverton Down Road, Bath BA2 7AY, UK

from a developed neoplasm [9], the vast majority die rather than initiating secondary tumours.

One potential solution to this evolutionary paradox is suggested by an analogy between metastasis and ecological dispersal [10]: an indirect benefit accrues to dispersers if the source population consists of closely related individuals competing for scarce resources [11]. This key prediction was tested in bacterial populations where relatedness between spreading and non-spreading mutants was experimentally manipulated [12]. This study concluded that populations of spreading cells that dispersed further increased distances between competitors and therefore reduced overall cell-cell competition. The consequence being that even under very high costs of dispersal, clonal populations of spreading bacteria were more fit compared to a mixed (low related) population. Therefore, the benefit to moving away from the primary tumour is two-fold: (i) the small proportion of dispersers that successfully colonise a new site will face less competition and reach untapped resources to facilitate rapid growth; and, (ii) by moving away, the cell is reducing competition between its clonemates at the primary tumour site. By increasing the fitness of its clonemates, who will leave more descendants, the disperser is indirectly increasing its own fitness – even if it perishes and fails to establish a metastatic tumour elsewhere [13].

Tumour cells are likely to face exploitation competition in growing neoplasms [14]. Competition will occur for resources including nutrients and oxygen in the early stages of cancer [15], as these can only diffuse approximately 1 mm into a tumour from surrounding blood capillaries alone [16–18]. There is some evidence that these hostile microenvironments favour motility. For example, in uterine cancer [19] and soft tissue sarcomas [20], hypoxia has been shown to be linked with greater likelihood of metastases. Therefore, as the primary tumour grows, resource competition between clonemates is likely to be quickly established. Evolutionary theory predicts that this will drive selection for dispersal.

Natural selection has been detected in clinical tumour samples by applying statistical techniques from population and evolutionary genetics to end-point data [21–23]. However, to gain a deep understanding of the quantitative effect of natural selection in cancer progression we must first go back to evolutionary basics. The power of an experimental evolution approach is that it enables causality of selection to be tested through hypothesis driven experiments.

Recent dispute over the importance of natural selection in tumour progression [24, 25] has highlighted the need for a quantitative understanding of the forces leading to cancer progression.

Viewing cancer progression as an evolutionary and ecological process is becoming more common practise; providing new insights into progression and treatment of cancers [26–29]. In particular, dispersal evolutionary simulation models have been utilised to explore the evolution of spreading behaviour revealing the role of metabolism and nutrient competition [30], the microenvironment [31] and resource heterogeneity [32] in driving the evolution of cell migration. A common theme is that the nutrient environment plays a critical role in selecting for increased cell motility, which provided the context for this study.

The next step is to experimentally validate these predictions for which we must develop experimental techniques that can accurately measure the effect of key selective drivers on the evolution of clinically relevant traits in tumour cell populations. In Taylor et al. [33], we advocated adapting the techniques of experimental evolution in microbes to cancer research (see also [34]). Here, we report the findings of the experiment we proposed to determine the role of cell-cell competition in causing increased spheroid spread, which models an early stage of metastasis seen within primary tumours (for a discussion of the advantages and limitations of spheroids spread as a model for metastasis, see [35]).

Methods

Aim

The aim of this study was to identify resource supply (high or low) as a driver for the evolution of spheroid spreading in a population of breast cancer cells. Six independent selection lines of MCF7 breast cancer cells were established; 3 replicate lines were maintained under low resource supply and 3 under high resource supply. Every 7 days, after the cells had become confluent, 10% were transferred to fresh media. Transfers were made each week for 12 weeks. Comparisons between lines, and with the ancestor, gives a measure of the effect of selection over time on cell phenotype.

Cell culture

Experiments were performed using MCF7 cells (ATCC® HTB-22™; passage number 17) [36], a relatively slow moving, non-metastatic cell line (although derived from a metastatic site). Cells were grown as monolayers in 25 cm^2 tissue culture flasks with non-phenol red Dulbecco's Modified Eagles Medium (DMEM)

containing 0.5% or 5% (depending on treatment group) foetal bovine serum (FBS), 1% Penicillin, 1% streptomycin, and 2 mM L-glutamine. Incubated at 5% CO_2, 37 °C.

Selection lines

MCF7 cell lines were maintained in low (0.5% FBS) or high (5% FBS) resources for 12 weeks. MCF7 cells have a generation time of approximately 24 h. Three independent replicate lines were maintained within each treatment group. Every 7 days a random subpopulation of 1000 - 4000 cells were transferred to fresh medium. 7 days were sufficient to allow cells to cover the base of the cell tissue flask, forcing cells to compete for space and resources.

Cells were removed from the incubator and the old media was discarded. To detach, cells were washed in 5 ml Phosphate Buffered Saline (PBS), treated with 2 ml Trypsin-EDTA and incubated at 37 °C for approximately 5 min. Cells were re-suspended in DMEM and 10% were transferred to fresh media. At each transfer a sample of each cell line was frozen down to allow resurrection for further post hoc phenotypic analysis. Cells were passaged as normal and the remaining solution was centrifuged at 1000 rpm for 3 min. The supernatant was discarded, cells re-suspended in freezing buffer (10% dimethylsulfoxide (DMSO) and 90% FBS) and stored at −80 °C. After 24 h, the vials were moved and placed in liquid nitrogen.

Spheroid spread assays

Spheroids offer a tangible in vitro model that more accurately reflect clinical expression profiles compared to monolayer cultures [35, 37], and have been previously used to study spreading behaviour of cancers (e.g. [38–40]). Cells from the selection lines were grown in non-adhesive flasks for 24 h, allowing them to form spheroids (small in vitro tumours), roughly spherical clusters of approximately 1000 cells. To avoid any initial responses to change in media, spheroids were cultured in the same media that they were to be tested in. When returned to flasks with a suitable surface, spheroids will adhere and the constituent cells will move outwards, eventually forming a monolayer. We allowed spheroids to adhere for 4 h to the surface of 12well tissue culture plates, and calculated the areas covered by cells dispersing from the spheroid by analysing images taken on a Zeiss A1 Inverted Epifluorescent microscope using Nikon NIS Elements and analysed with ImageJ software [41]. Six spheroids were measured within each well and 3 independent wells were measured. Spheroids within wells were randomly paired between photos taken at time zero and 72 h and the difference calculated. We were unable to measure the

same spheroids between time points as magnification adjustment and manual tracking was necessary to accommodate rapid spread to keep spheroids within the field of view.

Growth rate assays

Cells were transferred to 6-well plates and seeded at 2% confluency. The number of cells present were counted using a haemocytometer at time 0 and after 72 h to gain an initial cell count prior to cell adherence and total end cell count per well, respectively. Images were taken at 12 h intervals at the same locations within each well, and cells counted using ImageJ. The average for 3 independent replicates across all 3 cell lines was taken. Growth rate was calculated using the formula:

$$\frac{ln\left(\frac{N_1}{N_2}\right)}{\left(t_{2-t_1}\right)}$$

where t_2 is the time at the end of the experiment (72 h), t_1 is the time at the beginning of the experiment (12 h), N_2 is the number of cells present at t_2 and N_1 is the number of cells present at t_1. The first image was taken at 12 h to allow cells time at adhere to the well surface. Between 12 h and 72 h growth rate is assumed to be exponential as cells have not yet reached confluency.

Timelapse cell motility assay

Cells from the selection lines were seeded in a 12-well plate at approximately 5000 cells/well. Each cell line was cultured in the media it had been adapted to and each well had 5 points chosen at random from which to observe the cells. A Nikon TE200 Timelapse System with NIS Elements 3 was used to capture a bright-field image at the points chosen every 15 min for just over 48 h (actual, 52 h and 15-min). These images were collated to form a timelapse video allowing the individual cells within the field of view to be tracked using ImageJ and MtrackJ. This tracking allows observation of whether or not a cell divides and can be used to calculate the speed of individual cells. Speed was calculated as the total length (in microns) moved by the cell divided by the time (in hours) the cell was tracked for. As many cells as possible were measured and this was repeated for 3 independent replicates. In addition, distance to point was also measured using tracking data. Here, the distance a cell moves between two time-frames (every 15 min) is recorded against time. The benefit of this measurement is that it allows the proportion of cells across a population that are moving at a particular time point to be calculated, rather than tracking the motility of an individual cell. This, combined with the

cell speed data, gives an indication of population behaviour across time.

Statistics

Analyses and figures were produced on IBM® SPSS® Statistics 24.0. Significance of treatment (high or low resource supply) on phenotype was analysed parametrically using general linear models (GLMs). Terms used in the model are defined as the following: 'Spheroid area' [response variable], average area covered by spheroid spread after 72 h. Areas were square-root transformed to correct for right skew that is typical of area data, and divided by growth rate to correct for expansion of spheroid driven by differences in growth rate rather than spreading behaviour; 'Experimental media' [explanatory variable, factor], high or low resource supply in experimental conditions; 'Evolved environment' [explanatory variable, factor], ancestor and high or low resource supply maintained in evolving lines; 'Cell speed' [response variable], the average distance (429 Ancestor, 298 Evolved highresource supply and 356 Evolved low-resource supply) the cells moved in micrometres per hour over a 52-h period. A Kruskal-Wallis and skew test were used to measure the distribution of 'distance to point' data. In each case the effect of the well within the tissue culture plate was measured. This was to ensure there were no microenvironment differences between wells – this was non-significant in all cases and therefore removed from the final model. Replicate is treated as a random factor. Tukey tests were performed between treatment groups within cell line. In all cases the area dispersed was square-root transformed.

Results

Images of cells across different transfer points were taken over 72 h to measure the distance of spread of spheroids (clumps of around 1000 self-adhered cells; Fig. 1a) from an initial adhesion site. The change in area over time was used as a measure of spheroid spread. At transfer 0, there was no difference between spheroid spread in the different treatments for resource supply ($F_{1,9} = 1.426$; $P = 0.263$). Furthermore, there was no significant difference between initial spheroid size between selection lines ($F_{2,101} = 0.163$; $P = 0.848$). However, over time we find resource supply has a significant effect on spreading ($F_{1,105} = 103.61$; $P < 0.001$), with faster spreading emerging in lines maintained under low resource supply ($F_{3,105} = 23.872$; $P < 0.001$) (Fig. 1b). These effects are unlikely to be driven by the microenvironment as there was no significant difference between spreading distances of replicate spheroids between wells ($F_{2,89} = 1.423$; $P = 0.246$). After 12 transfers, spheroids from ancestor and both selection lines were allowed to grow and spread under high and low resource conditions (a "common garden" experiment in ecology). The ancestor showed no difference in spreading area under high or low resources supply ($F_{1,4} = 0.613$; $P = 0.477$). However, the evolved lines were found to respond differently; high-resource selection lines showed no difference in motility between the two resource conditions, but low-resource selection lines showed slower motility in high-resource medium compared to low-resource medium (Fig. 2; $F_{1,8} = 10.502$; $P = 0.012$).

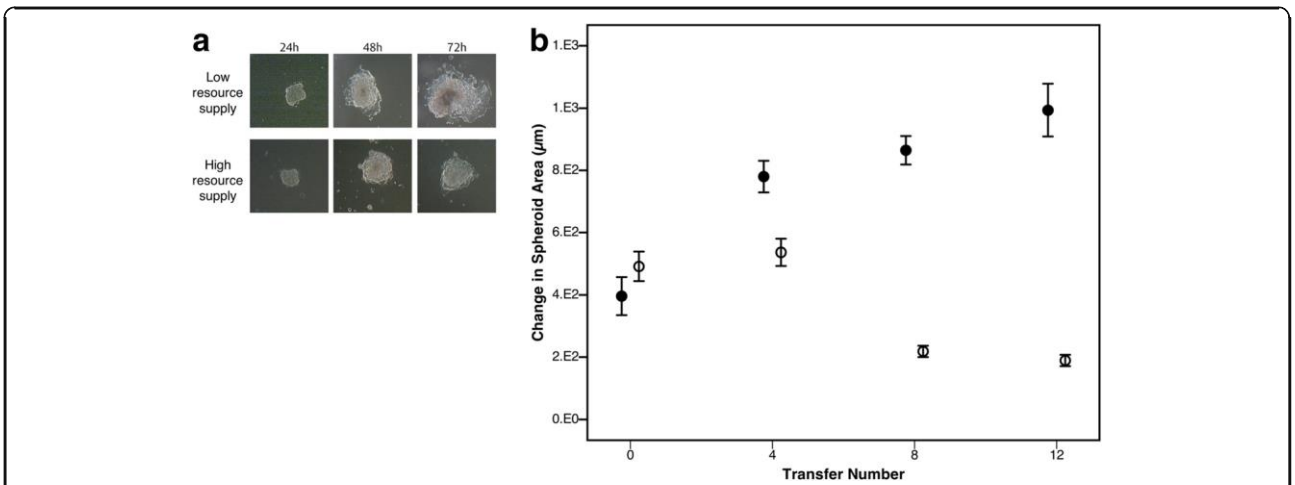


Fig. 1 Spheroid spread over time. a. Images of MCF7 cell populations demonstrating differences in spheroid (in vitro tumours of approximately 1000 cells) spread over a 72-h period. Cell motility was calculated by measuring the change in surface area covered by spheroids over time; b Average spheroid spread (across 3 replicate lines) for selection lines under high and low resource supply over 12 weekly transfers. Spreading was measured 72 h after adhesion of the spheroid, and in the selective environment for each line (i.e. the motility of the low resource supply line was measured in low resource medium). Closed circles represent lines evolved under low resources, open circles are lines evolved with access to high resources. Areas are square-root transformed (± 1 SE)

To correct for differences in spheroid spread that might be due to growth, spreading area is divided by growth rate (Fig. 3). We found that growth rates of selection lines and ancestor were not different with access to high resource supply (5.0% FBS) ($F_{2,6}$ = 1.171; P = 0.372), however under low resource supply (0.5% FBS) we find significant differences between ancestral and evolved lines ($F_{2,6}$ = 17.601; P = 0.003). When grown in 0.5% FBS (low-resource supply), lines evolved under high and low resource supply showed lower growth rates than the ancestor, despite the low-resource evolved lines showing higher spheroid spreading (Tukey: Ancestor v. 5.0% FBS P = 0.034; Ancestor v. 0.5% FBS, P = 0.003). This suggests that spreading behaviour is not explained by growth.

The speed of cell movement was measured using video tracking of motile cells. Individual cell motility in lines evolved in high and low resources were not significantly different from the ancestor, or each other (Fig. 4; Tukey test: 0.5% v. 5.0%, P = 0.968; 0.5% v. Ancestor, P = 0.282; 5.0% v. Ancestor, P = 0.377). In addition, we calculated the 'distance to point', which measures the distance a cell moves between time frames (Additional file 1: Figure S1). Populations evolved in a low resource supply show fewer non-motile cells at a given time point compared to those evolved with high resource supply and ancestral lines. In addition, the distribution of 'distance to point' values across different lines is not equal (Kruskal-Wallis, P < 0.05). Ancestors show the highest positive skew (4.164 ± 0.011) followed by lines evolved in high resource supply (3.176 ± 0.015) and lines evolved in low resource supply (2.485 ± 0.014). Finally, these trends hold true across time, when distance to point is measured after 15 min, 24 h and 48 h. Together, these data suggest that in low resource supply lines: there are fewer non-motile cells; a greater proportion of cells are spreading further between time points; and, these differences are sustained across time. This would mean that cells evolved in low resource supply would spread more quickly, as distances between cells increases, compared to ancestral and high resources supply lines.

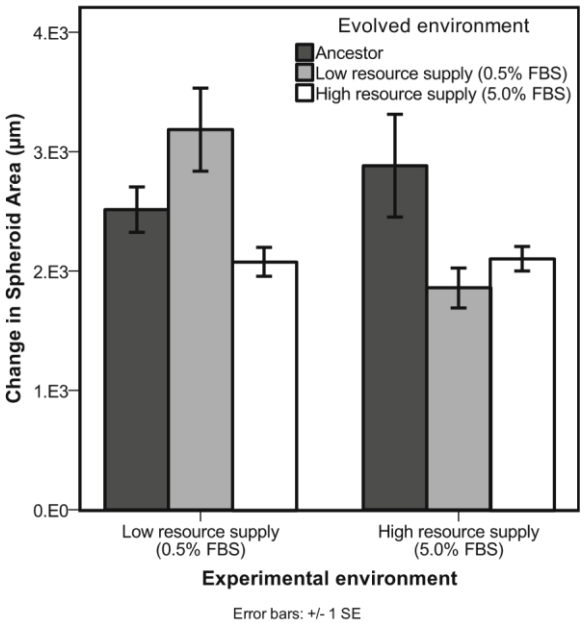


Fig. 2 Common garden experiment. Average change in spheroid area (square-rooted) over 72 h for ancestor and evolved high and low resource selection lines in high and low resource environments. Areas are square-root transformed (± 1 SE)

Discussion

We grew independent cancer cell populations over multiple generations to measure the effect of resource supply on spheroid spread. We find a significant difference in spreading area over time driven by resource supply. In particular, as predicted by dispersal evolutionary theory, when resources are scarce (driving competition between cells) spheroids are under selection to spread further compared to when they have access to a high resource supply in the same time period (Fig. 1b). Despite the energetic demands of cell motility, spheroids in high resource environments show reduced spreading, and spheroids with access to low resource supply evolve increased spreading. In addition, a common garden experiment – where

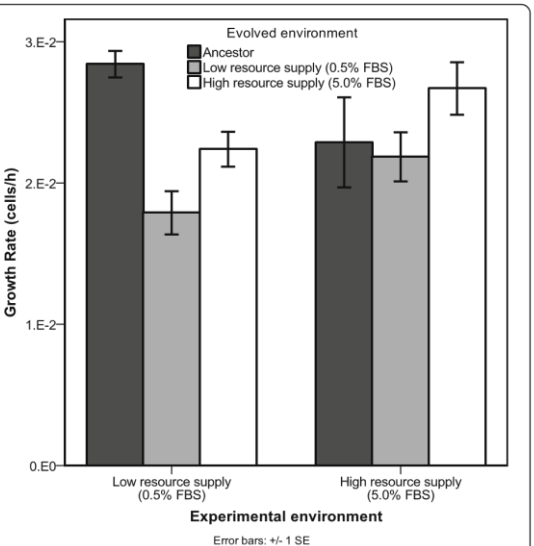


Fig. 3 Growth rate of ancestral and evolved lines. Average growth rate, represented as the increase in cell number per hour (± 1 SE). Growth rates of ancestral (dark grey bars), low resource supply evolved lines (light grey bars), and high resource supply evolved lines (white bars) are shown when grown in high resources (5.0% FBS) and low resources (0.5% FBS)

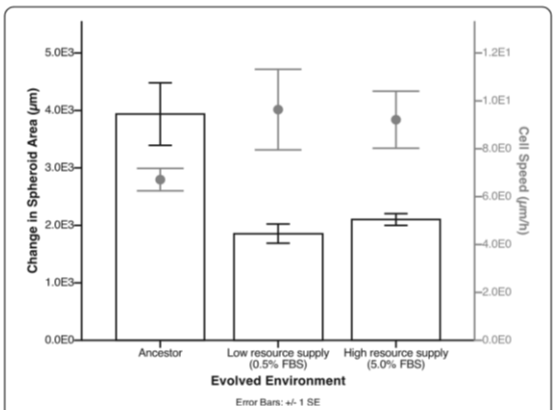


Fig. 4 Spheroid spread versus cell speed. Bars represent the average change in spheroid area over 72 h in high (5% FBS) resource media for ancestor and lines evolved in low and high resources (± 1 SE). Closed circles represent average cell speed of cells in high resource media for ancestor and evolved lines. Areas are square-root transformed (± 1 SE)

evolved lines are grown in both low and high resource environments – reveals that cell populations evolved with a low resource supply show phenotypic plasticity in spheroid spread, such that the rate of spread depends on their current nutrient environment, inducing faster spreading in low resources supply compared to the same population in high resource conditions (Fig. 2). Cell populations evolved with access to a high resource supply do not respond to changes in resource environment.

These results are consistent with competition selection as a driver of dispersal; as 40 generations is a relatively short timeframe, this would most likely be selection on standing genetic variation, which can be substantial in MCF7 due to genome instability [42]. However, the complexity of human gene regulatory pathways and the known sensitivity of cancer cell lines to subtle changes in environmental conditions [43], as well as uncharacterised experimental effects [44] necessitates caution in interpretation. Epigenetic responses to the environment that persist across cell generations could also cause stable changes in phenotype such as are seen here. Distinguishing between these alternative explanations should be a high research priority, not only out of academic interest in the precise level of adaptationism appropriate to cancer biology, but also because the difference may have clinical implications. If metastasis evolves as the result of genetic changes, effective preventative treatments might be those that focus on keeping the cancer cell effective population size and mutation rate low; however, if metastasis emerges as a response to environment, it is of greater importance to understand and control tumour microenvironments.

Hostile microenvironments may both directly cause and selectively favour tumour spread [45], and could pose particular danger in promoting metastasis. Human tumour cells contain the whole human genome, and so are capable of phenotypic change via complex physiological, epigenetic and developmental responses in addition to evolutionary response to natural selection [46]. Models from Waclaw [47] suggest that even shortrange cellular migratory activity can markedly increase the rate of tumour growth (i.e. fitness of the tumour cells), even in the absence of changes in cellular growth rates. We found that measurable, stable changes in spheroid spreading behaviour occurred quickly (within 4 weeks), suggesting this trait could evolve rapidly in vivo. Although, care must be taken when translating findings from the lab to in vivo, as the selective environments will greatly differ.

A cell's microenvironment is also a product of the population in which it resides. Within this experimental setup, we are considering motility at the level of the population (within a spheroid) rather than the individual (a cell). In other words, we are looking at the effect of resource supply in determining the spread of a tumour rather than individual cancer cells. We found spheroids that were evolved with access to a high resource supply showed reduced spread over time (Fig. 1b). This pattern was not driven by changes in cell motility, as no significant differences in cell motility were seen over time (Fig. 4). This suggests that access to resources during evolution is determining the patterns of spheroid spread independently of cell motility. One interpretation might be that cells evolved in starvation, rather than generally getting faster, are more fit if they move further from cells around them to reduce cell-cell competition – a key prediction in dispersal evolution [10]. Our data finds support for this hypothesis; we find that cells evolved under low-resource supply move consistently across time, and the distance they move slightly increases. In comparison, cells evolved in high-resource supply move the same distance between time points. This suggests that cells evolved in low-resource supply are moving further from the cells around them. The consequence for spheroid spread would be that cells towards the centre of the spheroid mass would have more space to move into if distances between cells at the periphery were greater, increasing spheroid spreading rate.

Dispersal plays a crucial role in a range of evolutionary and ecological processes and as such there has been a major effort to understand its evolution. A central factor that has been highlighted both theoretically [11, 48–50] and empirically [12] is that dispersal is likely to be favoured by selection if it reduces kin competition (here, competition between clonemates). These studies find that even under extremely high costs, dispersal is still favoured when populations are clonal due to the indirect fitness benefits gained from reducing competition between clonemates left in the natal patch. This is because these clonemates will pass on genes shared by the dispersing cell, even if the dispersing cell does not survive to do so itself.

Dispersal theory has previously been applied to cancers to try to predict the impact of the microenvironment on the emergence of motility and metastasis in cancer cell populations [30–32]. However, in these cases the evolution of dispersal was only considered from the perspective of individual cell fitness. Although these studies concur that competitive environments (low nutrient or hypoxic) select for increased tumour spread, they

do not consider the role of cell-cell (kin) competition as a driver for the evolution of cell dispersal. Considering the inclusive fitness of the cell – that is, taking into account not only its own reproductive success, but its effects on its relatives – may help solve the paradox as to why metastasis evolves despite high mortality rates of metastasising cells.

Dispersal can therefore be considered as a social behaviour. Social evolution – evolution under the consideration of inclusive fitness – is an area that has been extensively studied using experimental evolution and provides interesting opportunities for further study in cancers [51]. Cell behaviour is likely to be dependent on the social context (i.e. cellular behaviour changes when acting as an individual cell compared to within a tumour) [52, 53]. Our results find that individual cell speed and spheroid spread are not aligned, possibly because the social context is different between the experimental setups: cell speed is measured in a 2D monolayer culture, whereas spheroid spread starts as a 3D multi-cell aggregate. Incorporating social evolution into cancer evolution is likely to reveal new insights into the levels of selection across tumours and bridge gaps in our understanding of the evolution of multicellularity. Recent work has shown that even within a tumour, heterogeneity is established early and as such, evolutionary trajectories could be different across a very small scale [54–56]. In this study, however, while we do not measure a significant difference in cell speed, we do find that there are differences in the distance travelled between time frames between selection lines within a monolayer culture. This suggests that selection is having a measurable effect on the phenotype of cell motility within a 2D environment, as motility is a function of both speed and frequency of cell movement. Moreover, this suggests that selective pressures in one environmental structure (2D) can have important clinical consequences in alternative environments (3D) that are not clear until measured.

It is important to acknowledge that we do not claim to replicate in vivo conditions within our experimental design, in fact it is our aim to simplify the environment as much as possible. As such, there will be key differences, such as replicating cell behaviours from a structured 3D tumour to 2D in vitro assays. However, by simplifying experiments into a 2D environment we can develop methods that are easy to repeat and measure. Simplicity of design is a major strength of experimental evolution as it captures the influence of isolated selective drivers in the absence of biological noise – thus improving overall generality of results. Spheroids offer an

ideal in vitro model for studying tumour spreading. Spheroids are a very simple 3D culture; it allows cells to form cell-cell adhesions and then spread. This enables us to see whether cells change their behaviour between a 2D and 3D environment.

Translation of results from experimental cancer evolution studies into potential preventative or therapeutic approaches to cancer treatment is not a trivial task. A hypothesis driven approach, as is seen in experimental evolution studies, can only highlight key selective drivers of clinically relevant cancerous traits in the absence of in vivo noise. As this field develops this information, in combination with front-line research from cancer biologists and clinicians, will reveal novel treatment strategies – such as prevention approaches in patients with high-risk of cancer prior to tumour detection – and confer greater understanding and predictive power to the evolution of clinically relevant traits such as metastasis and drug resistance.

Conclusions

Cancer researchers face a daunting challenge – to harness evolutionary theory in a clinically meaningful way, and further our understanding of the progression of cancers. To meet this challenge, we must design hypothesis-driven experimental systems to effectively test theoretical predictions of cancer evolution [33, 34]. An experimental evolution approach can help systematically address central questions such as: what is the balance between ecological and evolutionary processes? Can we distinguish between genetic and epigenetic evolutionary changes? And, can we repeat the same evolutionary patterns across different environments and across different cancers? The next step would be to complement experimental evolution data with genomics and transcriptomics to allow changes in coding and regulatory regions to align with phenotypic changes over evolutionary time – giving a clearer indication of how genotype maps with phenotype.

This study uses experimental evolution to observe, in real time, the evolution of a key cancer trait and precursor to metastasis – tumour spreading. We find that low resource supply drives the evolution of spheroid spread. This result aligns with predictions from dispersal evolutionary theory. Seminal experimental evolution studies with microbes have fundamentally changed our understanding of evolution (for review see [57]) and there is strong potential for similar advancements in cancer biology. However, this task is not as simple as repeating existing experiments in a new system. While cancer cell populations share many similarities with microbes that make them amenable to

experimental evolution studies [33, 34, 54], they also present many new challenges. There is much greater potential for both stable and transient epigenetic effects on phenotype, and factors such as cancer types and genetic backgrounds will introduce further complexity. However, this study introduces a promising starting point for the development of experimental techniques to detect, measure and quantify key evolutionary processes in cancers.

Additional file

Additional file 1: Figure S1. Histogram showing distribution of distance to point measurements after 0.25, 24 and 48 h. Blue bars represent ancestral populations, green bars represent populations evolved in low resource supply (0.5% FBS) and pink bars represent populations evolved in high resource supply (5.0% FBS). (PDF 5 kb)

Abbreviations

DMEM: Dulbecco's modified eagles medium; DMSO: Dimethylsulfoxide;

FBS: Foetal bovine serum; MCF-7: Michigan cancer foundation-7;

PBS: Phosphate buffered saline; UR: University of Reading, Reading, UK; US: University of Sheffield, Sheffield, UK

Acknowledgements

We thank Khia Wilson for the photos in Fig. 1a; and Robert Jackson (UR) and Michael Brockhurst (US) for productive discussion and comments. We also thank 2 anonymous reviewers and Carlo Maley for their comments and suggestions.

Funding

This research was funded by the Leverhulme Trust (LJJ; F/00239/AL) and the Royal Society (TBT; DH150169). AVW was supported by a University of Reading studentship (PD).

Availability of data and materials

The datasets generated and analysed during the current study are available in the DRYAD repository [doi:https://doi.org/10.5061/dryad.cc34r].

Authors' contributions

TBT, AVW, LJJ and PD conceived and designed the study. TBT maintained the experimental evolution lines and was a major contributor in writing the manuscript. TBT and AVW performed phenotypic and statistical analysis of cell lines. All authors contributed to writing, reading and approving the final manuscript.

Ethics approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Received: 17 May 2017 Accepted: 14 December 2017Published online: 27 December 2017

References

1. Chambers AF, Groom AC, IC MD. Metastasis: dissemination and growth of

cancer cells in metastatic sites. Nat Rev Cancer. 2002;2(8):563–72.

2. Martin TA, Ye L, Sanders AJ, Lane J, Jiang WG. Cancer invasion and metastasis: molecular and cellular perspective. In: Rahul J, editor. Metastatic cancer clinical and biological perspectives. Austin: Landes Bioscience; 2013. p. 135–68.
3. Wells A, Grahovac J, Wheeler S, Ma B, Lauffenburger D. Targeting tumor cell motility as a strategy against invasion and metastasis. Trends Pharmacol Sci. 2013;34(5):283–9.
4. Mazzone M, Dettori D, de Oliveira RL, Loges S, Schmidt T, Jonckx B, et al. Heterozygous deficiency of PHD2 restores tumor oxygenation and inhibits metastasis via endothelial normalization. Cell. 2009;136(5):839–51.
5. Schiffman JD, White RM, Graham TA, Huang Q, Aktipis A. The Darwinian dynamics of motility and metastasis. In: Maley CC, Greaves M, editors. Frontiers in cancer research: evolutionary foundations, revolutionary directions. New York: Springer New York; 2016. p. 135–76.
6. Nguyen DX, Bos PD, Massague J. Metastasis: from dissemination to organspecific colonization. Nat Rev Cancer. 2009;9(4):274–84.
7. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000;100(1):57–70.
8. Bernards R, Weinberg RA. Metastasis genes: A progression puzzle. Nature. 2002;418(6900):823.
9. Butler LM, Agus DB, Scher HI, Higgins B, Rose A, Cordon-Cardo C, et al. Suberoylanilide hydroxamic acid, an inhibitor of histone deacetylase, suppresses the growth of prostate cancer cells in vitro and in vivo. Cancer Res. 2000;60(18):5165–70.
10. Amend SR, Roy S, Brown JS, Pienta KJ. Ecological paradigms to understand the dynamics of metastasis. Cancer Lett. 2016;380(1):237–42.
11. Hamilton WD, May RM. Dispersal in stable habitats. Nature. 1977; 269(5629):578–81.
12. Taylor TB, Buckling A. Competition and dispersal in Pseudomonas Aeruginosa. Am Nat. 2010;176(1):83–9.
13. Hamilton WD. The genetical evolution of social behaviour. II. J Theor Biol. 1964;7(1):17–52.
14. Pianka ER. Competition and niche theory. In: May R, MacLean A, editors. Theoretical ecology principles and applications. Oxford: Oxford University Press; 1981. p. 167–96.
15. Höckel M, Vaupel P. Tumor hypoxia: definitions and current clinical, biologic, and molecular aspects. J Natl Cancer Inst. 2001;93(4):266–76.
16. Boucher Y, Jain RK. Microvascular pressure is the principal driving force for interstitial hypertension in solid tumors: implications for vascular collapse. Cancer Res. 1992;52(18):5110–4.
17. Araujo RP, McElwain DS. A history of the study of solid tumour growth: the contribution of mathematical modelling. Bull Math Biol. 2004;66(5):1039–91.
18. Vaupel P, Harrison L. Tumor hypoxia: causative factors, compensatory mechanisms, and cellular response. Oncologist. 2004;9(Supplement 5):4–9.

19. Höckel M, Schlenger K, Aral B, Mitze M, Schäffer U, Vaupel P. Association between tumor hypoxia and malignant progression in advanced cancer of the uterine cervix. Cancer Res. 1996;56(19):4509–15.
20. Brizel DM, Scully SP, Harrelson JM, Layfield LJ, Bean JM, Prosnitz LR, et al. Tumor oxygenation predicts for the likelihood of distant metastases in human soft tissue sarcoma. Cancer Res. 1996;56(5):941–3.
21. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, et al.
22. Patterns of somatic mutation in human cancer genomes. Nature. 2007; 446(7132):153–8.
23. Woo YH, Li W-H. DNA replication timing and selection shape the landscape of nucleotide variation in cancer genomes. Nat Comms. 2012;3:1004.
24. Williams MJ, Werner B, Barnes CP, Graham TA, Sottoriva A. Identification of neutral tumor evolution across cancer types. Nat Genet. 2016;48(3):238–44.
25. Frank SA, Rosner MR. Nonheritable cellular variability accelerates the evolutionary processes of cancer. PLoS Biol. 2012;10(4):e1001296.
26. Sidow A, Spies N. Concepts in solid tumor evolution. Trends Genet. 2015; 31(4):208–14.
27. Merlo LM, Pepper JW, Reid BJ, Maley CC. Cancer as an evolutionary and ecological process. Nat Rev Cancer. 2006;6(12):924–35.
28. Crespi B, Summers K. Evolutionary biology of cancer. Trends in Ecol Evol. 2005;20(10):545–52.
29. Greaves M, Maley CC. Clonal evolution in cancer. Nature. 2012; 481(7381):306–13.
30. Korolev KS, Xavier JB, Gore J. Turning ecology and evolution against cancer. Nat Rev Cancer. 2014;14(5):371–80.
31. Aktipis CA, Maley CC, Pepper JW. Dispersal evolution in neoplasms: the role of disregulated metabolism in the evolution of cell motility. Cancer Prev Res. 2012;5(2):266–75.
32. Anderson AR, Weaver AM, Cummings PT, Quaranta V. Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. Cell. 2006;127(5):905–15.
33. Chen J, Sprouffske K, Huang Q, Maley CC. Solving the puzzle of metastasis:
34. the evolution of cell migration in neoplasms. PLoS One. 2011;6(4):e17933.
35. Taylor TB, Johnson LJ, Jackson RW, Brockhurst MA, Dash PR. First steps in experimental cancer evolution. Evol Appl. 2013;6(3):535–48.
36. Sprouffske K, Merlo LMF, Gerrish PJ, Maley CC, Sniegowski PD. Cancer in light of experimental evolution. Curr Biol. 2012;22(17):R762–71.
37. Hirschhaeuser F, Menne H, Dittfeld C, West J, Mueller-Klieser W, KunzSchughart LA. Multicellular tumor spheroids: an underestimated tool is catching up again. J Biotechnol. 2010;148(1):3–15.
38. Soule HD, Vasquez J, Long S, Albert S, Brennan M. A human cell line from a pleural effusion derived from a breast carcinoma. J Natl Cancer Inst. 1973; 51(5):1409–16.

39. Hamilton G. Multicellular spheroids as an in vitro tumor model. Cancer Lett. 1998;131(1):29–34.
40. Howes AL, Chiang GG, Lang ES, Ho CB, Powis G, Vuori K, et al. The
41. phosphatidylinositol 3-kinase inhibitor, PX-866, is a potent inhibitor of cancer cell motility and growth in three-dimensional cultures. Mol Cancer Ther. 2007;6(9):2505–14.
42. Fujiwara S, Nakagawa K, Harada H, Nagato S, Furukawa K, Teraoka M, et al. Silencing hypoxia-inducible factor-1α inhibits cell migration and invasion under hypoxic environment in malignant gliomas. Int J Oncol.
43. 2007;30(4):793–802.
44. Indovina P, Rainaldi G, Santini MT. Hypoxia increases adhesion and spreading of MG-63 three-dimensional tumor spheroids. Anticancer Res. 2008;28(2A):1013–22.
45. Abràmoff MD, Magalhães PJ, Ram SJ. Image processing with ImageJ. Biophoton Int. 2004;11(7):36–42.
46. Hampton OA, Den Hollander P, Miller CA, Delgado DA, Li J, Coarfa C, et al. A sequence-level map of chromosomal breakpoints in the MCF-7 breast cancer cell line yields insights into the evolution of a cancer genome. Genome Res. 2009;19(2):167–77.
47. Kim IS, Heilmann S, Kansler ER, Zhang Y, Zimmer M, Ratnakumar K, et al.
48. Microenvironment-derived factors driving metastatic plasticity in melanoma. Nature Comms. 2017;8:14343.
49. Baker M. Reproducibility: respect your cells. Nature. 2016;537(7620):433–5.
50. Joyce JA, Pollard JW. Microenvironmental regulation of metastasis. Nat Rev Cancer. 2009;9(4):239–52.
51. Li S, Garrett-Bakelman FE, Chung SS, Sanders MA, Hricik T, Rapaport F, et al. Distinct evolution and dynamics of epigenetic and genetic heterogeneity in acute myeloid leukemia. Nat Med. 2016;22(7):792–9.
52. Waclaw B, Bozic I, Pittman ME, Hruban RH, Vogelstein B, Nowak MA. A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity. Nature. 2015;525(7568):261–4.
53. Comins HN, Hamilton WD, May RM. Evolutionarily stable dispersal strategies. J Theor Biol. 1980;82(2):205–30.
54. Taylor PD, Frank SA. How to make a kin selection model. J Theor Biol. 1996;
55. 180(1):27–37.
56. Gandon S, Michalakis Y. Evolutionarily stable dispersal rate in a Metapopulation with extinctions and kin competition. J Theor Biol. 1999; 199(3):275–90.
57. West SA, Griffin AS, Gardner A, Diggle SP. Social evolution theory for microorganisms. Nat Rev Microbiol. 2006;4(8):597–607.
58. Nadell CD, Bucci V, Drescher K, Levin SA, Bassler BL, Xavier JB. Cutting through the complexity of cell collectives. Proc Royal Soc B: Biol Sci. 2013; 280(1755):20122770.
59. Vedel S, Tay S, Johnston DM, Bruus H, Quake SR. Migration of cells in a social context. Proc Natl Acad Sci. 2013;110(1):129–34.
60. de Bruin EC, Taylor TB, Swanton C. Intra-tumor heterogeneity: lessons from

microbial evolution and clinical implications. Genome Med. 2013;5(11):101.

61. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med. 2012;2012(366):883–92.
62. Carmona-Fontaine C, Deforet M, Akkari L, Thompson CB, Joyce JA, Xavier JB.
63. Metabolic origins of spatial organization in the tumor microenvironment. Proc Natl Acad Sci. 2017;114(11):2934–9.
64. Buckling A, Craig Maclean R, Brockhurst MA, Colegrave N. The beagle in a bottle. Nature. 2009;457(7231):824–9.

Appendix B

Cancer cell lines show high heritability for motility but not generation time

Article Type: Research Article

Title: Cancer cell lines show high heritability for motility but not generation time

Authors: Anastasia V. Wass[1], George Butler[1], Tiffany B. Taylor[1,2], Philip R. Dash[1] and Louise J. Johnson[1]*

[1]School of Biological Sciences, University of Reading, Whiteknights, Reading, RG6 6AH and [2]The Milner Centre for Evolution and Department of Biology and Biochemistry, University of Bath, Claverton Down, Bath, BA2 7AY

* Corresponding author

Contact info for corresponding author:

Email: L.J.Johnson@Reading.ac.uk

tel: +44 (0) 118 378 4432

ORCID: 0000-0002-0006-1511

Keywords: cancer evolution, heritability, phenotypic inheritance, experimental evolution

Abstract word count: 100

Main text word count including Methods: 1508

Display items: 1 Figure (colour essential); 1 Table.

Author contribution statement: AVW, PRD, TBT and LJJ conceived the project. AVW conducted laboratory work and collected data. GB and AVW conducted statistical analyses. All authors contributed to writing the manuscript. PRD and LJJ co-supervised the project and are joint senior authors. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

Data availability statement: All data generated and analysed in this study are available at http://dx.doi.org/10.17864/1947.167

Funding statement: This work was supported by the University of Reading, the Leverhulme Trust (reference: F/00 239/AL) and an anonymous charitable donation.

Competing interests statement: No competing interests to declare.

Abstract

Tumour evolution depends on heritable differences between cells in traits affecting cell survival or replication. It is well-established that cancer cells are genetically and phenotypically heterogeneous; however, the extent to which this phenotypic variation is heritable is far less well-explored.

Here we estimate the broad-sense heritability (H^2) of two cell traits related to cancer hallmarks - cell motility and generation time - within populations of four cancer cell lines *in vitro*, and find that motility is strongly heritable. This heritability is stable across multiple cell generations, with heritability values at the high end of those measured for a range of traits in natural populations of animals or plants. These findings confirm a central assumption of cancer evolution, provide a first quantification of the evolvability of key traits in cancer cells, and indicate that there is ample raw material for experimental evolution in cancer cell lines. Generation time, a trait directly affecting cell fitness, shows substantially lower values of heritability than cell speed, consistent with its having been under directional selection removing heritable variation.

1. Introduction

Evolutionary processes are acknowledged to play a significant role in the initiation and progression of cancer, as well as in the acquisition of traits such as chemotherapy resistance[1,2,3]. Cancer evolution as a field rests on the reasonable, but rarely directly tested, assumption that a variety of cellular traits linked to cancer development and progression are heritable at the level of the cell population: that is, there is phenotypic variation between cancer cells, and that at least some of that variation is due to factors passed on from mother cell to daughter cell, rather than being caused by environmental factors such as nutrient availability. For example, explaining the evolution of metastatic behaviour in terms of dispersal or foraging ecology requires that cells should vary heritably in their dispersal behaviours[4], and models of intra-tumour competition[5], by definition, assume a heterogeneous population with heritable differences between cells. The heterogeneity of cancer cells, whether in tumours or laboratory cell lines, is well established - for instance, a recent multi-omics study found high levels of both genetic and phenotypic heterogeneity between different populations of HeLa cells[6.] However, heritability - which is defined in quantitative genetics as the proportion of trait variance in a population that is due to genetic variation, and which determines the response to selection[7] - has never been directly measured for any trait in cancer cell populations.

Drastic reductions in the cost of sequencing technology over recent years have provided the cancer research community with an abundance of genome sequence data, including at a single-cell level[8]. Sequence analysis techniques

adapted from population and evolutionary genetics can detect and measure evolutionary processes in tumours[9], and signatures of natural selection can be detected by various methods including comparing synonymous and non-synonymous substitution rates[10], or estimating the relative contributions of neutral and adaptive evolution using distributions of allele frequencies within tumours[11]. However, while such analyses can show natural selection has occurred and sometimes identify target genes, they are not always informative as to which phenotypic traits are the targets of selection. Researchers increasingly recognise the need to complement sequence data with a clear and quantitative understanding of cell phenotypes, and - where possible - to link genotype to phenotype[12]. Theoretical models and simulations of cancer evolution, both population-level and agent-based, will undoubtedly guide this programme of phenotypic research[13], but these approaches could be even more valuable if informed by cell-level observational data.

Based on observations of cancer cells over multiple cell generations, we here present cell-level phenotypic data that allows us to estimate broad-sense heritability (H^2) of two traits within four clonally reproducing cancer cell lines. Generation time - the time elapsed between cell divisions - is closely linked to cell fitness. Cell motility is a key step toward metastasis, but confers an uncertain[14, 15], and likely context-dependent[16], selective advantage.

Both generation time and cell motility are observable and quantifiable using time-lapse video microscopy (Fig. 1a). Cell lineages must be observed over several generations to give reliable estimates of heritability, because

cytoplasmic factors cause transient similarity between sister cells for several hours after cell division[17], which will inflate estimates. We therefore estimate H^2 based on several different cell-cell relationships. By tracking and comparing second-generation clonal descendants of the same progenitor cell ("cousin cells"; Fig. 1b), which have never directly shared cytoplasm, we can provide estimates of H^2 that represent stably inherited differences between clonal lineages.

Methods

Cell lines

All cell lines used were lab-adapted. Three are adenocarcinoma in origin: MCF7 (ATCC® HTB-22™) [18], MDA-MB-231 (ATCC® HTB-26 ™) [19] and HeLa (ATCC® CCL-2™)[20]. HT1080 (ATCC® CCL-121™)[21] are fibrosarcoma derived. Cell lines were grown as a monolayer in 5% CO_2 at 37°C in minimum essential media containing 10% foetal bovine serum, 1mM sodium pyruvate and 2mM glutamine. MCF7, HT1080 and HeLa cell media also contained 1% non-essential amino acid solution. Passage numbers were 15-18 for MCF7, 41-44 for MDA-MB-231, and 5-8 for both HeLa and HT1080 cells. Typical laboratory cell culture maintains cell lines at population sizes of 10^5 to 10^6.

Timelapse microscopy and lineage tracking

In total, 9025 cells were tracked, giving a dataset of 471,573 cell positions at known time points. For each of 59 timelapse videos taken, a haemocytometer was used to plate 5000 cells per well onto a 24-well plate. Six wells per cell line were distributed around the plate. Five points within each well were

chosen at random and images were taken every 20 minutes over 72 hours with a Nikon TiE Time Lapse system. NIS software was used to convert images into a video file for each point. ImageJ and MtrackJ[22, 23] were used to analyse videos, recording cell positions at each timepoint and cell division events. 22.7% of cells tracked (2048 individual cells) met the requirements to be used in the analysis. Our final dataset consisted of cell families: groups of related cells for which we had whole-lifespan data on cell motility and generation time over three cell generations (Fig. 1b). These comprised 52 families for the cell line MCF7, 141 families for MDA-MB-231, 134 families for HeLa and 110 families for HT1080.

Statistical analysis

Statistical analysis was performed in R[24]. The generation time for a given cell is the time taken for a full cell cycle from cytokinesis to cytokinesis. Cell motility was calculated as the curvilinear speed of a cell over its entire individual lifetime: the total Euclidean distance of the path travelled from cell division to cell division, in microns, divided by its generation time in hours. All cells included in our analysis were observed for a complete cell cycle; the cells originally plated were therefore excluded, as were those that moved off-screen, died, or had not divided by the end of the tracking period. For this reason some cell families contained fewer than six cells. We could detect no significant differences in cell speed (t-test, $p>0.05$ in every cell line) between families containing different numbers of cells, suggesting that faster cells were not substantially more likely to move off-screen over the tracking period. There was also no significant difference between wells (Kruskal Wallis test;

p>0.05), suggesting that spatial effects - which could exacerbate similarities between related cells and inflate estimates of heritability - are minimal. We then estimated broad sense heritability as the slope parameter of an ordinary least squares regression of trait values between clonal cells and their relatives, for all three cell-cell relationships (Fig. 1c). This is a modification of standard parent-offspring regression techniques[25] for determining heritability in clonally reproducing cell families, and provides a straightforward way to compare heritability values calculated from different cell-cell relationships; the structure of our dataset, with some wells containing very few families, made this a more appropriate method of analysis than a linear mixed model. Where a cell had multiple daughters or "cousins" their mean value was used.

3. Results and Discussion

Cell motility

For all four cell lines, and for all cell-cell relationships considered, the broad-sense heritability (H^2) of motility is highly significant (Fig. 1c; Table 1). Values of H^2 ranged from 0.45 to 0.84, which is high compared to that of a range of traits in natural[26, 27] and agricultural[28] populations. Although no one statistic adequately describes evolvability, this result does imply that cancer cell populations contain substantial variation for motility on which natural selection could act. Mutations, stable epimutations, or both may contribute to this heritable variation[29], and potentially to evolutionary change[30].

Generation time

Generation time shows a very different pattern of heritability (Table 1). For all lines and all cell relationships, H^2 values are lower than those for motility. There is no consistent relationship between the two traits: in two cell lines, faster-dividing cells were slower-moving (for HeLa, Spearman's ρ = -0.068; p <0.001; for MCF7, ρ = -0.178; p <0.005); in HT1080, faster-dividing cells were faster-moving (ρ = 0.072; p <0.005); and no significant correlation was seen in MDA-MB-231 (ρ = -0.002; p >0.1).

Apart from in the HeLa cell line, only sister-sister heritability is consistently significant for generation time, consistent with variation in this trait being attributable to transient cytoplasmic or nutritional differences - the MDA-MB-231 cell line shows only marginally significant H^2 for generation time between cousins. As the low R^2 values indicate, the regression model explained the data poorly, suggesting there may be other factors not included in our model which would explain more of the variation.

Differences in heritability between traits could have a variety of causes: traits are likely to vary in sensitivity to environmental factors, and contributing genes may show different levels of standing variation. Generally, however, lower values of heritability might be expected for traits such as generation time which are highly correlated with fitness, as genetic variation is likely to be removed by natural selection[31]. Conversely, selection can maintain phenotypic variation as well as remove it, and one might expect to see high heritability of traits under fluctuating or frequency-dependent selection - which, under some models of evolution, would include motility traits.

The cell lines MCF7 and HeLa are both highly genetically unstable, showing extensive genome rearrangement[32], high levels of chromosome number variability[33] and many instances of regional copy number increase[34]. Such cell lines might be expected to show higher H^2 values, and greater evolvability, due to a greater level of standing genetic variation. This was not borne out for MCF7 in our data, although HeLa did show high H^2 values compared to the other cell lines tested.

Conclusions and future directions

Our results confirm an important but previously untested assumption of cancer evolution[35,36], and are encouraging for the prospect of experimental cancer evolution *in vitro,* an emerging field that is beginning to provide new insights into cancer biology and evolution[37,38,39] in the same way as microbial experimental evolution has advanced our understanding of adaptation more generally[40].

Similar methods could be used to estimate evolvability parameters in a range of traits, such as cell adhesion or extracellular matrix degradation, and a range of cell populations. It would be interesting to compare a range of cancer types with differing tendencies to metastasis - perhaps including transmissible cancers - and cancers from different patients, including recently-isolated lines and biopsy samples. Primary non-cancerous cells should also be tested: currently, beyond a few measures of proliferation rate and marker gene expression in Chinese hamster ovaries[41], almost no data are available

on the extent of heritable phenotypic variation between clonal somatic cells of multicellular animals, despite this question being highly relevant to the evolution of multicellularity[42].

Good data on trait heritability are available for a few model species of unicellular eukaryotes. Broad-sense heritability estimates obtained for 46 growth traits in *Saccharomyces cerevisiae*[43] had a median value of 0.77, and in *Neurospora crassa,* the heritability of six clock traits[44] ranged from 0.42 to 0.87. However, both of these experiments started with crosses between geographically disparate strains to maximise genetic variation. We obtain similar heritability estimates within individual clonal cell lines, suggesting that cancer cells undergo comparatively rapid phenotypic diversification.

Gene expression data from cells with varying trait values within the same cell line, or comparisons between selection and control lines after experimental evolution, could reveal possible targets of selection. Little is currently known about the molecular basis of phenotypic variation within tumours, although molecular genetic variation is known to be substantial[45, 46].

Because heritability is affected by levels of environmental variation[47], values of heritability *in vivo* are likely to differ from *in vitro* estimates. Further quantitative data on cancer cell populations either in culture conditions simulating particular conditions of interest in tumour microenvironments, or within real tumours, is needed to enable the evolutionary dynamics of cancers to be understood and translated into meaningful *in vivo* predictions for

personalised medicine, and to reveal new targets and therapies for future clinical interventions.

Acknowledgements

We thank Michael Brockhurst, Jonathan Gibbins, Robert Jackson, Michael Fry, Mark Pagel and Karen Poulter for comments. GB thanks tutors and participants at the 2018 Evolutionary Biology and Ecology of Cancer Summer School for helpful discussion. The authors have no conflicts of interest to declare.

References

1 Crespi, B., & Summers K. Evolutionary biology of cancer. *Trends Ecol Evol*. **20:** 545-552 (2005).

2 Klein, C. A. Selection and adaptation during metastatic cancer progression. *Nature* **501:** 365 (2013).

3 Chisholm, R.H., Lorenzi, T., Lorz, A., Larsen, A.K., De Almeida, L.N., Escargueil, A., Clairambault, J. Emergence of drug tolerance in cancer cell populations: an evolutionary outcome of selection, nongenetic instability, and stress-induced adaptation. *Cancer Research* **75:** 930-939 (2015).

4 Amend, S.R., Gatenby, R.A., Pienta, K.J., Brown, J.S. Cancer foraging ecology: diet choice, patch use, and habitat selection of cancer cells. *Current Pathobiology Reports* **6:** 209-218 (2018).

5 Zhang, J., Cunningham, J. J., Brown, J. S., Gatenby, R. A. Integrating evolutionary dynamics into treatment of metastatic castrate-resistant prostate cancer. *Nat. Commun.* **8:**1816. (2017)

6 Liu, Y., Mi, Y., Mueller, T., Kreibich, S., Williams, E. G., Van Drogen, A., Borel, C., Frank, M., Germain, P.L., Bludau, I., Mehnert, M. Multi-omic

measurements of heterogeneity in HeLa cells across laboratories. *Nat. Biotech*. **37:** 314-322 (2019).

7 Falconer, D.S. & Mackay, T. F. C. *Introduction to Quantitative Genetics* (Longman, Harlow, 1996).

8 Lawson, D.A., Kessenbrock, K., Davis, R,.T., Pervolarakis, N., Werb, Z. Tumour heterogeneity and metastasis at single-cell resolution. *Nat. Cell Biol.* **20:** 1349. (2018).

9. Williams, M. J., Sottoriva, A., Graham, T. A. Measuring clonal evolution in cancer with genomics. *Annu Rev Genomics Hum Genet* 20 (Review in Advance first posted online May 6, 2019)

10 Graham, T. A. & Sottoriva, A. Measuring cancer evolution from the genome. *J. Pathol.* **241,** 183-191 (2017).

11 Woo, Y. H. & Li, W. H. DNA replication timing and selection shape the landscape of nucleotide variation in cancer genomes. *Nat. Commun.* **3:** 1004 (2012).

12 Meyer, M., Reimand, J., Lan, X., Head, R., Zhu, X., Kushida, M., Bayani, J., Pressey, J. C., Lionel, A.C., Clarke, I. D., Cusimano, M. Single cell-derived clonal analysis of human glioblastoma links functional and genomic heterogeneity. *Proc Natl Acad Sci USA*. **112:** 851-856 (2015).

13 Gallaher, J., Anderson, A.R. Evolution of intratumoral phenotypic heterogeneity: the role of trait inheritance. *Interface Focus* **3:** 20130016 (2013)

14 Chen, J., Sprouffske K., Huang, Q. & Maley, C. C. Solving the puzzle of metastasis: the evolution of cell migration in neoplasms. *PLoS ONE* **6,** (4):e17933 (2011).

15 Arnal, A., *et al.* Evolutionary perspective of cancer: myth, metaphors, and reality. *Evol. Appl.* **8,** 541-544 (2015).

16 Boddy, A. M., Huang, W., Aktipis, A. Life history trade-offs in yumors. *Curr. Pathobiol. Rep*. **6,** 201-207 (2018).

17 Spencer, S. L., *et al.* Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* **459,** 428-432 (2009).

18 Soule, H. D., Vasquez, J., Long, S., Albert, S. & Brennan, M. A human cell line from a pleural effusion derived from a breast carcinoma. *J. Natl. Cancer Inst.* **51,** 1409-1416 (1973).

19 Cailleau, R., Young, R., Olive, M. & Reeves, W. J. Breast tumor cell lines from pleural effusions. *J. Natl. Cancer Inst*. **53,** 661-674 (1974).

20 Landry, J. J. M., *et al*. The genomic and transcriptomic landscape of a HeLa cell line. *G3: Genes, Genomes, Genetics* **3,** 1213-1224 (2013)

21 Rasheed, S., Nelson-Rees, W. A., Toth, E. A., Arnstein, P. & Gardner, M. B. Characterization of a newly derived human sarcoma cell line (HT-1080). *Cancer* **33,** 1027-1033 (1974)

22 Schindelin, J., *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9,** 676-682 (2012).

23 Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9,** 671-675 (2012).

24 R Core Team. R: A Language and Environment for Statistical Computing. https://www.R-project.org (R Foundation for Statistical Computing Vienna, Austria 2018).

25 Lynch, M., & Walsh, B. Genetics and analysis of quantitative traits. Sunderland, MA: Sinauer; 1998.

26 Kruuk, L. E., Slate, J. & Wilson, A. J. New answers for old questions: the evolutionary quantitative genetics of wild animal populations. *Annu. Rev. Ecol. Evol*. S. **39,** 525-548 (2008).

27 Kingsolver, J. G., Diamond, S. E., Siepielski, A. M. & Carlson, S. M. Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. *Evol. Ecol.* **26,** 1101-1118 (2012).

28 Visscher, P. M., Hill, W. G. & Wray, N. R. Heritability in the genomics era—concepts and misconceptions. *Nat. Rev. Genet.* **9,** 255-266 (2008).

29 Richards, E. J. Inherited epigenetic variation—revisiting soft inheritance. *Nat. Rev. Genet.* **7,** 395–401 (2006)

30 Skinner, M. K., et al. Epigenetics and the Evolution of Darwin's Finches. *Genome Biology and Evolution* **6,** 1972-1989 (2014).

31 Merilä, J. & Sheldon, B. C. Lifetime Reproductive Success and Heritability in Nature. *The American Naturalist* **155,** 301-310 (2000).

32 Hampton, O. A., *et al.* A sequence-level map of chromosomal breakpoints in the MCF-7 breast cancer cell line yields insights into the evolution of a cancer genome. *Genome Res.* **19,** 167-177 (2009).

33 Yoon, D. S., *et al.* Variable levels of chromosomal instability and mitotic spindle checkpoint defects in breast cancer. *Am. J. Pathol.* **161,** 391-397. (2002).

34 Kallioniemi, A., *et al.* Detection and mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization. *Proc. Natl. Acad. Sci. USA* **91,** 2156-2160 (1994).

35 Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23-28 (1976).

36 Aktipis, C. & Nesse, R. M. Evolutionary foundations for cancer biology. *Evol. Appl*. **6**, 144-159 (2013).

37 Taylor, T. B., Johnson, L. J., Jackson, R. W., Brockhurst, M. A. & Dash, P. R. First steps in experimental cancer evolution. *Evol. Appl.* **6,** 535-548 (2013).

38 Lai, P.-L. et al. Selection of a Malignant Subpopulation from a Colorectal Cancer Cell Line. bioRxiv 578823; doi: https://doi.org/10.1101/578823 (2019)

39 Acar, A. et al. Exploiting evolutionary herding to control drug resistance in cancer. bioRxiv 566950; doi: https://doi.org/10.1101/566950 (2019)

40 Lenski R. E. Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations. *The ISME Journal* 11:2181 (2017)

41 Davies, S. L., et al. Functional heterogeneity and heritability in CHO cell populations. *Biotechnol. Bioeng*. **110,** 260-274 (2013)

42 Grosberg RK, Strathmann RR. The evolution of multicellularity: a minor major transition?. *Annu. Rev. Ecol. Evol. Syst.* **38,** 621-654 (2007)

43 Bloom J. S., Ehrenreich, I. M., Loo, W.T., Lite, T. L., & Kruglyak, L. Finding the sources of missing heritability in a yeast cross. *Nature* **494,** 234. (2013)

44 Kim, T.S., Logsdon, B.A., Park, S., Mezey, J.G. and Lee, K. Quantitative trait loci for the circadian clock in Neurospora crassa. *Genetics* **177,** 2335-2347 (2007)

45 Swanton, C. Intratumor heterogeneity: evolution through space and time. *Cancer Res.* **72,** 4875-4882 (2012).

46 Burrell, R. A. & Swanton, C. Re-evaluating clonal dominance in cancer evolution. *Trends in Cancer* **2,** 263-276 (2016).

47 Houle, D. Comparing evolvability and variability of quantitative traits. *Genetics* **130,** 195-204 (1992).

Table and Table Legend:

Table 1

Mean trait value (± standard deviation) and broad-sense heritability of motility and generation time, in four cell lines, for three relationships between cells in tracked clonal families. R^2 is also given for each regression.

Cell line	No. of families	Mean speed (μm/hour)	Mean generation time (hours)	Heritability of cell motility H^2 (R^2)			Heritability of generation time H^2 (R^2)		
				Sister: sister	Mother: daughter	Cousin: cousin	Sister: sister	Mother: daughter	Cousin: cousin
MCF7	52	10.76 ± 3.74	24.78 ± 5.88	0.5*** (0.28)	0.46*** (0.33)	0.58** (0.27)	0.49*** (0.35)	0.15 (0.01)	0.27 (0.03)
MDA-MB-231	141	26.82 ± 11.37	22.46 ± 6.62	0.82*** (0.65)	0.59*** (0.55)	0.56*** (0.41)	0.47*** (0.23)	0.10 (0.00)	0.27* (0.06)
HT1080	110	37.55 ± 10.03	14.22 ± 4.15	0.62*** (0.44)	0.45*** (0.3)	0.56*** (0.3)	0.45*** (0.11)	0.15* (0.02)	0.26 (0.02)
HeLa	134	12.48 ± 5.23	24.43 ± 4.46	0.82*** (0.6)	0.69*** (0.61)	0.85*** (0.60)	0.56*** (0.27)	0.4*** (0.08)	0.62*** (0.36)

*p<0.05, **p<0.01, ***p<0.001

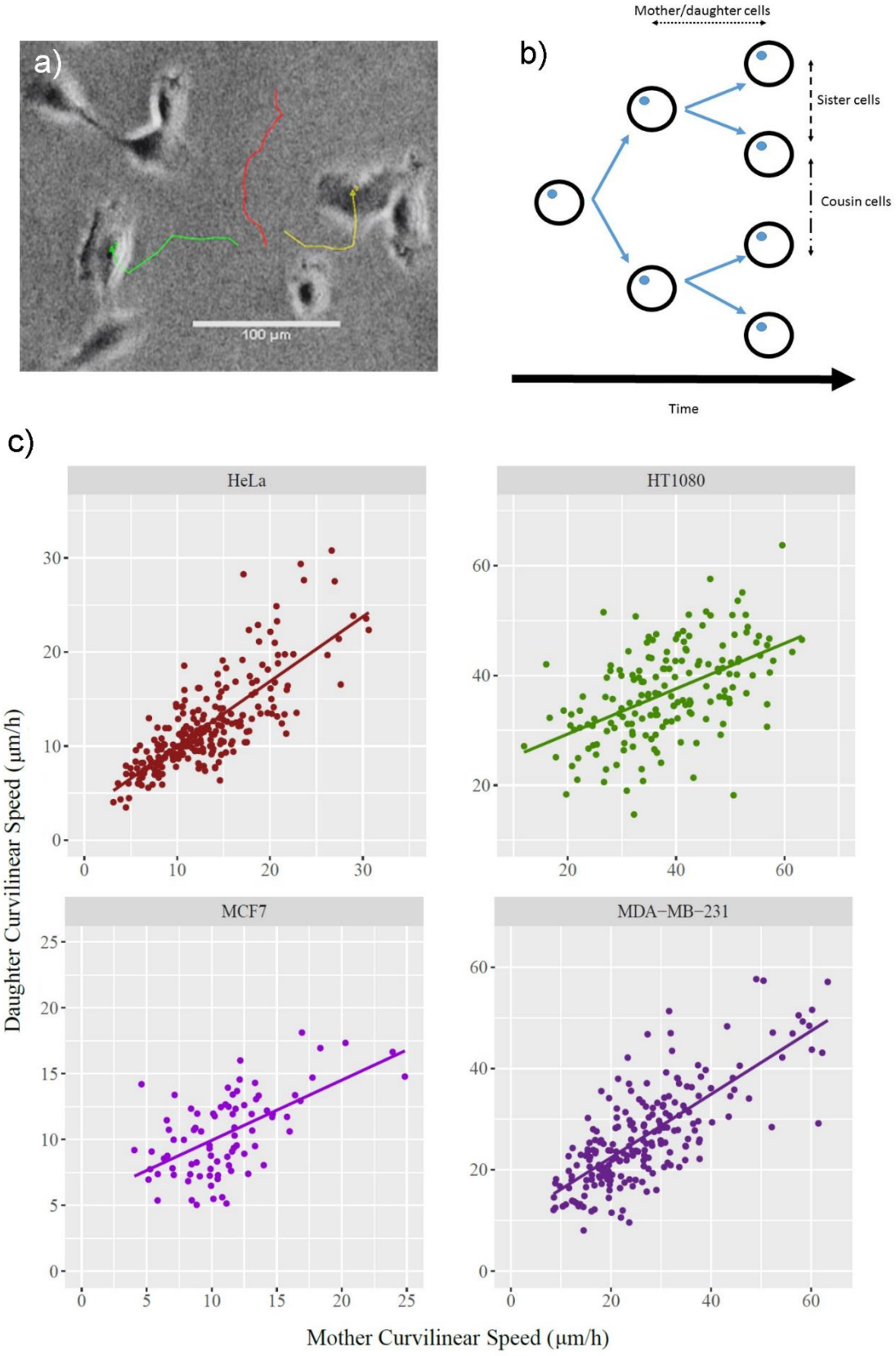
Figures and Figure Legends:

Fig. 1

a) Example image from a time-lapse video, showing one HT1080 cell lineage as tracked for 11h 40min during the 72h observation period. Here, a mother cell moved from the upper middle part of the frame to the lower middle (red line) and divided into two daughter cells, which have since moved outwards to the left and right (green and yellow lines). See Supplementary Files for video.

b) Schematic diagram of a cell family over three cell divisions showing the three cell-cell relationships used to calculate broad-sense heritability.

c) Example of parent-offspring regression for motility, shown here for the mother cell/daughter cell relationship, in all four cell lines tested. Each point represents a mother cell and the mean speed of her daughters, and the slope of the regression line is the estimate of broad-sense heritability. Note that axes differ between graphs due to differences in speed between cell lines. As the cells reproduce clonally, the same method is applicable to other cell-cell relationships. H^2 values for all cell-cell relationships are shown in Table 1.


a)
100 µm
b)
Mother/daughter cells
Sister cells
Cousin cells
Time
c)
HeLa
HT1080
MCF7
MDA−MB−231
Daughter Curvilinear Speed (µm/h)
Mother Curvilinear Speed (µm/h)