The molecular basis of adaptive evolution in yeast: response to ethanol

A thesis submitted to the University of Manchester for the degree of PhD in Bioinformatics in the Faculty of Life Sciences

Daniel Patrick Smith

2013

1. Preface

Contents

Word Count: 82,598

1.1 Declaration

No part of this thesis has been submitted in support of an application for any degree or qualification of The University of Manchester or any other University or Institute of learning

1.2 Copyright statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University's policy on presentation of Theses

1.3 Autobiographical statement

Since my first degree I have gained experience as a programmer supporting Humanities producing web-accessible database research tools, and gained an MSc. in Bioinformatics, both in the University of Manchester. My Master's thesis stored and mined ultra-deep HIV sequence data to mine signals of recombination in emerging drug resistance.

1.4 Acknowledgements

I am profoundly grateful for the unfailing support of my wife and family.

I have also benefitted greatly from the support of Bharat Rash in the laboratory and for the subsequent work he has done on my evolved strains. I would also like to thank Mike Jackson of the fluorescence cytometry facility.

1.5 Abbreviations

CFP	Cyan fluorescent protein
FACS	Fluorescence activated cell sorting; an acronym widely used to denote fluorescent flow cytometry regardless of whether sorting and cell recovery is implemented
GFP	Green fluorescent protein
FM 4-64	The dye pyridinium dibromide
His+	Able to grow in minimal media deficient in histidine.
HO	The HOmothallic site-specific endonuclease gene involved in gene conversion at the MAT locus (homothallic switching)
Indel	An insertion or deletion in a genetic sequence
PCR	Polymerase chain reaction
RHO-	Respiratory deficient
ROS	Reactive oxygen species
SNP	Single nucleotide polymorphism (in a genetic sequence)
SNPQ	Single nucleotide polymorphism quality. An indicator of the likely accuracy of a SNP variant call.
YFP	Yellow fluorescent protein

1.6 Figures and tables

1.6.1. Figures

1.6.2. Tables

1.7 Abstract

Ethanol tolerance in *Saccharomyces cerevisiae* is a complex polygenic trait. As a toxin, ethanol damages multiple cell constituents as well as being both a substrate and product of the metabolism of *S.cerevisiae*. This complexity has made ethanol tolerance difficult to study. Deletion screens have identified hundreds of genes that impair ethanol tolerance by their absence and hence might help survival in high ethanol environments. Similarly, expression studies have revealed genes that respond to ethanol shock, but it is unclear whether those genes are likely targets for improvement of ethanol tolerance in strains growing normally. In addition, those yeasts that are currently commercially exploited for their high ethanol tolerance in the brewing and bioethanol industries are commonly aneuploid or polyploid which makes it difficult to correlate particular features of their genotype with the ethanol tolerant phenotype.

Experimental evolution can reveal genetic changes that change competitive fitness. It is practical to run numerous competitions in parallel between isogenic S.cerevisiae strains for hundreds of generations under ethanol stress, after which whole genome sequencing can identify the genetic changes. Fluorescent tagging of those strains can reveal small changes in population dynamics.

We propagated 144 populations in batch culture for between 100 and 200 generations under 4 ethanol regimes (0%, 4.5%, 6.5% and ramped 0-10%). We monitored the progress of evolution by using mixtures of two fluorescently tagged, but otherwise isogenic, haploid, *hoΔ* (site-specific endonuclease deletant) founder strains (DeLuna et al 2008). Population dynamics measured using fluorescently labelled strains indicated that changes had occurred in competitive fitness due to adaptive evolution. Cell-size measurement and flow cytometry showed that evolved populations were diploid or triploid and the transition to higher ploidy occurred more rapidly with increasing ethanol stress. During the experimental evolution three strains evolved the capacity to grow on organic acids. We have sequenced the complete genomes of eight evolved strains. These strains are confirmed as being diploid, but not aneuploid. Sequencing of evolved strains revealed mutations that have not been previously characterised in deletion or expression studies of ethanol or organic acid tolerance in *S.cerevisiae*.

Both increasing ploidy, to produce triploids as well as diploids, and the acquisition of organic acid tolerance under ethanol stress are unexpected outcomes that have implications for future work.

2. Introduction

This introduction is in four sections. The first, fermentation and ethanol tolerance, concerns the nature of *S. cerevisiae*, the potential benefit of high rates of fermentation and good ethanol tolerance in securing a competitive advantage as well as the physiology of ethanol toxicity and tolerance.

The second section concerns the evolution of *S. cerevisiae*, the whole genome duplication (a paleo-diploidy) of its ancestors and its correlation with the key characteristics of *S. cerevisiae*. Starting with haploid ancestors, all the experimentally evolved strains isolated in this work were diploid or polyploid so this section concludes with an introduction to ploidy in *S. cerevisiae* and a discussion of the potential relevance of ploidy to ethanol tolerance.

The third section concerns the experimental methods used to determine the genetics of ethanol tolerance, with a consideration of the strengths and weaknesses of each: gene deletion studies, expression analysis and experimental evolution.

The fourth part concerns how ethanol tolerance might manifest in changed growth patterns of *S. cerevisiae* strains and how growth is measured.

2.1 Fermentation and ethanol tolerance

2.1.1. The nature of S. cerevisiae

Yeasts are fungi that grow predominantly as individual cells rather than as a mycelium. There are more than a thousand yeast species among the 80,000 fungi described and the yeasts have repeatedly evolved from widely divergent ancestors (Dujon 2010). Amongst the yeasts, the behaviour and properties of *S. cerevisiae* are extreme; the ability to rapidly produce ethanol, tolerate its toxicity and consume it efficiently is critical for the success of *S. cerevisiae* in brewing and baking and in the rotting fruit ecological niche. Ethanol in brewed beverages acts to inhibit the growth of pathogenic microbes, rapid ethanolic fermentation releases carbon dioxide which leavens bread, and the volatile ethanol plume attracts insects such as drosophila which feed and disperse yeast.

The addition of its preferred substrate glucose has a profound and rapid effect on gene expression in *S. cerevisiae,* culminating in strong repression of respiration in the presence of oxygen (the Crabtree effect) and an increase in growth rate (De

Deken 1966; Kresnowati et al. 2006). If intra-cellular levels of glucose are lowered by transporter modification then fermentation is reduced and the respiration rate increases (Otterstedt et al. 2004). When the type species were tested, only 23% of 75 yeast genera were able to grow anaerobically and *S. cerevisiae* had the fastest anaerobic growth rate (Visser et al. 1990).

2.1.2. Ethanol secretion, competition and niche construction

S.cerevisiae has standard metabolic pathways for aerobic and anaerobic metabolism. The pyruvate produced by glycolysis from glucose and other 'fermentable' substrates has two potential fates. In the presence of oxygen, standard strains of *S.cerevisiae* can respire pyruvate to carbon dioxide and water by the citric acid cycle and oxidative phosphorylation and, when constrained by the absence of oxygen, *S.cerevisiae* can ferment pyruvate to ethanol and carbon dioxide. What is unusual is that *S.cerevisiae* preferentially ferments glucose to ethanol despite the presence of oxygen.

Superficially it appears a poor strategy to ferment glucose to ethanol rather than respire it to carbon dioxide. Not only is ethanol extremely toxic (see section 2.1.3) but its production is extremely energetically inefficient. Per mole, glucose fermentation can produce 2 moles ATP against a theoretical maximum yield of 36 moles of ATP for aerobic respiration – hence the maximum biomass achievable by fermentation alone is much smaller than for aerobic respiration. Ethanol secreted by fermentation may subsequently be re-absorbed for respiration once glucose has been exhausted - this has been termed the 'make-accumulate-consume' strategy (Piskur et al. 2006). Whilst such consumption of ethanol can improve the energetic balance, the energy returns remain inferior to solely employing aerobic respiration and much secreted ethanol may be lost through evaporation.

This inefficient conversion of glucose, a universally useful substrate, into the inaccessible and toxic ethanol has been seen as an example of a 'tragedy of the commons' in which *S. cerevisiae* is pursuing an extreme 'cheater' strategy (MacLean & Gudelj 2006). This 'tragedy' has two aspects: firstly, the rapid consumption of glucose by S.cerevisiae deprives competitors of a critical resource, and secondly the ethanol produced by that rapid and inefficient metabolism is a growth inhibiting toxin.

Merico (Merico et al. 2007) grew a number of *Saccharomyces* yeasts in glucose and measured their biomass and ethanol production. Figure 2.1 (adapted from Merico et al. 2007) shows a strong inverse correlation between biomass and ethanol produced;

S. cerevisiae produces the highest concentration of ethanol and less than a third of the biomass per gram of carbon substrate of *Kluyveromyces wickerhamii*.

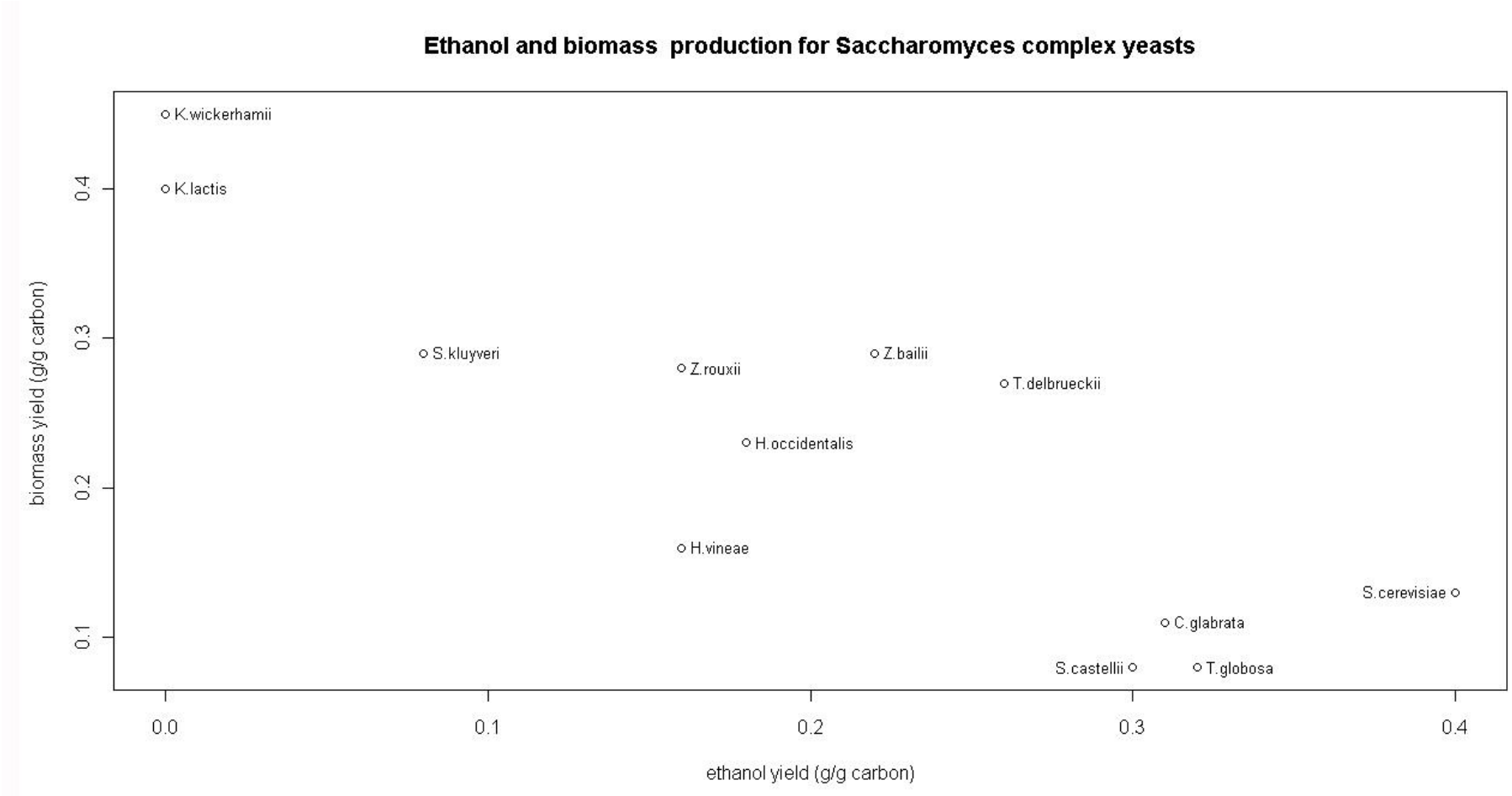


Figure 2.1 Biomass produced against ethanol yield for a number of *Saccharomyces* complex yeasts, data is from Merico et al. (Merico et al. 2007)
Units are grams of ethanol or biomass produced per gram of carbon feedstock. The data shows a strong negative correlation between ethanol yield and biomass produced.

However, fermentation permits faster metabolic rates and the accumulation of toxic ethanol may inhibit growth of other species with lower ethanol resistance which are thereby competitively excluded (Boulton et al. 1996). Ethanol secretion for this purpose can be seen as niche construction, remodelling the environment in a way that recursively changes the selective pressures upon it (Laland et al. 2004).

The toxic suppression of competitors is supported by studies of fermentation ecology. Fruit pulps like grape must are a good growth medium for many other microorganisms; bacteria, protozoa and yeasts such as *Kloeckera apiculata* and members of the *Hanseniaspora* and *Candida* genus are readily found in early fermentations (Margalith 1981; Mills et al. 2002). *Saccharomyces* such as *S. cerevisiae* and *S. bayanus* are hard to isolate from early fermentations but predominant by the end (Pretorius 2000; Fleet 2003; Goddard 2008). A key part of the emergence of *Saccharomyces* appears to be their greater tolerance of the ethanol that has been secreted. Hence the energetic inefficiency of fermentation may be balanced by the fitness advantage over competitor species, a fitness advantage that has been estimated as 7% (Goddard 2008). This is illustrated in Figure 2.2 (from Goddard, 2008) which shows the species, ethanol levels and temperature of four barrels of wine fermenting naturally. The first yeasts to emerge are non-*Saccharomyces* and, despite substantial growth, the level of ethanol remains low at less than 2% and the temperature remains close to ambient. Once *S. cerevisiae* is detected at day 7-8 the ethanol concentration begins to rise as does the temperature, a rise likely exacerbated by high rates of relatively inefficient glucose metabolism by *S.cerevisiae*. The non-*saccharomyces* yeasts decline sharply once the ethanol level exceeds 4% while *S. cerevisiae* growth accelerates and the temperature increases steeply, both indicative of very high fermentation rates by *S. cerevisiae*.

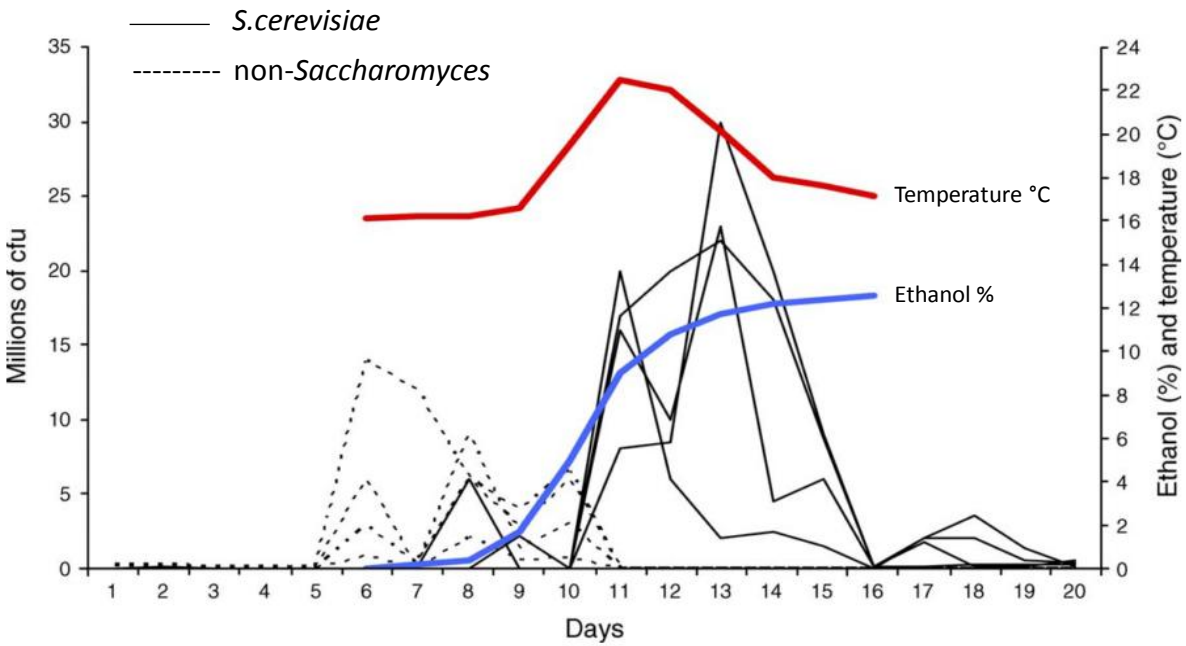


Figure 2.2 The change in yeast community composition, temperature, and ethanol concentration during a traditional wine fermentation (from Goddard 2008)
The x-axis is fermentation time in days. The black lines show growth of non-*saccharomyces* yeasts (dotted lines) and *S. cerevisiae* (solid lines), the units (scale on left axis is in millions of colony forming units). The red line shows the temperature (scale on the right axis), the blue line shows the ethanol percentage.

It is not solely the toxicity of ethanol that is problematic for the competitors of *S. cerevisiae*. Without an alcohol dehydrogenase optimised for consumption of ethanol it will remain inaccessible as a resource. *S. cerevisiae* has two alcohol dehydrogenase enzymes in the cytosol, the homologues *ADH1* and *ADH2*. *ADH1* is constitutively expressed at high levels and has a high K_m for ethanol (17,000– 20,000 mM) favouring the reaction producing ethanol. The *ADH2* K_m for ethanol is low (600–800 mM) favouring the reaction consuming ethanol and producing acetaldehyde. *ADH2* expression is variable and low in the presence of glucose but de-repressed following the diauxic shift (Rolland et al. 2002). Comparative sequence analysis has been used to construct possible ancestral alcohol dehydrogenase enzymes (Thomson et al. 2005). Eleven hypothetical ancestral sequences produced functional enzymes and these favoured the acetaldehyde to ethanol reaction, suggesting that the consume pathway arose after the duplication of an ancestral alcohol dehydrogenase approximately 80mya, probably after the whole genome duplication and the divergence of the common ancestors of the *Saccharomyces* and *Kluyveromyces* but before the divergence of the *Saccharomyces sensu strictu*.

Ethanol secretion may also assist in *S. cerevisiae* dispersal; the yeast is not thought to travel far through the air in the wild, though they are readily isolated from buildings (and people) associated with baking and brewing. The abundant sugar at high concentration which *S. cerevisiae* exploits so readily in ripe fruit, is a rich but isolated and transient food source and there has been speculation that the generation of a volatile plume of ethanol by yeast might attract herbivores important for seed dispersal (Dudley 2004). *Drosophila* species are one possible vector for yeast to travel between flowers and fruit and they are attracted by the smell of ethanol (Hutner et al. 1937; Reed 1938) and will preferentially consume foods containing it, subsequently displaying symptoms of addiction (Devineni & Heberlein 2009). Yeasts can be isolated in large numbers from *Drosophila* (Phaff et al. 1956) and *Drosophila* have been shown to spread yeasts in woodland (Gilbert 1980). Ingestion of *S. cerevisiae* by *Drosophila* also promotes out breeding between spores by digestion of the ascus (Reuter et al. 2007).

Drosophila have also evolved substantial ethanol tolerance (Montooth et al. 2006) which could provide a competitive advantage for eggs and larvae against competitors in the rotting fruit niche. As in *S. cerevisiae*, a key component of ethanol tolerance is the modification of alcohol dehydrogenase (*ADH*) which arose in *Drosophila* 65Mya, around the time of the emergence of fleshy fruits (Ashburner 1998). A number of interactions between yeasts and insect vectors have been described, (see Lachance

et al. 2001) and insects can be a rich source of yeast biodiversity; just one survey of beetles (largely species feeding on *basidomycete* fungi) isolated 650 yeasts, mostly *ascomycete* budding yeasts of which 200 were claimed as new taxa (Suh et al. 2005).

2.1.3. The physiology of ethanol toxicity and tolerance

Membrane damage

The toxicity of alcohols for *S. cerevisiae* is strongly correlated with their lipophilicity, which varies with chain length, suggesting that membranes are key sites of alcohol damage (Fujita et al. 2004) by changing the lipid properties of the membrane or denaturing membrane proteins. Ethanol and heat have very similar, and synergistic, deleterious effects on membranes (Hazel & Williams 1990) and fermentation can produce large amounts of both. Accordingly, ethanol reduces the temperature range over which *S. cerevisiae* can grow (Sa-correia & van Uden 1983; Lucas & van Uden 1985). Ethanol and heat increase membrane fluidity (Jones & Greenfield 1987) increasing the mobility of fatty acids and membrane bound proteins and the membrane permeability. Increased plasma membrane permeability in response to ethanol causes a damaging efflux of nucleotides, nucleosides, Mg^{2+}, and amino acids (Dombek & Ingram 1986) and an exponentially increased influx of protons both in vitro (Zeng et al. 1993), and in vivo (Jiménez & van Uden 1985). This lowers the intracellular pH and dissipates the electrochemical gradient required for active transport of amino acids and glucose (Leao & van Uden 1982). In the absence of ethanol the H^+-ATPase which maintains the pH gradient across the plasma membrane is extremely abundant, constituting 25% of membrane protein (Toulmay & Schneiter 2007). Despite a reduction in the amount of H+-ATPase in response to ethanol (Monteiro et al. 1994) a dramatic increase in its activity in response to increased permeability (Piper 1995) substantially increases ATP consumption (Rosa & Sá-Correia 1991). Ethanol therefore has a direct energetic cost to the cell. Ethanol tolerance, defined as the ethanol concentration required to halve the growth rate, correlates strongly with the rate of proton influx (Jiménez & van Uden 1985).

Other internal membrane associated structures and processes are changed by ethanol stress. Endocytosis is impaired by ethanol and heat shock, the dye pyridinium dibromide (FM 4-64) remaining trapped in endocytotic intermediates whilst the vacuoles become fused into a single large organelle (Meaden et al. 1999). A screen of a deletion mutant library identified 14 vacuolar genes (see section 2.4.1)

that increased ethanol sensitivity (Fujita et al. 2006) suggesting that it may play a role in sequestration of H+ ions to maintain a raised cytoplasmic pH.

Petite strains of *S. cerevisiae* form in high ethanol concentrations. Such strains have smaller cells and form smaller, slower growing colonies which are unable to grow on non-fermentable substrates because they are unable to respire. Mitochondrial DNA damage (see next section) appears to be a key cause of petite formation but petites also have aberrant mitochondrial morphology; lacking the dynamic, complex lamellar structures of normal mitochondria (Bereiter-Hahn & Voth 1994) that are critically dependent on permeability barriers for function and integrity (Cabeqa-Silva et al. 1982). Mitochondrial membrane integrity is damaged by ethanol; leakage of damaging reactive oxygen species (ROS) from respiring mitochondria is increased by ethanol (Du & Takagi 2007). Peak damage occurs after the diauxic shift when glucose is exhausted and the rate of mitochondrial respiration increases when there is an increase in the expression of antioxidant defensive genes such as superoxide dismutase and catalase. Post-diauxy, the mitochondrial superoxide dismutase, encoded by the enzyme *Sod2*, enhances ethanol tolerance a requirement which is abolished in respiration deficient mutants (Costa et al. 1997) and the N-Acetyltransferase *Mpr1* enhances ethanol tolerance by reduction ROS (Du & Takagi 2007).

In response to increased ethanol concentration, *S. cerevisiae* changes its membrane lipid composition, increasing the concentration of the unsaturated fatty acids palmitoleic ($C_{16:1}$) and oleic acid ($C_{18:1}$), and the sterol ergosterol (You et al. 2003; Alexandre et al. 1994) . The *OLE1* gene (Delta9 Fatty Acid Desaturase, YGL055W) encodes the desaturase enzyme required for production of palmitoleic acid and oleic acid. Supplementation of an *OLE1* deletion mutant with oleic acid to twice the wild-type levels substantially increased the ethanol tolerance (Swan & Watson 1998; Swan & Watson 1999) as does over-expression of *OLE1* (Kajiwara et al. 2000). *URA7* (CTP Synthase, YBL039C) and *GAL6* (Cysteine Aminopeptidase, YNL239W) deletion mutants had higher levels of membrane oleic acid and had an enhanced ethanol tolerance compared to their parent strain (Yazawa et al. 2007).

Sterols and unsaturated fatty acids require oxygen for synthesis. Rather than import ergosterol, under aerobic conditions *S. cerevisiae* synthesises it using an energetically expensive pathway of 22 enzymes (Wilcox et al. 2002) but in anaerobic conditions ergosterol must be present in the medium for growth (Visser et al. 1990; Snoek & Steensma 2007) and the expression of ergosterol carriers is accordingly

enhanced (Wilcox et al. 2002). The aerobic culture of rice with *koji* (*Aspergillus oryzae*) provides a culture medium naturally containing ergosterol (Wu et al. 2006) which contributes to the high ethanol tolerance of sake fermentations which can achieve ethanol concentrations of 21%. An ethanol-sensitive mutant sake strain unable to grow in 7% ethanol, was found to have a mutation in the *ERG6* gene of the ergosterol biosynthetic pathway, and had correspondingly low membrane ergosterol levels. Restoration of the gene function restored wild type ethanol tolerance though over-expression did not enhance it (Inoue et al. 2000)

Mitochondrial DNA damage

Exposure to ethanol does not appear to increase the mutation rate in the nuclear genes of wild-type *S. cerevisiae*. Acetaldehyde, the primary metabolite of ethanol, causes GG to TT tandem mutations in human fibroblasts (Matsuda et al. 1998) and has been shown to induce single strand DNA breaks in repair deficient *S. cerevisiae* (Ristow et al. 1995) but it is assumed that DNA repair prevents unresolved nuclear genome breaks in wild-type yeast cells.

However, the mitochondrial genome appears more susceptible to damage. Petite strains lacking functional mitochondrial DNA can be reliably induced by exposure to 24% ethanol or by exposure to a wide variety of other mutagens such as ethidium bromide or ultra-violet light (Goldring et al. 1971; Chen & Clark-Walker 2000). Mitochondrial DNA breaks cause restriction fragment length polymorphism (RFLP) in the 'flor' yeasts found as a velum on the surface of sherry which has an ethanol concentration of over 15% (Ibeas & Jimenez 1997a; Castrejón et al. 2002).

The transfer of wine yeast mitochondria to lab strains increases their resistance to both heat and ethanol and reduces the incidence of the respiratory deficient (rho^-) mutation (Jimenez & Benitez 1988) suggesting that wine yeasts are less vulnerable to the mutagenic activity of ethanol and a key component of tolerance may lie outside the nuclear genome.

Changed trehalose metabolism

Trehalose is a reserve disaccharide of glucose that resists the damaging effects of a wide variety of stressors: dehydration and thermal stress, either freezing or heat. Trehalose improves the survival of *S. cerevisiae* exposed to ethanol concentrations of ethanol greater than 10%. Such concentrations cause the viable cell population to decrease, but the decrease is greater in otherwise isogenic strains unable to make trehalose (Stanley et al. 2009). The trehalose genes (*TPS1*, *TPS2*, *TSL1*, *PGM2* and *NTH1*) are also induced 30 minutes after a 7% ethanol shock (Alexandre et al.

2001).Two experimentally evolved mutants with enhanced ethanol tolerance were shown to have higher levels of trehalose and higher expression of trehalose genes (Abe et al. 2009) and ethanol tolerant strains bred from sake yeasts have been shown to have high levels of catalase, trehalose and glycerol (Ogawa et al. 2000). A deletion mutant lacking the trehalase enzyme responsible for hydrolysing trehalose has enhanced ethanol tolerance (Kim et al. 1996). Trehalose reduces the ethanol induced leakage across membranes in vitro and in vivo (Mansure et al. 1994).

S. cerevisiae grown in high concentrations of solutes have less plasma membrane leakage and better viability when exposed to potentially lethal (16%) ethanol concentrations (Sharma 1997) and this elevated tolerance is correlated with increased intracellular trehalose. Osmotically stressful environments are found in common yeast fermentations; grape juice has a high sugar concentration of 200–350 g/L (Ribereau-Gayon et al. 2006). Trehalose also increases the resistance of proteins to denaturation under heat stress (Kaushik & Bhat 2003; Zancan & Sola-Penna 2005) which may be produced by high rates of fermentation. Ethanol also promotes protein denaturation.

2.2 Evolution of S. cerevisiae and ploidy

2.2.1. Whole Genome Duplication

Fig.1 shows the diverse phylogenetic background of 40 sequenced yeast species and the likely position of a whole genome duplication (WGD) before the divergence of the modern *Saccharomyces* which is critical to the evolution of *S. cerevisiae.* This ancestral duplication occurred approximately at the end of cretaceous period 100Mya (Smith 1987; Wolfe & Shields 1997; Kellis et al. 2004) which coincides with fruiting angiosperms becoming ecologically prevalent (Barrett et al. 2001; Davies et al. 2004), though the dense glucose and fructose sources in fleshy fruits did not emerge until 65Mya (Manchester et al. 1994).

Ohno (Ohno 1970) speculated that gene loss and rearrangement after a whole genome duplication promotes speciation, and the WGD is thought to have played a significant role in the divergence of the *Saccharomyces* (Scannell et al. 2007). However, duplication of an entire genome is a problematic genetic change. Autopolyploids of modern *S. cerevisiae* are extremely genetically unstable and fragile (Mayer & Aguilera 1990). Even if the ancestral WGD was less immediately problematic, such duplication necessarily has significant costs: it increases the energy burden of gene expression (Wagner 2005), and causes genetic isolation

(Greig, Louis, et al. 2002; Greig, Borts, et al. 2002; Greig 2009). Despite such problems, gene duplications comprise 50% of the current *S. cerevisiae* genome (Wolfe & Shields 1997).

An increased maximum glycolytic flux through gene duplication may have conferred an immediate benefit to set against these costs in environments in which glucose is in excess. Against a background of a widespread loss of duplicate genes following the WGD, it appears significant that many duplicate genes involved in glucose acquisition and fermentation have been retained or amplified further. These include *SNF3* and *RGT2*, plasma membrane bound glucose sensors; the additionally amplified *HXT* genes, hexose transporters of varying affinities; the glycolytic genes *HXK/GLK* (hexose kinase/ glucokinase), *TDH* (glyceraldehyde-3-phosphate dehydrogenase), *ENO* (enolase, phosphopyruvate dehydratase) and *PYK* (pyruvate kinase). These transporters and glycolytic enzymes play a significant role in determining the maximal flux through the glycolytic pathway (Conant & Wolfe 2007).

Many descendants of the whole genome duplication do not just display faster fermentation but also a greater independence from their mitochondria. This may be an adaptation towards greater ethanol tolerance, since ethanol is a potent mitochondrial mutagen (Ibeas & Jimenez 1997b; Castrejón et al. 2002). In 24% ethanol the rate of respiration deficient (rho^-) mutants increases from the background rate of 0.4% of the population to approximately half (Bandas & Zakharov 1980). The extent of the mitochondrial damage varies in petite strains of *S. cerevisiae* but some are able to grow on rich media with no mitochondrial genome that can be detected by caesium chloride density centrifugation; though the growth of such petites is poorer than the wild type (Hutter & Oliver 1998; Moller et al. 2001)..

Petite strains would be expected to have impaired or no growth after the diauxic shift when glucose is exhausted, but the inability to synthesise unsaturated fatty acids and ergosterol by oxygen-dependent mitochondrial pathways is also thought to be a factor; both ergosterol and unsaturated fatty acids are required supplements for anaerobic growth of *S. cerevisiae* in minimal media (Visser et al. 1990). The absence of mitochondrial retrograde signalling may also be damaging (Butow & Avadhani 2004). The ability to form petites is relatively rare amongst the yeasts and in the phylogeny of the *Saccharomyces* (see Figure 2.3). However, the known species able to form petites (in red) predominantly share the ancestor with the whole genome duplication (Piskur et al. 1998; Merico et al. 2007) suggesting a correlation

between high potential ethanol fermentation flux and tolerance of mitochondrial damage.

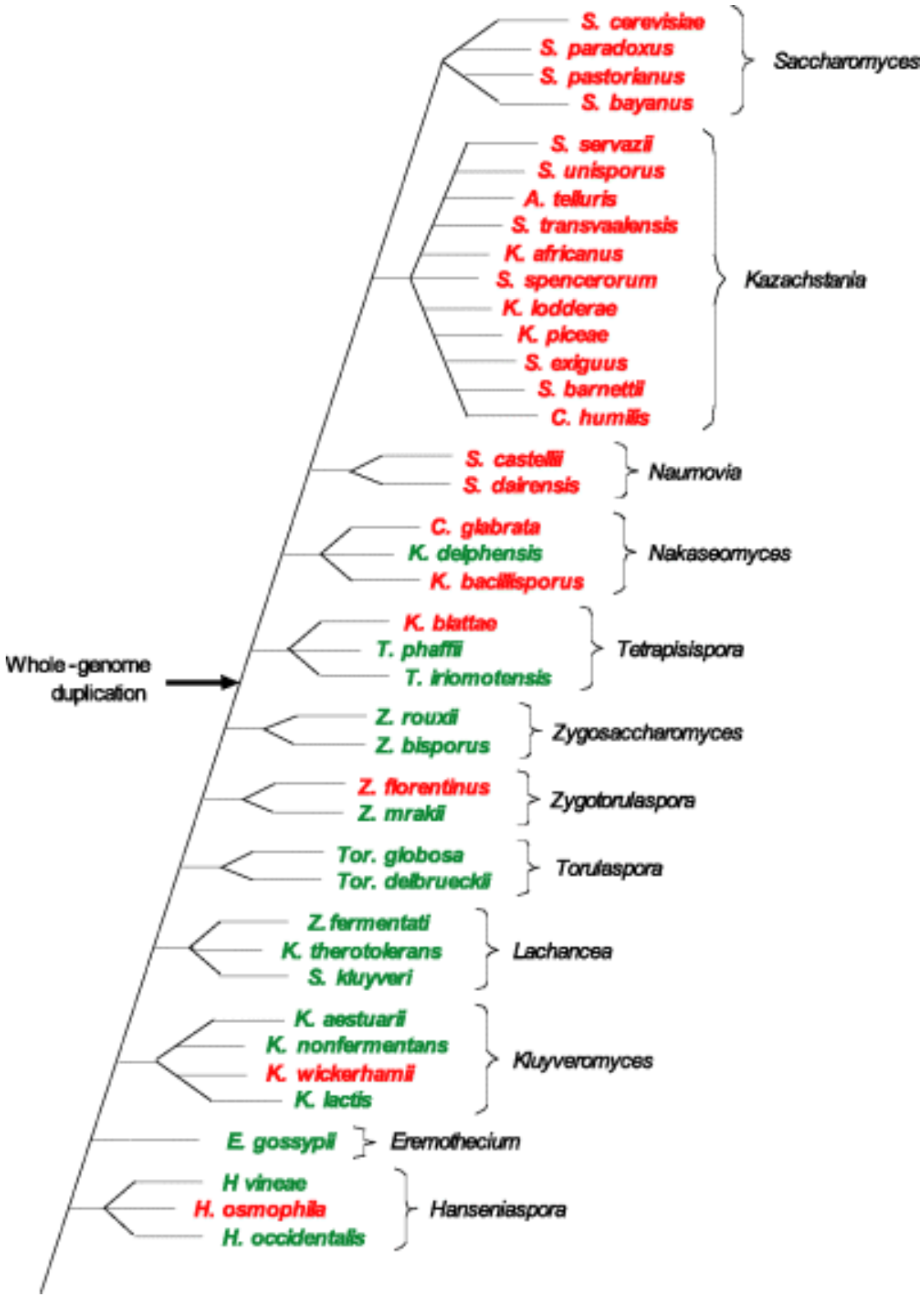


Figure 2.3 The Distribution of petite-positive (in red) and petite-negative (green) species in a phylogenetic tree of the *Saccharomyces* complex (from Merico et al. 1997).

2.2.2. Yeast ploidy

‘‘Polyploidy may be preserved because it protects against mutations in some thousands of loci at once...’’

(Haldane 1933)

“Two *are* better than one; because they have a good reward for their labour” Ecclesiastes 4:9

The whole genome duplication (WGD) undergone by the distant ancestor of *S. cerevisiae* could be considered an ancient increase in ploidy (see Introduction, section 2.2.1,”Whole Genome Duplication”). Modern *S. cerevisiae* can grow indefinitely as genetically stable haploids and diploids and higher ploidy values are found. Wild haploid, diploid and triploid *S. cerevisiae* strains have been isolated from samples from ‘Evolution Canyon’ in Israel (Ezov et al. 2010), including diploid strains from a single tetrad i.e. from a tetraploid parent.

In the sexual alternation of generations *S. cerevisiae* may favour diploidy over haploidy more than other yeast strains. In common with many other yeast species, *S. cerevisiae* will mate whenever haploid cells of opposite mating type are in proximity. However, whereas other fungi such as *Schizosaccharomyces* enter meiosis and sporulate soon after mating *S. cerevisiae* and its close relative *S. paradoxus* do not enter meiosis unless they experience starvation (Gerstein & Otto 2009).

Ploidys higher than diploid are generally less fit and less genetically stable. Galitski reported that the growth of haploids and diploids in an isogenic series produced from the strain was “nearly indistinguishable” but triploids and tetraploids showed a longer lag phase and a 10-15% growth rate reduction (T. Galitski 1999). As ploidy increases from diploid to tetraploid the frequency of chromosome loss to produce aneuploids increases (Mayer & Aguilera 1990). Meiosis in triploid and pentaploid strains produces many aneuploid offspring due to poor segregation of odd numbers of chromosomes. This results in low spore germination since aneuploids are usually less fit than the parent strain (Pavelka et al. 2010). All tetraploids treated with the alkylating mutagen EMS decreased in ploidy (Mable & Otto 2001).

Tetraploids have generally poorer fitness in both rich and minimal media than haploids and diploids (Mable 2001). Andalis et al, found that while haploid cells stopped budding and remained viable for weeks in stationary phase, isogenic

tetraploids remained in bud and began to die after 4 days in stationary phase and were completely dead after 10–15 days (Andalis et al. 2004). There are 39 gene deletions that are lethal in triploids and tetraploids but not in haploids or diploids. These are principally associated with division, both mitotic and meiotic; the assembly of the mitotic spindle and the adhesion and recombination of sister chromatids (Storchová et al. 2006). Gunge and Nakatomi were able to produce strains with all ploidys up to eight but found that cultures with high ploidys contained many small and dead cells (Gunge & Nakatomi 1972).

There is no clear adaptive benefit of haploidy or diploidy in all environments. Paquin and Adams (in a chemostat experiment) found that diploid strains evolved canavanine and cycloheximide resistance faster than haploids (Paquin & Adams 1983) and diploid *S. cerevisiae* were found to evolve fluconazole resistance faster than haploids (Anderson et al. 2004). Whilst tetraploids reduced ploidy when treated with the mutagen EMS, Mable and Otto found that haploids treated similarly became diploid (Mable & Otto 2001); presumably because diploids are better able to repair or mask deleterious mutations by EMS than haploids.

However, a study with seven different environments over 200 generations suggested that haploids adapt faster than diploids(Zeyl et al. 2003). As part of a gene expression study, Li et al showed that a diploid strain demonstrated much higher ethanol tolerance than two of its haploid offspring. In 10% ethanol the diploid exited the lag phase within 30 hours, whilst one haploid (*MATa*) took 80 hours and the other (*MATα*) took more than 130 hours (Li et al. 2010).

Theoretically, diploids might mask deleterious mutations (Orr 1995; Zeyl 2004) and homologous chromosomes also enable recombination repair, the likely cause of the greater resistance to UV and gamma irradiation of diploids when compared to haploids (Heude & Fabre 1993). Diploids might also generate greater variety through heterozygosity, and effectively have duplicated genes which have double the chance of generating beneficial mutations. However, diploids will accumulate a larger mutational load because deleterious recessive mutations are masked from selection and could mask key beneficial mutations that are recessive (Gerstein & Otto 2009).

S. cerevisiae appears to be able to gain or lose complete sets of chromosomes by rare events that differ from the normal sexual alternation of generations(Gerstein et al. 2006). Such individuals with changed ploidy may proliferate in a population for physiological reasons, possibly linked to altered cell size, or due to enhanced genetic

fitness. Having more copies of genes may also confer a greater reserve of genome innovation or evolvability that confers the capacity to make successive adaptive genotypic changes without confining the organism to highly specialised niches.

Lahr (pondering the ploidy plasticity of predominantly asexual amoebae) speculates that such cyclic polyploidy could be a means of avoiding Muller's ratchet, the grave accumulation of deleterious mutations in small asexual populations (Lahr et al. 2011).

A series of 200-generation chemostat experimental evolutions with either glucose, phosphate or sulphate limitations found that whilst single gene amplification (typically of transporters) was found equally in haploids and diploids, structural genetic changes, such as large amplifications, rearrangements, and deletions were more common in diploids than haploids (Gresham et al. 2008). It is possible that such structural changes are more harmful in haploids and hence diploids might have a greater 'structural plasticity'.

Work by Sia et al indicates that diploids have a higher mitochondrial DNA content and lower mitochondrial mutation rates than haploids; regardless of the *MAT* locus genotype, the frequency of frameshift mutations was 47-74 fold lower in diploids than haploids and erythromycin resistance fluctuation tests, which select for mutations in a mitochondrial gene, showed a 2.1 fold decrease in the point mutation rate (Sia et al. 2003).

2.2.3. Mechanisms of ploidy change

Ploidy can increase by mating or by endoreduplication; replication of the genome without cell division. S. cerevisiae has two mating types, designated 'α' and 'a', and mating can only occur between strains of contrasting types. However, wild type strains are homothallic, they can switch mating type once a generation. The mating type is determined by a cassette of active genes at the *MAT* locus but there are silenced copies of each mating type on the same chromosome which can replace the *MAT* locus by recombination after targeted *MAT* locus damage caused by the endonuclease encoded by the *HO* ('HOmothallic') gene. *HO* expression is suppressed in diploids, preventing further mating type switching (and hence further rounds of mating to produce ploidies higher than diploid). For a recent review of yeast mating type switching see Haber (Haber 2012). However, the S. cerevisiae strains in these experimental evolution experiments were heterothallic (both mating type *MATα*) and generally incapable of mating, because their *HO* locus had been

deleted and replaced by a fluorescent protein gene and constitutive promoter (see section 2.6 “Experimental evolution starting strains”).

Normal mating can only increase ploidy from haploid to diploid because conventional diploids are heterozygous for the mat mating locus which suppresses production of mating pheromones and receptors. However, heterozygous diploids can become homozygous at the *MAT* locus by mitotic recombination or DNA repair. Gunge and Nakatomi estimated the frequency of these mating competent diploids as $<1in10^{-6}$ but were able to use selective media to isolate higher ploidy offspring up to octoploid from *MAT* locus heterozygotes (Gunge & Nakatomi 1972). Roman and Sands observed diploids and likely tetraploid spores (of 7 micron diameter) arising from haploid heterothallic crosses (Roman & Sands 1953); they speculate that these arose by fusion of nuclei after meiosis but before spore formation.

A failure to separate daughter chromatids during mitosis results in endoreduplication; if the cell survives it has double the chromosome number. Drugs that disrupt the mitotic spindle such as colcemid, colchicin and concanavalin A promote endoreduplication (Rizzoni & Palitti 1973) as does topisomerase II mutations and inhibitors which prevent chromatid separation (Pastor 2003). Pressure can also induce endoreduplication; one in ten of surviving colonies are polyploid relative to the starting *S. cerevisiae* cultures when subjected to pressures between 200-250MPa pressure (Hamada et al. 1992). This is also likely due to disruption of the mitotic spindle since pressure induces cell cycle arrest, causing protein denaturation and aggregation and a cessation of protein synthesis (Fernandes 2005).

2.2.4. Physiology of ploidy change

"'The first thing I've got to do,' said Alice to herself, as she wandered about in the wood, 'is to grow to my right size again…'" – Lewis Carroll, Alice's Adventures in Wonderland (1865).

If there are physiological benefits of cell enlargement, increase in ploidy may be an efficient, reversible genetic change with minimal fitness costs. As Edgar and Orr-Weaver put it, "endoreplication constitutes an effective strategy of cell growth"(Edgar & Orr-Weaver 2001). The size of a cell can affect its potential fitness in certain environments. Larger cells have a smaller surface area to volume ratio. This may favour smaller cells if maximal rate of nutrient absorption determines fitness. Conversely an increased volume of cytoplasmic storage or rate of synthesis may favour larger cells.

There is a well-established correlation in eukaryotes between nuclear DNA content, nuclear volume and cell volume (Wilson 1928; Jorgensen et al. 2007). Increases in ploidy correspondingly increase the volumes of nucleus and cytoplasm in rich media. This has been investigated in *S. cerevisiae* by producing isogenic ploidy series by rounds of mating following mating-type switching, see Figure 2.4, (Galitski 1999).

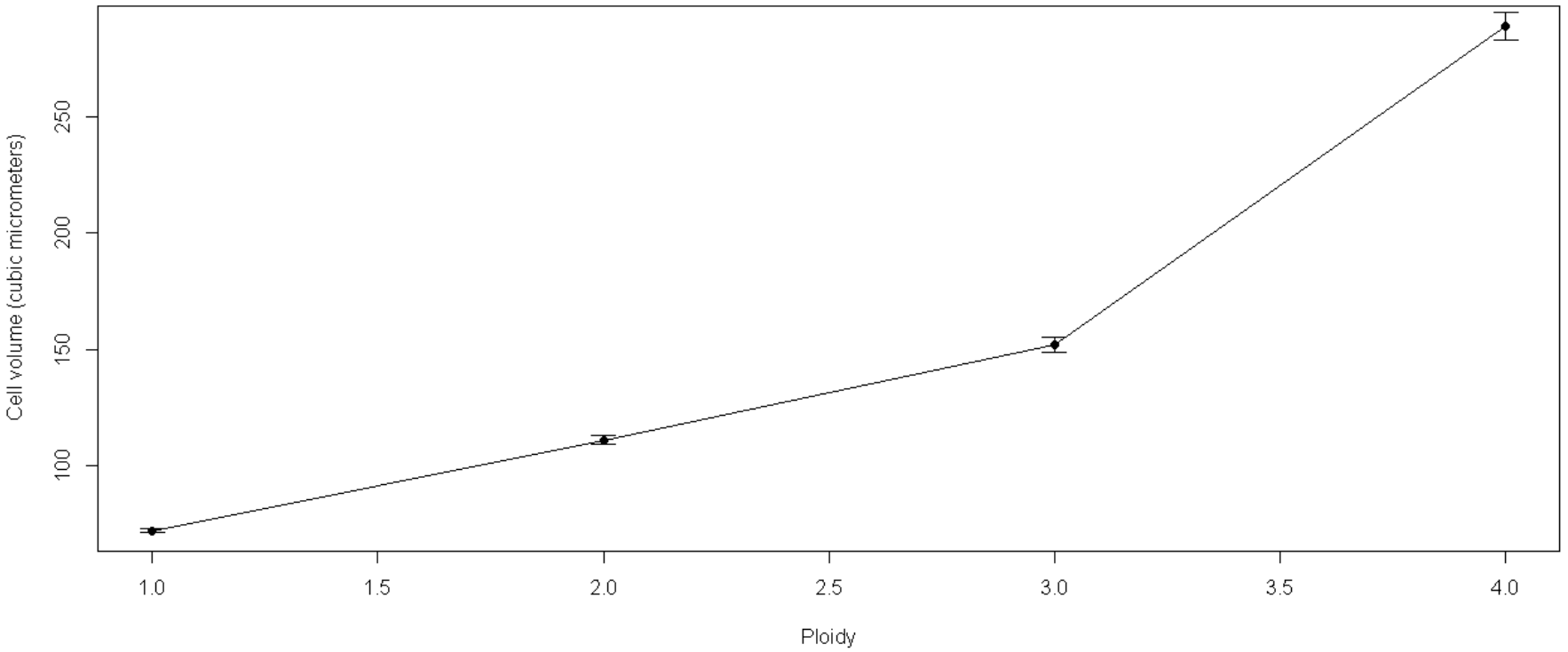


Figure 2.4 Ploidy vs cell volume for an isogenic series derived from strain Σ1278b, adapted from Galitski (1999).
Volumes of exponential phase cells were calculated from 50 observations of length and width of budded mother cells using Nomarski optics. Calculations assume rotational symmetry about the long axis. Errors are 95% confidence intervals

Diploids are also 1.57 times the volume of haploids on minimal media with unlimited carbon (dextrose) but if carbon is limited by chemostat growth then the haploid and diploid sizes are very similar (Weiss et al. 1975; Adams & Hansche 1974). The nature of the carbon source also affects cell size; *S. cerevisiae* are bigger when grown on glucose rather than on non-fermentable carbon sources such as ethanol or glycerol (Cipollina et al. 2005).

Gerstein et al found that not only did diploids spontaneously emerge and dominate experimental *S. cerevisiae* evolutions with haploid starting populations when grown in YPD with or without salt stress (Gerstein et al. 2006), but also that the haploid strains remaining in those evolutions became significantly bigger (Gerstein & Otto 2011), suggesting a fitness advantage linked to larger size, regardless of ploidy.

Under constant growth conditions *S. cerevisiae* cells have a consistent size, this means that there is co-ordination between growth and division (Carter & Jagadish 1978). Regulation of cell size occurs in G1 (Johnston et al. 1979) and the expression of the G1 cyclin genes, *CLN1* and *PC11*, are repressed as ploidy increases which may cause cells to pass through G1 START at a larger size (Galitski 1999). The volume of the mother cell for bud initiation is 1.7 times larger for diploids than haploids regardless of growth rate (Lorincz & Carter 1979).

Deletion screening identified several cytoskeletal genes important for ethanol tolerance (Auesukaree et al. 2009). Ethanol causes cell cycle delay by dispersing F-actin from the normal growth regions close to the bud with a corresponding increase in cell size (Kubota et al. 2004). The cell size of mother cells increases with ethanol (Kubota et al. 2004; Dinh et al. 2008) whilst the size of daughter cells is not substantially increased. Dispersal of actin appears to be a general stress response which arrests the weakening of the cell wall and delivery of wall material required for budding; such depolarisation also occurs with hypertonic osmotic (Chowdhury et al. 1992) and temperature shocks (Yeh & Haarer 1996; Lillie & Brown 1994). It has been claimed that actin depolarisation can halt the cell cycle in G2, i.e. there is a morphogenesis checkpoint in *S. cerevisiae* (Lew & Reed 1995).

The *RPI1* (Ras-cAMP pathway inhibitor 1) gene is an inhibitor of the *RAS/*cAMP stress pathway and overexpression increases cell viability under ethanol stress (Puria et al. 2009). Increasing cAMP can raise the critical cell mass for initiating DNA synthesis by inhibiting the G1-S 'start' transition for slow-growing cells (Baroni et al. 1992). *URA7* (CTP Synthase, YBL039C) deletion mutants, not only increase ethanol tolerance (Yazawa et al. 2007) but also increase cell size (Jorgensen et al. 2002).

There is phenotypic plasticity in the size of the yeast cell as it commits to budding and the size of the bud produced. The time variation from committing to budding to casting off the bud is relatively small so the faster the rate of growth the larger the bud will be (Rupeš 2002) and the smaller the disparity in size between mother and daughter cells (Porro et al. 2003). When grown in YPD with glucose, only 12% of log phase buds are smaller than the 'start' size for producing buds of their own, but this fraction rises to approximately 29% on glycerol (McNulty & Lew 2005). Since the addition of ethanol to the growth medium slows the rate of growth, this would be expected to reduce the size of buds and increase the size disparity between mother cells and buds.

Lorincz and Carter found that in a rich YPD medium in which the time to double the mass of the haploid strain C4,2 is 2.4h, the daughter cells initiated their first bud at a cell volume of 38.7μm^3. However, in a minimal ethanol medium, in which the mass doubling time is 6.84 h, the critical volume was smaller at 28.5μm^3 (Lorincz & Carter 1979).

2.3 Growth of S. cerevisiae in organic acids

Benzoic, sorbic, propionic and acetic acids are widely used for food preservation. These weak acids do not kill micro-organisms outright but greatly extend the lag phase of growth (Fernandes et al. 2005) extending the shelf life of food. *S. cerevisiae* has a high tolerance of acidity; lager fermentation typically starts at a pH of ~5.5 and falls to pH4.1 (Coote & Kirsop 1976) and a standard growth medium is basal medium A which has a pH of 5.5 (Fiedurek et al. 2011). Wine yeasts readily ferment Sauvignon Blanc grape juice with pH of 3.1 (Harsch et al. 2010) and yeast can withstand washing with acid (pH 2.2) to kill bacterial contaminants (Boulton & Quain 2001).

S. cerevisiae produces acetic acid during fermentation, which can rise to inhibitory concentrations and fermentation arrest, so called 'stuck fermentations' (Mira et al. 2010) but *S. cerevisiae* can also consume an excess of acetic acid in wine by a second fermentation (Vasserot et al. 2010; Vilela-Moura et al. 2010). Acetic acid can be the sole carbon source for *S. cerevisiae* (Schüller 2003) but it is a non-fermentable substrate that requires functional mitochondria and aerobic conditions (Flores et al. 2000). Enhancing the organic acid tolerance of *S. cerevisiae* is an area of active research; lignocellulose derived hydrolysates are a potential feedstock for bioethanol production but contain inhibitory concentrations of weak acids including acetic acid (Palmqvist & Hahn-Hägerdal 2000).

Tryptophan auxotrophs are extremely sensitive to inhibition by the weak organic acids acetate and sorbate, an inhibition that is supressed by tryptophan supplementation (both acids) or (sorbate only) by overexpressing the *TAT2* tryptophan permease (Bauer et al. 2003), suggesting that such acids impair tryptophan uptake. Lactic acid stress greatly reduces the intracellular pools of all the amino acids (Suzuki et al. 2012), probably due to reduced active transport dependent on a proton gradient.

2.4 Methods for analysis of ethanol tolerance

Since ethanol is both a metabolic product and a toxin which damages many cellular components and processes (see section 2.1.3), it is not surprising that ethanol tolerance is a complex trait with contributions from many genes. In *Saccharomyces cerevisiae,* ethanol tolerance is critical to its ecological (and industrial) niche (see section 2.1.2). Valuable insights into ethanol tolerance have been gained using traditional approaches which have looked at genes associated with ethanol metabolism or damaged cellular systems, screened deletion libraries for sensitive strains, crossed inbred strains of differing tolerance to look for quantitative trait loci, and looked for changes in gene expression between strains of varying tolerance or in response to an ethanol shock. However, it remains a challenge to find the genes underlying such complex traits, and elucidate how they interact.

2.4.1. Deletion screening and quantitative trait locus studies

Traditional methods of genetic analysis have provided a great deal of information about the genes involved in ethanol tolerance. There have been a number of screens of deletion libraries for increased ethanol sensitivity in *S. cerevisiae* in the belief that those genes might be involved in the greater tolerance of the wild type. Table 2.1 illustrates in red the number of genes found in each study. By definition these deletion libraries are of non-essential genes but whilst not lethal, deletion of many genes produces strains with poor growth. The smaller numbers in brackets in table 1 (numbers provided for each experiment by Yoshikawa et al. 2009) exclude gene deletions from the screen that grow poorly without added ethanol. The numbers in black show the overlap between different screens identifying the same genes. The relatively poor correspondence between these studies can be partly explained by differing experimental conditions, particularly the use of different levels of ethanol.

	Kubota et al. (2004)	Van Voorst (2006)	Fujita et al. (2006)	Yoshikawa et al. (2009)
(Kubota et al. 2004)	**256 (114)**	21 (5)	61 (27)	183 (63)
(Van Voorst et al. 2006)		**46 (20)**	11 (5)	38 (14)
(Fujita et al. 2006)			**137 (56)**	102 (33)
(Yoshikawa et al. 2009)				**864 (446)**

Table 2.1. Genes found by two gene deletion studies (adapted from Yoshikawa et al. 2009). Totals are in red. Numbers in brackets exclude gene deletions where growth was impaired in the absence of ethanol

Deletion of two genes, *URA7* and *GAL6* increased tolerance in 8% ethanol (Yazawa et al. 2007); these genes cause cell wall changes inducing resistance to K1 killer toxin and increase the mono-unsaturated oleic acid (C18:1) content of the plasma membrane. Work on deletion strains showing widespread aneuploidy, especially of regions encoding homologous genes to the deletion, suggests that might be another confounding factor (Hughes et al. 2000).

A large recent quantitative trait locus study identified 5 loci which explain ~50% of phenotypic variation and one locus on chromosome 9 which explains ~25% (Hu et al. 2007).

2.4.2. Gene expression change

There have been a number of studies of gene expression changes in response to ethanol shocks (Alexandre et al. 2001; Chandler et al. 2004). Such studies have pointed to the key role that trehalose plays in ethanol shock resistance (see section 0). The response of yeast to heat or ethanol shock is very similar, ethanol stress causes the expression of heat shock proteins such as *HSP104* (Piper 1995), and both shocks cause reductions in the export of polyadenylated mRNA from the nucleus shutting down gene expression (Izawa et al. 2008). It is unclear the extent to which stress responses which prevent cell death with an ethanol shock are relevant to general ethanol tolerance, but nuclear accumulation of mRNA is observed at high ethanol concentrations in sake brewing and wine making (Izawa et al. 2007). Such a shut-down of gene expression suggests that there might be an antagonism between stress tolerance promoting cell survival and growth rate. Three strains of

HSP104 deletion mutants grew to higher stationary phase culture densities on the non-fermentable substrate acetate than the wild-type (Sanchez et al. 1992) suggesting that *HSP104* might confer enhanced fitness in the fermenting log phase at the expense of growth via respiration after the auxotrophic shift.

The transcription factors *Msn2p* and *Msn4p* are localised in the nucleus in response to many stressful conditions, including ethanol (Görner et al. 1998) and over-expression of *MSN2* during sake fermentation promotes ethanol tolerance (Watanabe et al. 2009). However, there are two gene products known that reversibly accumulate in the nucleus in response to ethanol but not a heat shock: *Asr1p* accumulates in the nucleus under ethanol stress (Betz et al. 2004), and sake yeast progressively accumulates *Rat8p* in the nucleus up to an ethanol concentration of 12% before it returns to normal late in fermentation (Izawa et al. 2005).

2.4.3. Enhancing ethanol tolerance by experimental evolution

Yeast is an ideal organism for experimental evolution because it grows rapidly to form large populations that allow replication and can be cultured asexually for many generations as a haploid or a diploid. The starting population can be closely defined because it originates from a single colony, and intermediate strains can be archived by freezing at -80°C which allows competition between evolved strains and ancestors to quantify fitness changes. Asexual reproduction means replicates do not diverge due to mating and recombination and markers (such as auxotrophies or fluorescent proteins) are not recombined during experimental evolution or fitness competition. *S. cerevisiae* was the first eukaryotic genome sequenced and its genome has subsequently been intensively re-sequenced by many project groups. Accordingly there is a good reference genome for laboratory strains, good expression profile datasets and a well-characterised deletion library. These resources, coupled with the falling cost and greater throughput of next generation sequencing make it now feasible and practical to sequence multiple yeast genomes to look for genotypic changes during adaptive evolution.

Despite the crucial role that ethanol tolerance appears to have played in the evolution of *S. cerevisiae* (see section 2.1.2) a number of experiments have shown it is possible to experimentally evolve strains with enhanced tolerance within 100-200 generations. There may be several reasons for this. There may be a trade-off between tolerance of ethanol and growth rate. Equally, a relaxation of selection pressure for ethanol production and tolerance in lab cultured strains may have caused such strains to drift from the well-adapted tail of ethanol tolerance.

Regardless, the genetic analysis of such strains can provide a valuable additional approach to elucidate the genetics of ethanol tolerance. Much remains to be learned about adaptive evolution and exploration of this polygenic trait could produce more general insights about evolutionary processes.

There are two widely used methods of experimental evolution: continuous culture or batch culture with serial transfer to new media of a portion of a population. In continuous culture there is an influx of liquid media which both dilutes the culture and supplies new resources. The dilution rate may be fixed, a 'chemostat', or be dependent on culture density, a 'cytostat' or 'turbidostat' (Bryson & Szybalski 1952). In most continuous cultures resource limitation means that the growth rate is substantially below the maximum, although there have been reports of successfully deriving evolved strains using a cytostat with the high rates of dilution and low cell density required for maximal growth rates (Gilbert et al. 2009). Growth rates may also be limited by a toxic selection pressure (e.g. by ethanol addition). Ethanol tolerant strains have been developed using cytostats with *Saccharomyces bayanus* (Brown & Oliver 1982), hybrids of wine and laboratory strains of *S. cerevisiae* (Jiménez & Benitez 1987) and using mutagenised and non-mutagenised W303-1A, a haploid laboratory *S. cerevisiae* strain (Stanley et al. 2010).

In batch culture populations are grown to stationary phase before part is transferred to new media. The oldest and longest running batch ethanol tolerance study is the production of traditional sherry which is stored for years in casks which are filled with a mixture of new sherry and old, a high ethanol batch culture regime that has been running for centuries, and has produced the *flor* yeasts, found as a floating biofilm on the sherry surface (Ibeas & Jimenez 1997). Dinh et al. used varying (and stepwise increasing) ethanol concentrations on batch cultured *S. cerevisiae* strain FY834 to produce strains showing better growth rates in 10% ethanol (Dinh et al. 2008).

Study	Methods	Findings
(Brown & Oliver 1982)	Continuous culture (turbidostat)	Increased fermentation rates under ethanol stress, increased viability plated on 12% ethanol.
(Jiménez & Benitez 1988)	Continuous culture of hybrid strains derived from crosses of wine and lab strains	Strain able to grow in 16% ethanol.
(Cakar et al. 2005)	Continuous and batch culture with ethyl methyl sulphonate (EMS) mutagenesis	Increased viability with 20% ethanol stress.
(Dinh et al. 2008)	Batch culture	Increased growth rate of strains exposed to more than 6.5% ethanol. Membrane palmitic fatty acid (C16:0) decreased (see section 2.1.3 “Membrane Damage”) and cells were larger (see section 2.2.4 “Physiology of ploidy change”)
(Abe et al. 2009)	Batch culture (proofreading-deficient DNA polymerase δ)	Produced strains able to grow in 13% ethanol with enlarged vacuoles and increased expression of trehalose biosynthetic enzymes(see section 2.1.3 “Changed Trehalose Metabolism” .
(Stanley et al. 2010)	Continuous culture with and without ethyl methane sulphonate (EMS) mutagenesis	A shorter lag, higher growth rate and higher viability under high levels of ethanol stress, with increased production of glycerol.

Table 2.2 A summary of successful experimental evolution experiments to increase ethanol tolerance with methods and key findings

Batch culture is closer than continuous culture to the industrially important brewing process in which the genetics of ethanol tolerance is of critical importance. Batch culture is also closer to growth in the rotting fruit niche which may resemble the environment in which *S. cerevisiae* evolved. Batch culture, unlike continuous culture, has stages of growth: lag, log and slow growth. Adaptation for ethanol tolerance may arise in each stage; it may reduce lag, increase the duration or rate of growth in log phase and may maintain higher cell viability and growth after log phase growth has ceased. Continuous cultures that do not put ethanol stresses on the organism in each of these growth phases may not evoke the same range of adaptations. Furthermore, an aim of this project was to identify ethanol tolerant adaptive changes that occur in independent experimental evolutions. Each continuous culture requires monitoring equipment and delivery systems to administer nutrients or ethanol as required. This currently limits the minimum volume of culture, and hence the number of parallel evolutionary experiments that can be carried out. However, batch cultures can be run in parallel on 96 well plates in 120µl volumes (with bottleneck population sizes that can be varied) and this parallelism permits the search for rare mutations. Accordingly, for the experiments in this project experimental evolution was carried out in batch culture.

2.4.4. Mutation rate, population size and adaptive trajectories

The true rate of mutation without selection is hard to experimentally determine. Models of adaptation have traditionally assumed that mutations are rare and mostly of small effect (Orr 1998). Experimental estimates in *S. cerevisiae* by Lynch et al have suggested that there are approximately 0.32 mutations per cell per division of which 0.1% have discernible (good or bad) fitness effects (Lynch et al. 2008).

In asexually reproducing organisms lack of mating and recombination means that if two beneficial mutations arise in a population they will compete. Miller et al. concluded (from work with the bacteriophage ID11) that selective sweeps are unlikely for microbial populations with an effective population size greater than 10^4 (Miller et al. 2011). Such clonal interference means that the mutations fixed are likely both to have a larger fitness effect and take longer to become prevalent (Gerrish & Lenski 1998; Orr 2000; Rozen et al. 2002). Modelling has suggested that in the rugged fitness landscape that might arise due to epistasis, such clonal interference may prevent adaptation being trapped in local optima (Jain & Krug 2007).

A series of experimental evolutions have suggested that beneficial mutation rates in asexual yeast populations are large enough not only for clonal interference to be

widespread (Joseph & Hall 2004; Hall & Joseph 2010) but also for competing beneficial mutations to arise rapidly in genetic backgrounds that have already diverged from the starting genotype. Such adaptive mutations may arise in a background of relatively poor fitness (Kao & Sherlock 2009) increasing the likelihood of clonal interference or in a high fitness background making a selective sweep more likely. Lang et al observed the trajectory of a type of beneficial mutation (sterility in asexual culture) arising in hundreds of yeast populations for 1000 generations (Lang et al. 2011). They observed very few selective sweeps (8 out of 221 examples), more commonly seeing a succession of alleles arising without fixing or alleles stalling at intermediate frequencies before rising after a second mutation.

If an experimental population is too small then the (random) time between beneficial mutations may be too long for practical observation and fitness will not change. The emergence of several DNA repair deficient 'mutator' strains in the long-running Lenski experimental evolution of *E.coli* (Barrick et al. 2009) suggested that an elevated mutation rate might speed up adaptation by reducing the average time between beneficial mutations. According to Orr, there is an optimal mutation rate for adaptation in asexual populations (Orr 2000b); mutators or mutagens may help in medium-sized populations by increasing the supply of beneficial mutations but it will increase the likelihood that competing beneficial mutations will arise within the population – i.e. the severity of clonal interference will increase. High rates of mutation in small populations may increase the genetic load of deleterious mutation to the point where the population degenerates rather than adapts because purifying selection is not able to remove deleterious mutations.

Two experimental evolution studies used ethyl methane sulphonate (EMS) as a mutagen in the isolation of ethanol tolerant strains (Cakar et al. 2005 Stanley et al. 2010). Abe et al. utilised a proofreading-deficient DNA polymerase δ which introduces mutations only in the lagging DNA strand to enhance diversity (Abe et al. 2009).

As an organism becomes better adapted, the number of remaining mutations that will increase fitness decline as does the size of their fitness benefit. The precise shape of the distribution of beneficial fitness effects for mutations in generally well-adapted organisms is a matter of debate (Orr 2003; Orr 2005a; Rokyta et al. 2008) and is practically difficult to resolve.

The neutral theory suggests that all mutations, regardless of how beneficial, are vulnerable to loss at low frequency or in small populations. Lenski et al. estimate

that in the long running *E.coli* experiment (Lenski et al. 1991; Elena & Lenski 2003) a mutation with a fitness benefit of 10% will, on average, be eliminated at low frequency by random drift five times before it becomes prevalent and a mutation with only a 0.1% advantage would need to arise more than 500 times. Smaller populations are more vulnerable to loss of mutations that are not fixed and hence mutations of smaller fitness change are less likely to be seen. However, if the experimental population is enlarged, not only will the frequency of clonal interference rise, but the 'transit time' for a mutation to sweep from rare to dominant or fixed is longer. This has some advantages, since mutations can be observed increasing fitness stepwise.

In a glucose limited chemostat culture with a population size of $4\text{-}5x10^9$, one clone is replaced by another every 40 generations (Paquin & Adams 1983). But longer transit time means that a mutation spends longer at low frequency vulnerable to loss by neutral drift. This problem is accentuated if the mutation is of small fitness benefit since the increase of an adaptive mutation from rare to prevalent is inversely proportional to the fitness benefit. Lenski et al. estimate that in their *E.coli* experiment a mutation that has a 10% fitness increase will take 250 generations to increase from inception to becoming dominant but a mutation with a fitness benefit of 0.1% takes 25,000 generations (Lenski et al. 1991; Elena & Lenski 2003).

Increasing the size of the selection pressure i.e. exposing a population to higher ethanol concentrations could speed up the fixation of beneficial mutations if the effect size is increased at higher ethanol concentrations. However, if the selection pressure reduces the population size then it will reduce the number of beneficial mutations generated. It could also fix different mutations; the beneficial mutations at low ethanol concentrations might be a subset of those at high concentration but they might be completely different with a different adaptive landscape. For example, more extreme selection pressures may encourage the fixation of mutations which have antagonistic pleiotropic effects, followed by an adaptive walk of compensatory mutations that reduce the deleterious effects of the first (Rozen et al. 2007).

A compromise is to use a stress that progressively increases with the tolerance of the evolving population to maintain population size. The use of such ramped stresses in continuous culture is known as a cytostat or turbidostat and has been used to increase ethanol tolerance using an ethanol stress (Brown & Oliver 1982). Enhanced ethanol production has also arisen as a by-product of adaptation to a ramped acetate stress (Gilbert & Srienc 2009) in a cytostat.

It is an open question whether the genetic changes conferring most ethanol tolerance are in regulatory or structural genes. Changed regulatory factors can produce radical changes in phenotype because they can affect the expression of multiple genes. Such changes are likely to be more pleiotropic than changes to structural genes but despite the strongly held belief that the more pleiotropic a gene is the more likely changes to it are to be deleterious (Orr 2000), the pleiotropy of a gene appears to have little effect on its rate of evolution(Salathe et al. 2006; Razeto-Barry et al. 2011), indeed antagonistic pleiotropy may help to escape from local maxima on the adaptive landscape in much the same way that predation by phage may promote host diversity and evolvability(Williams 2013). Selection following targeted mutagenesis of transcription factors produced enhanced ethanol tolerance in a strain with a changed TATA binding protein *Spt15p* (Alper et al. 2006).

It is also an open question whether most adaptive genetic changes are mostly substitutions, or insertions/deletions in single genes, or large scale chromosomal changes. Eight continuous culture experiments by Dunham produced six aneuploids, three with a shared common breakpoint adjacent to a gene relevant to citrate synthase and three with amplifications of a region containing a high affinity glucose transporter (Dunham et al. 2002). Both of these genes are relevant to the enhanced glucose transport and aerobic respiration found in strains with enhanced fitness in such cultures.

2.4.5. The experimental design of this project

By running a large number of parallel evolutions it is possible to look at the diversity of the initial steps that increase fitness, giving an insight into the probability of parallel evolution of ethanol tolerance. If there are changes to multiple genes with radically different functions this may suggest a rugged fitness landscape with many peaks (Colegrave & Buckling 2005; Rozen et al. 2008) and numerous evolutionary solutions to ethanol tolerance. Conversely if there is no such diversity in multiple parallel experimental evolutions then it suggests that there are few accessible genetic changes leading to higher fitness with this strain background in these conditions.

There are practical limits on the number of experimental evolutions that can be run in continuous culture. Using the current methods, continuous culture requires larger volumes and larger population sizes which in turn cause greater clonal interference (see Section 2.4.4.) to obscure the diversity of initial steps. Greater tolerance of ethanol stress may manifest as a swifter exit from lag, greater maximum growth rate or enhanced viability during slow growth. These growth stages do not normally

manifest in continuous culture. Methods were devised to detect these phenotypic changes in batch culture, a research goal, see Section 2.7 “Research aims and goals” and Chapter 5). This makes it desirable to select for ethanol tolerance with a culture method that pass through multiple stages of growth which also favours batch rather than continuous culture (see Section 2.4.3). Batch cultures can be performed with small volumes (120µl in these experimental evolutions) and cell counting can be used to subject populations to a measured bottleneck which moderates the degree of clonal interference (see Section 2.4.4.) whilst allowing normal growth to a dense culture. Reduced clonal interference and population size can encourage a greater diversity of adaptive mutations to sweep in fewer generations. However, there is a lower limit to the size of the population bottleneck; below this limit selection will be swamped by neutral drift and may degenerate due to stochastic amplification of deleterious mutations and loss of beneficial alleles (Muller’s ratchet). Modelling was used to establish the safe lower limits to bottleneck size (see Chapter 4, section 4.2.2.). A key research aim is establishing effective methods for experimental evolution.

Detecting adaptive evolution by detecting changed growth patterns in single populations is much harder (i.e. effect sizes must be bigger) than monitoring change in relative population sizes of initially isogenic competing populations. Some experimental evolutions have estimated relative population sizes by looking for metabolic markers visible when the competing populations have been grown on solid media (Lenski et al. 1991). However fluorescent proteins enable each population to be monitored throughout the competition enabling greater phenotypic characterisation.

If two genes are fixed in a significant number of experiments then new questions can be addressed. Is the fitness effect of the second mutation smaller than the first as described by Schoustra in *Aspergillus nidulans* (Schoustra et al. 2009)? Alternatively, does the second mutation mitigate pleiotropic fitness costs of the first mutation? Such a cost could be assessed by performing fitness competitions without added ethanol. Such controls also establish whether genetic changes are an effect of ethanol or adaptive in that genetic background regardless of ethanol stress.

Accordingly, the experimental methods chosen in these experiments had several characteristics. They used batch culture on 96 well plates enabling 48 strains evolving in parallel (see section 3.5 for further details) for up to ~200 generations

which is likely to fix one or possibly two genes, creating a library of initial steps rather than an adaptive walk.

2.4.6. Size and ploidy change in experimental evolution

Some experimental evolution experiments have produced strains with a ploidy that differs from the starting strains. An experimental evolution with an isogenic ploidy series of starting strains showed a general convergence towards diploidy accelerated by salt stress; haploid starting strains become diploid, diploids remained unchanged, and without saline stress tetraploids became diploid (sometimes with triploid intermediates). The exception to this pattern was salt stressed tetraploids which became stably triploid (Gerstein et al. 2006). Gerstein later concluded that *S. cerevisiae* is able to change ploidy whilst remaining euploid (Gerstein et al. 2008). Subsequent experimental evolutions with haploid *S. cerevisiae* under saline stress showed diploids arising and sweeping through the populations despite having no measurable competitive advantage over haploids in assays (Dhar et al. 2011; Gerstein & Otto 2011) and other work by some of the same authors showing that haploids adapt faster than diploids when subjected to a range of stressors (Gerstein et al. 2011).

Adams and Hansche found that a haploid *S. cerevisiae* strain that outcompetes a diploid when one nutrient is scarce may have parity of competitive fitness when all nutrients are in excess (Adams & Hansche 1974). This may be because the smaller size of haploid cells gives them a larger surface area to volume ratio and hence more efficient nutrient uptake with scarcity. In a 260-generation, glucose-limited, continuous culture experiment with a diploid starting strain, cell sizes consistently shrank (Adams et al. 1985). Such results should be interpreted with caution since small cell size can be an indicator of poor growth rather than adaptation. Adams et al did not consider ploidy reduction as a possible cause of cell size shrinkage. An isogenic ploidy series from haploid to triploid of *S. cerevisiae* (strain Σ1278b) showed that the initial uptake rate of arginine, lysine, or uridine taken up per cell mass declined with increasing ploidy (Hennaut et al. 1970). Hennault et al speculated that this was due to limitations of permease transport.

Cell size changes in experimental evolution are not just confined to yeast. Cell volume approximately doubled in 8,000 generations in all twelve lines in Lenski's long-term experimental evolution of *E. coli,* decreasing the surface area to volume ratio by 27% (Lenski & Mongold 2000). As in *S. cerevisiae*, cell size in *E.coli* is positively correlated with growth rate but evolved cells in Lenski remain larger than

their ancestors when growth rate constrained in a chemostat (see figure 4b, p41) (Lenski et al. 1998). However, it is unclear in *S.cerevisiae* or *E.coli* whether cell size changes are adaptive or only side-effects of genetic change. In both organisms, cell size is a complex phenotype that arises out of the interaction between growth and division. Conclusively showing that the optimal cell size varied with particular environmental conditions would require the capacity to systematically change cell size and measure biomass.

2.4.7. Drawing conclusions from experimental evolution

A recent paper by Blank et al. states that "Determining the molecular changes that give rise to functional innovations is a major unresolved problem in biology" and "The paucity of examples has served as a significant hindrance in furthering our understanding of this process (Blank et al. 2014). However, most mutations are neutral or deleterious and competing theories of adaptation still contend (Nei 2005) . Distinguishing adaptive mutations from neutral or deleterious is non-trivial and highly dependent on environmental conditions; fitness determinations are strictly only valid for the environment in which the determination takes place. When *S. cerevisiae* is cultured in the laboratory on rich media there are a large number of 'non-essential genes' that can be deleted without destroying cell viability and some such deletions may leave the competitive fitness of mutants apparently undiminished.

For well-characterised structural genes it is possible to speculate on the effect on function of truncations or non-conservative amino acid changes to protein sequences by non-synonymous nucleotide polymorphisms, insertions, or deletions. It is easier to predict that a protein will be unable to function (e.g. because a critical domain is deleted or lacking key residue) than that its function might be augmented. Some null mutations may be adaptive, e.g. the lack of the uracil biosynthetic enzyme *ura3*Δ has been reported to increase ethanol tolerance (Swinnen et al. 2012) although the inability to synthesise uracil may be antagonistically pleiotropic (see Introduction section 2.6).

It is not possible to guarantee, certainly with current whole genome sequencing methodology, that a strain isolated from an experimental evolution has only one mutation, other mutations may be present that have not been detected (for a discussion of this, see Section 6.1). For synonymous structural coding changes (which alter the tRNA required to construct the protein) or changes to regulatory sequences such as promoters or non-coding RNAs it is even less straightforward to draw conclusions solely from sequence alterations. Accordingly robust determination

of the fitness impact of a mutation requires the introduction of the mutation into the ancestral background.

Because causality is difficult to demonstrate and individual mutations must arise from random and rare events then common changes in parallel evolutions are required to conclude with confidence that particular (recurrent) genetic changes are an adaptive response to a particular environment or stress. Such parallelism may readily be found in some experimental evolutions; e.g. mutations in the *rpoB* (RNA polymerase beta subunit) are found when bacteria evolve resistance to the antibiotic rifampicin, preventing the binding of the antibiotic to the polymerase (Feklistov et al. 2008). However, ethanol damages multiple cellular targets including compromising membrane integrity, denaturing proteins, and mutating DNA (see Introduction section 2.1.3); suggesting that there may be many effective genetic targets for adaptive evolution.

Adequately determining ecosystem selection pressures is problematic. The key research aim is to isolate and characterise the phenotypes and genotypes of *S. cerevisiae* strains with enhanced ethanol tolerance using experimental evolution (see Section 2.7 “Research aims and goals”). However, the competitive fitness in experimental evolution is only a proxy for ecological fitness if the microbial strain and environment are close to that found in the ecosystem, or at least comparable to a desirable industrial use. While relatively constant and experimental conditions are a desirable experimental feature of continuous culture, batch culture is much closer to the current industrial uses of S. cerevisiae in brewing or bioethanol production

Continuous culture may also make interpretation of genotypic changes more difficult since there may be accumulation of hitchhiking mutations in regions of genome subject to relaxed selection pressure – the fitness effects of single mutations in otherwise isogenic backgrounds needs to be assessed completely to exclude this. Batch selection that subjects microbes to feast then famine might be expected to more closely resemble ecosystem selection pressures.

The generally poor fermentation performance of laboratory strains of yeast suggests they may initially be poorly adapted. This may be due to relaxed selection pressure for ethanol tolerance in laboratory strains; whilst a 15% v/v ethanol content of wine is common, Marks et al. claim that, “in standard laboratory media the ethanol concentration never exceeds 1% (v/v)” (Marks et al. 2008). This is not a barrier to evolving improved ethanol tolerance, but laboratory strains are also derived largely from S288C, a narrow genetic base which has led to the claim that such strains are

“recessive mutants at multiple loci” (Kron 1997) and it may therefore not be possible to develop very high levels of ethanol tolerance with laboratory strains. Flocculation confers ethanol stress resistance and is promoted by increasing ethanol concentration (Smukalla et al. 2008). Breeding effort was expended to reduce the flocculation of S288, the ancestor of most laboratory strains, to make it more suitable for laboratory use (Mortimer & Johnston 1986).

Auxotrophies may also have unintended effects, changing the adaptive genetic changes made and hence the perception of gene function. Acid growth inhibition is increased by deletion of the genes for the ATP-binding cassette *PDR12* (P. Piper et al. 1998), and *AZR1* (Tenreiro et al. 2000) plasma membrane transporter genes. The transcription factor *War1p*, dramatically increases *PDR12* expression in response to acid stress (Kren et al. 2003) which pumps out carboxylate anions (Piper et al. 2001). However, according to Bauer et al, deletions of *PDR12* and *AZR1* only increase acid sensitivity in media deficient in tryptophan (Bauer et al. 2003), which are more sensitive to acid stress than the wild-type, a sensitivity that can be overcome by tryptophan enrichment of the medium, or (for sorbate stress) by over-expressing the *TAT2* tryptophan permease (Bauer et al. 2003). Tryptophan import appears to be impaired by acid stress causing poor growth in strains unable to synthesise tryptophan; such mutants are not in the well-adapted tail of the acid-stress fitness distribution.

A wide range of brewing strains with proven (and varying) ethanol tolerance could be studied. However, there are problems with using such strains in the laboratory; many have a strong flocculation response which clarifies brewed beer or wine but makes assessment of growth by optical density harder. Adhesion to well surface could create heterogeneity in the population in which sub-populations or demes increase clonal interference (Campos et al. 2008). It is possible that growth with added galactose (El-Behhari et al. 2000) or mannose (Smukalla et al. 2008) might inhibit flocculation enabling study of these strains. However, such industrial strains are frequently aneuploid and hybrid, which would also make genetic analysis of evolved strains more difficult.

However, experimental evolution may continuously or repeatedly subject microbes to environments that differ from those in which they evolved. Prolonged chemostat cultivation, in which the medium is constantly diluted, subjects microbes to a constant nutrient shortage in which the rate of substrate assimilation is limiting the rate of growth. Such continuous culture favours increases in the efficiency of absorption

and utilisation of nutrients. Whilst starvation is likely to be a common microbial state, the *S. cerevisiae* strains studied appear to have adaptations to episodic plenty rather than consistent paucity; a strong Crabtree effect with a high flux of inefficient fermentation rather than efficient respiration. *S. cerevisiae* lacks the high affinity active glucose transporters found in other yeasts (Van Urk et al. 1989). Although *S. cerevisiae* may also fill other niches, this is consistent with the rotting fruit ecological niche.

Accordingly continuous culture may adapt *S. cerevisiae* in radical ways that may conflict with its evolutionary history. Evolution of *S. cerevisiae* in glucose-limited chemostat culture produces strains with substantially reduced maximum rates of glycolysis (Jansen et al. 2005). Similarly chemostat culture for more than 25 generations with limited maltose as the sole carbon source, produces *S. cerevisiae* strains with raised maltose active transport capacity and affinity which are killed by high concentrations of maltose that are safe for the starting strains(Jansen et al. 2004). Both these examples suggest that low metabolic flux cytometer selection may impair the performance of *S. cerevisiae* under conditions of high flux.

2.5 Yeast growth with ethanol stress

2.5.1. Phases of yeast growth

Yeast growth in typical laboratory culture can be considered in different stages. There will be a lag before cells transferred from exhausted media into fresh media will commence growth. This lag is followed by a phase of logarithmic growth with high fermentation rates. When the glucose concentration drops below 0.2% the diauxic shift occurs; the rate of fermentation declines and cells stop dividing for a few hours. The higher affinity sugar transporters *HXT6* and *HXT7* are mainly expressed post-diauxy (Santos et al. 2008). After the diauxic shift, cells enter the post-diauxic or slow-growth phase which, depending on genetic background and environment, may last a week in which cells may undergo one to three doublings while they consume ethanol or other non-fermentable carbon sources (Westerbeek-Marres et al. 1988; Fuge et al. 1994; Guidi et al. 2010). After approximately seven days, depending on the strain, the post-diauxic phase is followed by the stationary phase, characterised by a cessation of doubling (Gray et al. 2004).

2.5.2. Ethanol and growth

Each phase of growth is affected by ethanol stress. The duration of the lag phase consistently increases under ethanol stress (Medawar et al. 2003). Similarly ethanol

stress is one cause of logarithmic growth ceasing despite sugar remaining in the growth medium; in brewing this is called a 'stuck fermentation' (Blateyron & Sablayrolles 2001). If two strains are competing in a typical laboratory culture medium in which there are abundant substrates for growth (sugars, nitrogen) but no ethanol there is a selection pressure for early emergence from slow growth/quiescence. The strain which begins to grow first gains an advantage. Davey et al. (Davey et al. 2012) modelled competitions between strains in batch fermentation whose growth differed solely in their lag durations, one with a 15 minute lag, the other with 100 minutes; they found that the strain with the shorter lag rapidly outcompeted the tardier strain and would have constituted more than 90% of the population within 9 rounds of growth. A priori, interaction between strains would be expected to be minimal during the lag phase since the cell count is small. However, it has been shown that yeast cultures in the lag phase, subject to ethanol stress, secrete acetaldehyde into the growth medium and this can shorten the lag phase (Walker-Caprioglio & Parks 1987). Accordingly the emergence from lag and growth of one yeast strain may stimulate the growth of a competitor, minimising the difference in lag emergence.

Cells that are transferred into a rich medium with added ethanol have conflicting benefits and stressors. Stationary phase cells have a higher ethanol tolerance than log phase cells (Santos et al. 2008) so emergence into active fermentative growth has risks as well as benefits, fresh resources offer an opportunity for a number of generations of growth, but ethanol toxicity may prevent the utilisation of that resource.

Since most growth occurs during the log phase, the duration of log phase and the rates of growth during it will largely determine competitive advantage. Delayed exit from lag or a premature diauxic shift will impair fitness in competition. An ethanol stress of 12%(v/v) reduces log phase glucose transport by 80% (Santos et al. 2008) and it has been suggested that this might cause a stuck fermentation by initiating a premature diauxic shift due to depletion of intracellular glucose (Leao & Van Uden 1982).

2.5.3. Ethanol and quiescence

The median life span of most laboratory *Saccharomyces cerevisiae* strains is about 25-35 generations (~3 days) forming 2-50 buds (Jazwinski 1990). Percoll density gradient centrifugation of yeast cultures in the post-diauxic phase show two fractions which form after 36 hours of growth (Allen et al. 2006). The upper fraction of lower

density cells vary greatly in size. There are mother cells with bud scars and a significant number of cells in bud, they lack storage carbohydrates, have a low respiration rate and are still fermenting. The dense lower fraction is composed of unbudded daughter cells with a low variance in cell size (4.6+/-0.33µm), large amounts of storage carbohydrates (glycogen, trehalose), and a good respiration rate (Benbadis et al. 2009). Quiescence arises in the dense fraction of respiratory competent cells. Accordingly the dense fraction of daughter cells undergoes two phenotypic changes before entering quiescence; they switch from respiration to fermentation and accumulate storage reserves rather than budding.

The reproductive capacity of non-quiescent cells declines faster than quiescent cells – 50% of non-quiescent cells are apoptotic or necrotic and 90% have lost the capacity to reproduce after 14 days and have accumulated substantial reactive oxygen damage (Allen et al. 2006). The quiescent cells are probably in a state akin to G_0 because they synchronously enter the cell cycle when their growth restarts (Allen et al. 2006). Recent work suggests that sake yeast, capable of very high ethanol fermentations, have a defect in establishing quiescence and fermentation continues after growth ceases, unlike normal fermentations in which fermentation is linked to cell division, a possible by-product of strain selection for high ethanol brewing (Urbanczyk et al. 2011).

Cells in the dormant or quiescent state have a low metabolism and do not divide but have a greater tolerance of environmental stress. Modelling studies (Malik & Smith 2008) suggest that strains capable of dormancy have a higher fitness when favourable growth conditions are rare. However, a laboratory environment that predictably provides favourable growth conditions within a short period of the onset of diauxy is unlikely to favour quiescence.

More than 50 mitochondrial mutants were unable to form a high density band of quiescent cells (Martinez & Roy 2004) suggesting that mitochondrial function is essential for entry into dormancy. Many non-quiescent cells that can still reproduce lose mitochondrial function and become petites (Aragon & Rodriguez 2008), suggesting mitochondrial DNA damage has occurred. It has also been suggested that cell death under high levels of ethanol stress has features such as mitochondrial fission akin to apoptosis (Kitagaki et al. 2007).

2.5.4. Measures of growth

There are several measures of yeast growth and each has strengths and limitations. The optical density of the growth medium at 600nm (OD_{600}) and the fluorescence of endogenously synthesised fluorescent proteins are both population growth indicators, and cells counts and sizes can determined by light cytometry.

Whilst there is a clear positive correlation between yeast growth and OD_{600}, the extent to which cultures of yeast in suspension absorb or scatter light is more complex than that of homogeneous fluids. Optical density can fall with cell lysis (Jones et al. 2001) and might also be affected by cell size and shape (which is in turn affected by ploidy and the fraction of cells in bud) and by the adhesion of cells to the plate.

The ancestral strains used in the experimental evolutions described in this thesis (DeLuna et al. 2008) were tagged with cyan fluorescent protein (CFP) or yellow fluorescent protein (YFP). The fluorescence from these proteins is a proxy for protein synthesis and growth. Their expression is under the control of the constitutive glyceraldehyde-3-phosphate dehydrogenase (*TDH3*) promoter which is identical for each fluorophore (DeLuna et al. 2008) and presumed to be expressed as a consistent proportion of total growth as ploidy changes; genes that are differentially regulated with ploidy tend to encode proteins localised to the cell surface reflecting changes to the surface area to volume ratio as cells enlarge (Wu et al. 2010). It has been claimed that measurement of fluorescence can detect differences in fitness as small as 1% (DeLuna et al. 2008).

These fluorescent proteins are derived from green fluorescent protein (*GFP*). These fluorophores derived from *GFP* are surprisingly durable, they have a soluble β-barrel that encapsulates and protects the modified hexapeptide chromophore (Ormo et al. 1996). *GFP* resists denaturation, remaining stable in pH 5.5-12, temperatures up to 70°c, 6M guanidine HCL, 8M urea or 1% sodium dodecylsulphate up to 45°C (Bokman & Ward 1981). Fluorescence is unaffected by two day incubations in trypsin, chymotrypsin, papain, subtilisin, thermolysin or pancreatin at protease concentrations up to 1mg/ml *GFP* (Bokman & Ward 1981). This durability is an asset since fluorescence is unlikely to be affected by protein denaturation in living cells. There are some indications that cell stress and lysis may provoke oxidant fluorophore damage (Greenbaum et al. 2000) but the exclusion of bulk solvent from the region of the chromophore gives *GFP* and its derivatives substantial protection from oxidation damage and in general the remains of dead cells will continue to

fluoresce. Green fluorescent protein variants such as CFP and YFP are bleached at half the rate of fluorescein (Swaminathan et al. 1997).

YFP is a variant of *GFP* in which a key substitution of tyrosine for threonine at position 203 not only shifts the excitation and emission wavelengths, but also makes YFP more susceptible to quenching by increases in acidity (Elsliger et al. 1999), halide ions and nitrate (Wachter et al. 1998). When chloride ions increase from 0 to 150mM, the fluorescence decreases by 40% at pH 7.0 and by 60% at pH 6.4 (Wachter & Remington 1999). It is unclear whether there is a significant change in halide ions or nitrate during fermentation but ethanol acidifies the cytoplasm by a well-documented increase in cytoplasmic membrane permeability (see section 2.1.3). The extent of the effect of pH on YFP in this experiment is unclear; there is no significant reduction in the YFP signal when strains with this fluorophore are grown with 10% ethanol added to the growth media (see section 5.3). However, a fall in the YFP fluorescence is seen after the diauxic shift (see Figure 2.5) that is absent from any of the CFP fluorescence plots. This YFP fall may be due to a drop in cytosolic pH as metabolic rate falls at diauxy.

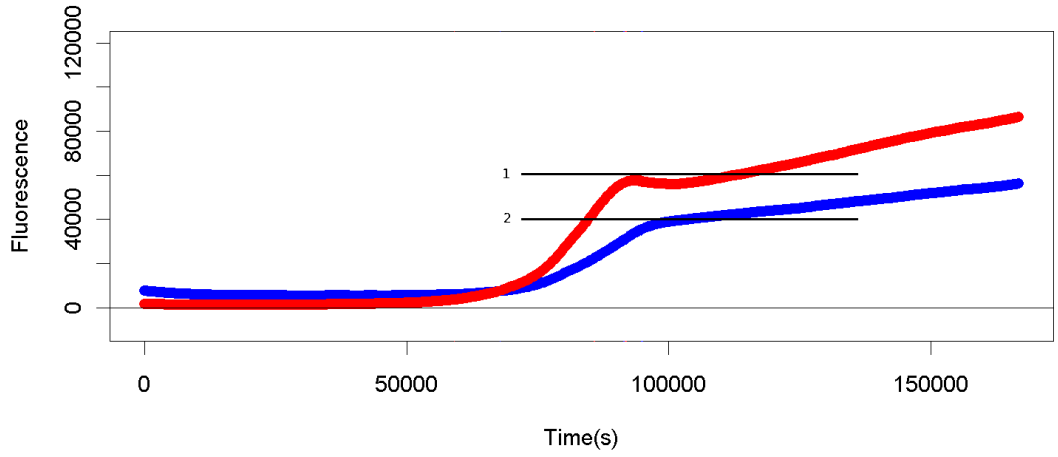


Figure 2.5 CFP and YFP fluorescence during a competition between ancestral CFP and evolved YFP strain Evo10_20_YFPe_G10_2 in 6% ethanol.
The CFP fluorescence is blue and the YFP fluorescence is red. The horizontal black lines in the figure show the fluorescence close to diauxy for YFP (line 1) and CFP (line 2). The value of YFP fluorescence is lower at $1x10^5$s than at diauxy with fluorescence falling below line 1. However, while the rate of increase of fluorescence falls for the CFP measurement, CFP fluorescence does not fall below line 2. The fall in YFP fluorescence may be due to lowered cytosolic pH when metabolic rate falls at diauxy.

Fluorescent proteins related to *GFP* not only require oxygen in order to develop fluorescence (Pouwels et al. 2008) but YFP needs to be activated in two steps whereas CFP has only one; accordingly YFP has twice the oxygen requirement of CFP (Remington 2006). This raises two confounding possibilities. Firstly, if oxygen becomes depleted in the medium then the fluorophores may not complete their

maturation (Zhang et al. 2006) and the fluorescence will underestimate growth, particularly at high cell density. Secondly any such shortage may under-report growth with the YFP reporter more than the CFP reporter. This possibility cannot be wholly discounted. However, lids on the plates were not sealed to facilitate gaseous exchange and the 120µl culture volumes used were both small, with short diffusion paths, and intensively shaken; every five minutes the plates were shaken for a minute with a 200 rpm double orbital pattern prior to readings being taken (see Section 3.3).

Whilst the intensive shaking that accompanied the sampling routines may minimise the risk of failure to develop fluorophore fluorescence it does increase the risk of another potential confounding factor. *GFP*-derived fluorophores produce toxic reactive oxygen species including singlet oxygen $^1O_2$ when they photobleach (Greenbaum et al. 2000). Such toxicity is exploited in the fluorescent protein, 'KillerRed', a photosensitiser which destroys illuminated cells (Dixit et al. 2006). Whilst reactive oxygen toxicity from this source is known to have a significantly adverse effect on *Caenorhabditis elegans* (Remington 2006) it is unclear how damaging such photo-bleaching is in *S. cerevisiae*. The intensive shaking and sampling regime in these studies may have unintentionally caused a reactive oxygen species stress in these studies.

Staining procedures may be used in both light and fluorescence cytometry to count the cells in a population that have damaged membrane integrity and are likely to be dead. Counting dilutions of liquid cultures spread on agar plates provides a measure of the cells in a population that can be cultured, although there have been some claims that the difference between microbial 'culturability' and viability may be substantial (Kell et al. 1998). Optical density does not give a measure of the dead cells in a population.

2.6 Experimental evolution starting strains

The extent to which results from experimental evolution can be generalised to draw conclusions about evolution of a characteristic in a species depends on how typical the starting strains are. Conclusions drawn from laboratory strains which may have diverged from wild type may not apply to wild strains (for further discussion of this see, this chapter, section 2.4.7). However, the parental strains used for experimental evolution in this work had also been engineered with multiple deletions and gene fusions to enable selective culture prior to acquiring their fluorophore genes. The Kishony lab placed the fluorescent protein genes for YFP (yEVenus) or CFP (yECerulean) at the *hoΔ* locus of the Y7092 synthetic genetic array (SGA) strain which is in turn derived from the *MATα* strain BY4742 (Tong et al. 2001; Hin et al. 2007). The fluorophore expression is driven by the promoter for the constitutive glycolytic enzyme Glyceraldehyde-3-phosphate dehydrogenase (*TDH3*) (DeLuna et al. 2008).

The deleted genes in strain Y7092 were required for synthesis or absorption of certain amino acids and the synthesis of pyrimidine nucleobases:

1) *met15Δ*, Methionine and cysteine synthase (O-acetyl homoserine-O-acetyl serine sulfhydrylase), required for sulfur amino acid synthesis
2) *leu2Δ* Beta-isopropylmalate dehydrogenase (*IMDH*), catalyzes the third step in the leucine biosynthesis pathway
3) *can1Δ* plasma membrane permease for absorption of arginine. The *CAN1* gene is replaced by the *Ste2pr::SP_HIS5* fusion
4) *lyp1Δ* (YNL268W) plasma membrane permease for absorption of lysine.
5) *his3Δ*. The *HIS3* gene catalyses the seventh step in histidine biosynthesis. However, the function of this gene in strain Y7092 is conditionally complemented by a fusion between *STE2*, which codes for the alpha factor mating receptor fused to the *S. pombe HIS5* gene (which encodes an enzyme catalysing the same reaction as *S. cerevisiae HIS3*). The *Ste2pr::SP_HIS5* fusion is inserted at the locus of the deleted arginine permease (*CAN1*). The *S. pombe* gene is used because it cannot restore wild-type *HIS3* function by gene conversion with small deletions (Daniel et al. 2006) and extensive deletions of *HIS3* were found to additionally impair the adjacent *MRM1*/*PET56* gene leading to a higher rate of petite formation (Brachmann et al. 1998).

6) *ura3Δ* encodes the enzyme orotidine-5'-phosphate (OMP) decarboxylase, which catalyses the sixth enzymatic step in the de novo biosynthesis of pyrimidines. This deletion can be used as a selectable marker since it makes the cell resistant to 5-Fluoroorotic acid (5-FOA) which is lethally converted into the toxic 5-fluorouracil in wild-type cells.

Deleting endogenous synthesis of cysteine, methionine, leucine, histidine and pyrimidines makes them essential nutrients that must be absorbed from the medium. The fact that SGA strains such as Y7092 grow well on rich media such as yeast extract peptone dextrose (YPD), suggests that the amino acid and pyrimidine pools in these strains are not greatly compromised by these deletions. However, there are a number of reports not just of impaired active transport but leakage of amino acids from the cell under ethanol stress (see Introduction, section 2.1.3). Accordingly, the evolutionary landscape of Y7092 under ethanol stress may differ from the wild type due to shortage of amino acids or pyrimidines.

Davey et al. found that a heterozygous *lyp1Δ* deletant strain had decreased viability during seven days of starvation, suggesting haploinsufficiency at this locus in (nitrogen limited) grape juice medium, but this effect was not detected during a batch culture regime with 48 hours in stationary phase (Davey et al. 2012). This study also identified two arginine biosynthesis mutations, *Arg2* (YJL071w) and *arg5-6* (YER069w), as haploinsufficient in the same medium in batch culture despite a wild-type *CAN1* permease background and arginine being the most abundant amino acid in grape juice (Austin & Butzke 2000).

Although it appears that there are no reports in *S. cerevisiae*, in *E.coli* absorption of lysine and arginine (as well as glutamate) from the growth medium is important for acidic tolerance (Lin et al. 1996; Diez-Gonzalez & Karaibrahimoglu 2004). The *E.coli* arginine and lysine decarboxylases are inducible by low pH (Gale 1946; Auger et al. 1989) and it is thought that decarboxylation of imported lysine and arginine by these enzymes raises the pH of the cytoplasm. Mechanisms which oppose acidification of the cytoplasm under ethanol stress are thought to be a component of tolerance (see this chapter, section 2.1.3).

There are conflicting reports of the effect of *ura3Δ* deletion on fitness under ethanol stress in YPD medium. On one hand it has been suggested that *ura3Δ* deletion increases ethanol tolerance (Swinnen et al. 2012) but uracil can also be a limiting nutrient in YPD medium; in YPD with 25% glucose; the cell density of the *ura3Δ* strain is half that of the wild type prototroph (Chopra et al. 1999). Whilst carbon in

YPD with 2% glucose may become limiting long before uracil, absorption of uracil from the medium may not always be possible. Uracil permease is rapidly degraded as cells approach stationary phase or if they are starved of carbon, nitrogen or phosphate (Volland et al. 1994).

The Y7092 strain has the mating type *MATα*. Putting the *HIS3* complementing gene fusion under the alpha factor promoter in the *Ste2pr::SP_HIS5* fusion means that *MATα* are unable to synthesise histidine (his-) because they repress the alpha factor receptor *STE2*. Accordingly *MATα* Y7092 cells have very poor fitness on synthetic defined medium without histidine (Hin et al. 2007). However, if *MATα* Y7092 cells were able to switch mating type to *MATa*, the alpha factor receptor would no longer be repressed, the fusion protein would be expressed and they would become able to synthesise histidine and gain a likely growth advantage on media deficient in histidine.

The selective advantage of mating type switching may be significant even on media containing histidine. When a *HIS3* cassette was introduced into the haploid *his3Δ* mutant strain BMA41-1B and competed with its parent in a continuous culture medium supplemented with amino acids including histidine, the his+ strain increased from 50% to approximately 80% within 10 generations, an average growth rate advantage of 23% (Baganz et al. 1997).

The rate of mating-type switching in *hoΔ* heterothallic yeast, is estimated to be between 10^{-6}-10^{-7} (Klein & Wintersberger 1988). This is thought to be due to a DNA break in the *MAT* locus that is repaired using intra-chromosomal gene conversion with the silent locus of the opposite mating type. Accordingly, DNA damage due to X-rays (Schiestl & Wintersberger 1982), UV light, or chemical carcinogens (Schiestl & Wintersberger 1983) increases the rate of mating-type switching. It is possible that ethanol stress could increase the rate of mating type switching, but although the primary ethanol metabolite acetaldehyde has been shown to induce single strand nuclear DNA breaks this has only been seen in repair deficient yeast strains (Ristow et al. 1995), (see this chapter, section 2.1.3).

The compound 3-amino-triazole targets the product of the *HIS3* locus (imidazoleglycerol-phosphate dehydratase). Yeast heterozygous at the *HIS3* locus grow less well than the homozygous wild type, but only when 3-amino-triazole is present, i.e. the drug stress induces haploinsufficiency at this locus and shortage of histidine synthesis despite growth in synthetic complete media containing histidine

(Shoemaker et al. 1996; Giaever et al. 1999). Several other examples of drug-stress induced haploinsufficiency are known (Giaever et al. 1999).

Aside from any fitness benefits of gaining histidine biosynthetic competence, any mating-type switchers would be able to mate with the rest of the *MATα* population thereby increasing their ploidy from haploid to diploid and this may confer benefits (see this chapter, section 2.2.2).

2.7 Research aims and goals

The core research aim was to isolate and characterise the phenotypes and genotypes of *Saccharomyces cerevisiae* strains with enhanced ethanol tolerance using experimental evolution. In order to achieve this aim, it was necessary to devise and carry out experimental evolutions with effective selection pressures and batch protocols; in which adaptive evolution might ideally occur within weeks and months rather than the many years of the long-term experiment with *E. coli* (Barrick et al. 2009). Further, it was necessary to identify those experimental evolutions in which adaptive evolution had occurred under ethanol stress amongst dozens run in parallel, and to isolate strains with enhanced competitive fitness from them (see Chapter 4). These isolated strains needed to be phenotypically characterised using ancestral competitions to identify the characteristics associated with enhanced ethanol tolerance, such as shortened lag, enhanced duration of growth, greater maximum growth rate or greater overall growth (see Chapter 5). In turn, strains were genotyped using next generation sequencing with the goal of identifying the initial mutations conferring a fitness benefit under ethanol stress and thereby give an estimate of the extent to which common patterns of genetic variation might arise in experimental evolutions subjected to ethanol stress (see Chapter 6). Such common patterns could then be used as an indicator of the diversity of accessible adaptive targets under an ethanol selection pressure, i.e. a measure of the ruggedness of the adaptive landscape, and provide genetic targets for further investigation to demonstrate genetic changes responsible for the changed phenotype.

2.8 Structure of thesis

The thesis structure follows the sequence of research aims and goals. This introduction (chapter 2) has covered the growth and nature of *S. cerevisiae*, its ethanol tolerance and methods of experimental evolution and this is followed by a discussion of the common experimental methods (chapter 3). In chapter 4, the experimental evolution protocols are described together with the Dirac curve fitting

and population genetic modelling used to characterise experimentally evolving populations. Chapter 5 describes the phenotypic characterisation of isolated strains, growth patterns, including the verification that they could out-compete their ancestors and some could gain organic acid tolerance. Chapter 6 concerns the genotypic characterisation of isolated strains by whole genome sequencing as well as the ploidy determination of strains and populations using fluorescence cytometry of cells with the Sytox Green stain bound to its DNA. Chapter 7 is a general discussion.

3. General Methods

This chapter contains the methods that were in general use for these experimental evolution experiments. These methods include how plate readers were used to make measurements of fluorescence and optical density in dozens of parallel experimental evolutions to measure growth and look for changing ratios of fluorescence from the two fluorophores as indicators of sweeps in the competing populations driven by adaptive evolution (this is discussed in detail in Chapter 4). These general methods also include the protocols followed for counting cells and determining their size using a Nexcelom Cellometer

3.1 Software versions

A PC version of Perl, version 5.12.3, was used for data processing and database storage. R, version 2.15.2 unless otherwise stated in the text, was used for data modelling and producing figures. A PC edition of MySQL v5.5 was used for plate reader data storage. Some Bash (v 3.2) scripting was run on Mac computers for analysis of sequence data.

3.2 Sizing and counting cells with the Nexcelom Cellometer

Cell suspensions were counted using the Nexcelom Auto T4 Cellometer (Nexcelom Bioscience, 360 Merrimack Street, Building 9,Lawrence, MA 01843, USA) which counts cells in a 20µl suspension drawn into disposable slides by capillarity. The cytometer takes micrographs of the suspension and uses image recognition to identify and count the cells and then determines the mean of the largest diameter and the diameter at 90° to it.

For counting and measuring a stationary phase culture in YPD with 2% glucose an initial dilution of ~100x is required to produce images that contain hundreds of cells but which nevertheless lack dense overlapping cell clusters.

Overnight cultures were vigorously vortexed before cell counting. However, there was found to be high risk of cross contamination when 96-well plates were vortexed to suspend cell cultures. Aggregation and biofilm formation can be a problem in both continuous and batch experimental evolution since it may lead to selection pressures varying throughout the experimental environment. In batch culture, cells that adhere more strongly to the well sides may also be selected against if they are not suspended prior to batch transfer. Enhanced adhesion of *S.cerevisiae* cells to each

other (flocculation) and to the apparatus occurs normally when cells are in the slow-growth phase in dense cultures, despite the early selection for minimal flocculation when yeast strains were first introduced into the lab (Mortimer & Johnston 1986). Accordingly, great care was taken that the populations in 96-well plates were suspended by stirring and repeatedly (30 times) drawing up the culture with a pipette. Particular attention was paid to the middle and bottom right-angled corners of the wells with the contents drawn up to ensure that the visible adherent flocs that collect there were properly suspended. Confidence that the repeated drawing of the suspension through the pipette tip was successful in breaking up flocs was reinforced by the cell images viewed in the Nexcelom Cellometer, which showed none of the cell aggregates which can readily be detected in inadequately broken up dense cultures. The image in Figure 3.1 is typical, showing no large aggregates of cells although groups of two cells, one larger and one smaller, likely cells with buds, are common even in slow growth cultures that are post-diauxy. These may be separate daughter cells that have remained adherent to the mother cell. The Nexcelom cell counting software has a mode showing the parts of the image recognised as cells, showing that buds are recognised and measured as independent cells, something that plating and colony counting would be unable to do. Sonication to separate mother and mature but adherent daughter cells was considered but might have caused cell damage, particularly to cells with immature buds.

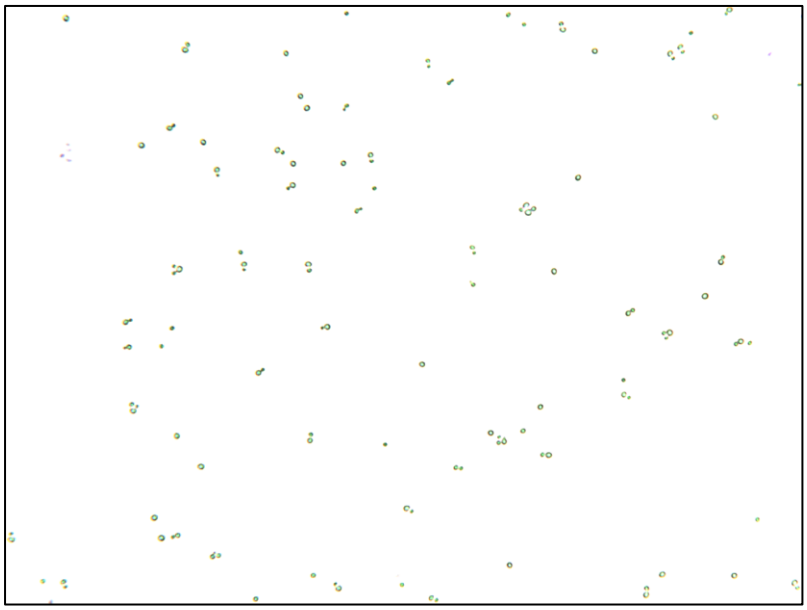


Figure 3.1 Nexcelom Cellometer micrograph of 1/100 dilution of a slow growth phase ancestral CFP culture grown in YPD with no added ethanol.
Figure shows few clumps of cells but many cells in pairs of unequal size, likely to be mother cells with buds/daughter cells.

3.3 Reading 96-well plates: OD_{600} and fluorescence

For experimental evolution and strain analysis the 96-well plates used were Greiner sterile black fluorescence microplates (catalogue number 655077) with sterile lids (Greiner Greiner Bio-One Ltd., Brunel Way,Stroudwater Business Park,GL10 3 SX,Stonehouse). The plate choice was constrained by the availability of sterile plates suitable for fluorescence measurement, but these plates may not be optimal for experimental evolution with free circulation of cells in suspension since they are made with a polystyrene resin which is claimed to have "enhanced protein binding properties" (Product description: Greiner 96 "F" well Polystyrene microplate, 2013 http://www.greinerbioone.com/en/england/articles/catalogue/article/68_4/16664/) which may have encouraged cell adhesion.

Plates were read in one of three BMG plate readers (BMG Labtech GmbH, Allmendgruen 8, D-77799 Ortenberg/Germany): a PolarStar Omega or a FluoStar Omega plate reader.

The minimum practical incubation temperature for the plate readers used is 1-2°C above ambient. Each reader was able to maintain an incubation temperature of 30°C for all experimental evolutions and strain evaluations. OD_{600} and fluorescence were measured at five minute intervals with a minute of 200rpm double orbital shaking prior to reading. Each machine was calibrated to show 90% of its maximum fluorescence value when measuring a dense stationary culture of cells of the appropriate strain. The data was output in the form of text files for each OD or fluorescence reading. These were merged and the data parsed and loaded into a MySQL database using custom Perl scripts.

3.4 Cell growth calibration curves

3.4.1. Optical density calibration

A calibration curve of OD_{600} against cell count was used to calculate bottleneck sizes for evolutions 8 and 10. The calibration curve was produced from saturated overnight cultures of the haploid CFP and YFP strains grown in YPD medium. The serial dilution was carried out in a black fluorescence plate from the same batch used in the evolution and read on the same plate reader (the BMG Omega for evolution 8 and the BMG Fluostar for evolution 10). The layout of the serial dilution for evolution 8 is shown in Figure 3.2. For evolution 8, dilution series were carried out for each ethanol treatment (see Figure 3.2).

120µl of culture medium was placed in all but the first well in the dilution series and the first well received 255µl of vortexed saturated culture. The culture was mixed by taking up repeatedly into a pipette and then 120µl was placed in the next well in the series. Once all the dilution transfers had been made, 135µl was discarded from the final well and 15µl was taken from the first (undiluted) well in the series for x100 dilution and cell counting using the standard method (see this chapter section 3.2) so that all wells had the standard final volume of 120µl

	1	2	3	4	5	6	7	8	9	10	11	12
A	CFP YPD 255µl	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Medium (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Medium (YPD)	120µl Media (YPD)
B	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Medium (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Medium (YPD) Discard 135µl	120µl Media (YPD) BLANK
C	YFP YPD 255µl	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Medium (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Medium (YPD)	120µl Media (YPD)
D	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Medium (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Media (YPD)	120µl Medium (YPD) Discard 135µl	120µl Media (YPD) BLANK
E	CFP YPD + 4.5% EtoH 255µl	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH Discard 135µl	120µl YPDE 4.5% EtoH BLANK
F	CFP YPD + 6.5% EtoH 255µl	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH Discard 135µl	120µl YPDE 6.5% EtoH BLANK
G	YFP YPD + 4.5% EtoH 255µl	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH	120µl YPDE 4.5% EtoH Discard 135µl	120µl YPDE 4.5% EtoH BLANK
H	YFP YPD + 6.5% EtoH 255µl	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH	120µl YPDE 6.5% EtoH Discard 135µl	120µl YPDE 6.5% EtoH BLANK

Figure 3.2 Plate layout for evolution 8 calibration curve.
The layout includes dilution series for CFP and YFP ancestral cells in three treatments 4.5% and 6.5% ethanol and a zero ethanol control. Rows A and B are a serial dilution of CFP in YPD.

This produced an estimate of the OD_{600} and fluorescence for each halving of the culture density from a saturated population. The completed calibration plate was subjected to the standard regime of shaking and reading used for each evolution for 30 minutes (six cycles of shaking). This length of time is insufficient for division in stationary phase cultures but does allow the cells to disperse and the optical density

to settle to a consistent value. The readings at 30 minutes were used to prepare the calibration curve.

No difference in the OD_{600} values for the CFP and YFP cultures was detected, neither were there changes due to the ethanol content of the medium, which suggested no substantial cell lysis in YPD containing 4.5% or 6.5% ethanol (see Figure 3.3).

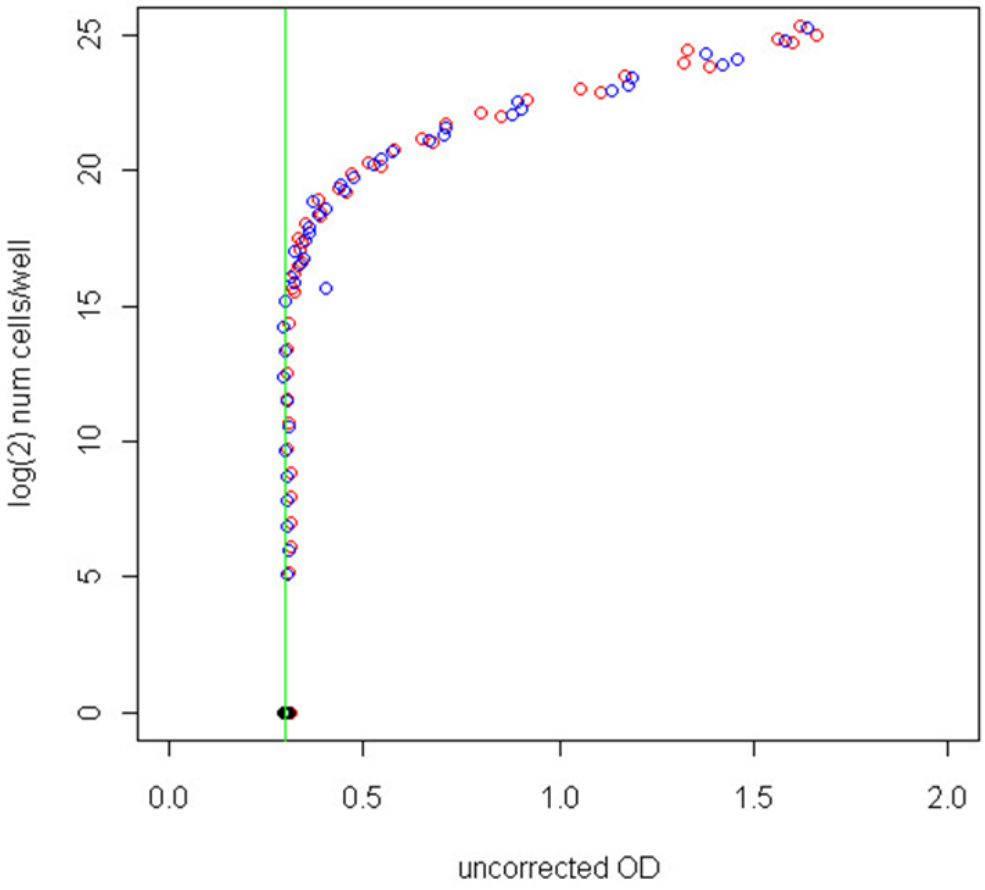


Figure 3.3 Calculated cell count against OD_{600} for evolution 8 calibration.
The cell count is a log to the base 2. The plate was obtained by a BMG Omega plate reader. Red circles represent the ancestral YFP strain and blue circles represent ancestral CFP. The vertical green line indicates an OD_{600} of 0.3 which is close to the blank value of the plate and medium with no added cells. A population in excess of ~$2^{\wedge 15}$ cells (~$3.3x10^4$ cells) is required for detecting the optical density of cells over the background value and the OD_{600} shows a consistent relationship with cell count up to an OD600 value of at least 1.6. The optical density of both strains is very similar as demonstrated by the substantial overlap of their calibration data

When the OD600 data was blank corrected and both the corrected OD_{600} and the log (base 2) cell count were log transformed then a linear relationship emerged for log corrected OD_{600} greater than -4 (a corrected OD_{600} of 0.18). A linear model of this data was made using R, see the green best fit line in Figure 3.4.

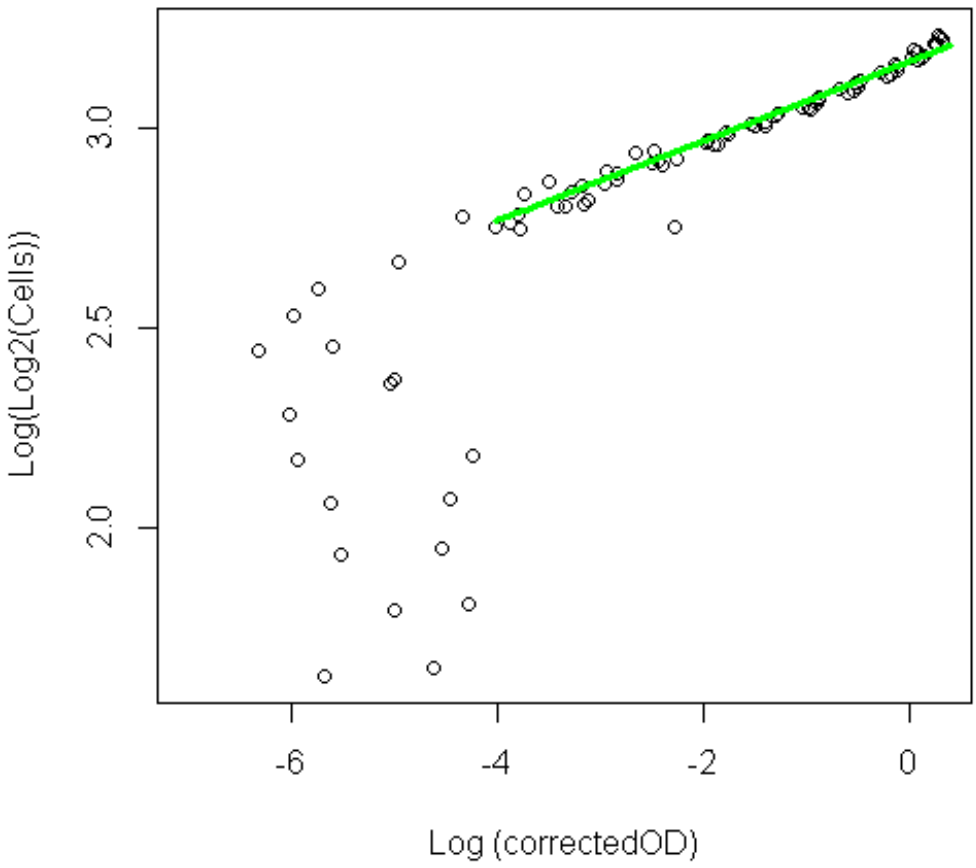


Figure 3.4 A plot of the log transformed (log2) cell count versus the log corrected OD600 for evolution 8 calibration data.
Calibration was done on the BMG Omega used for evolution 8. A linear fit was made using R for log (corrected OD600) greater than -4 (equivalent to a corrected OD600 of 0.18). The best fit line is shown in green (Intercept = 3.167+/-0.003 s.e., slope 0.100+/- 0.002 s.e.).

Fitting this transformed data gave the relationship shown in Equation 3.1. The same procedure was followed for the BMG Fluostar used for evolution 10. The fitted parameters for each evolution are given in Table 3.1

$$log2\ cells\ = e^{(a+b*\log(correctedOD))}$$

Equation 3.1

Figure 3.5 shows the fit using these parameters to the evolution 8 calibration data.

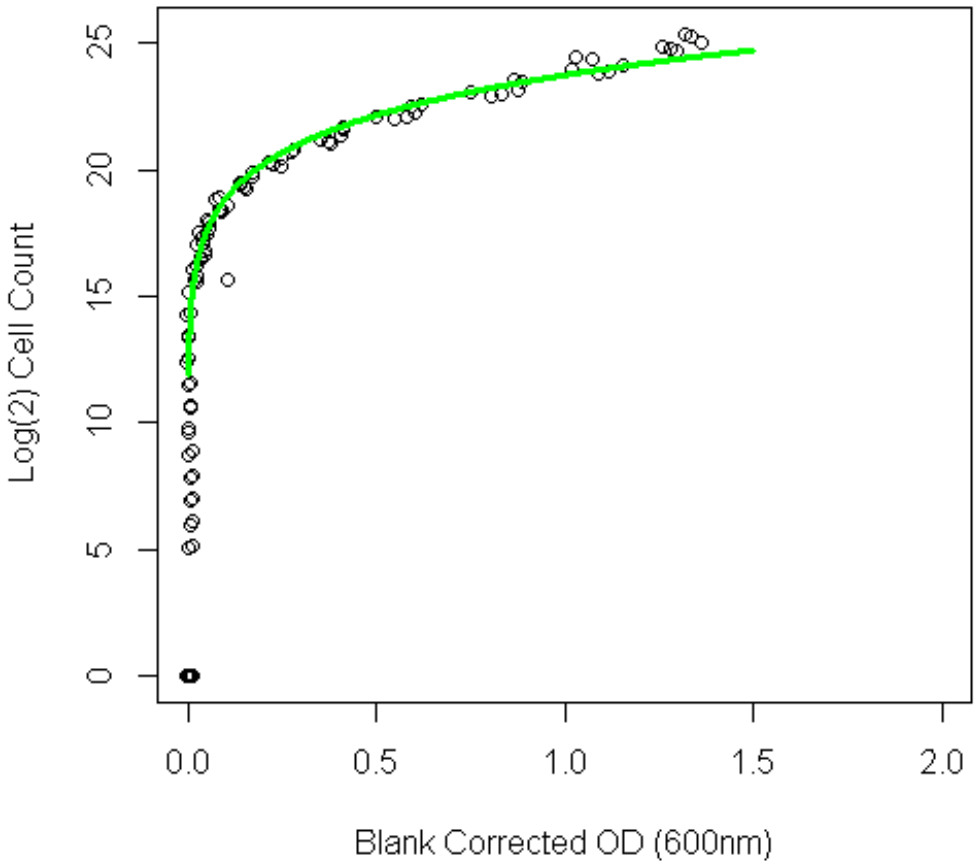


Figure 3.5 Log (base 2) cell count against blank corrected OD_{600} for the BMG Omega calibration data for evolution 8.
The best fit line, in green is derived from Equation 3.1 where the parameters a=3.167+/-0.003s.e. and b= 0.100+/- 0.002 s.e.

Evolution	Machine	Min OD_{600}	A	b
8	BMG Omega	0.18	3.1671 +/-0.0034 s.e	0.0997 +/-0.0020 s.e
10	BMG Fluostar	0.09	3.0926 +/-0.0054 s.e	0.1660 +/-0.0053 s.e

Table 3.1 Calibration curve parameters for cell count against OD_{600} for BMG plate readers.
Showing the minimum reliable OD_{600} and parameters a and b in Equation 3.1

The parameters derived from the curve fitting were used in a Perl script to convert the median peak OD_{600} curves into a log to the base 2 of the cell concentration. This concentration was used to calculate the serial dilution required for the correct bottleneck in the evolution series.

3.4.2. Fluorescence Calibration

There is fluorescence from the plate and medium, particularly in the CFP band. The blank fluorescence of the plate and medium measured in the CFP band is higher than that for YFP (see Figure 3.6). Typical fluorescence values for blank wells in the BMG FluoStar are ~$3.3x10^4$ for CFP and ~$7x10^3$ for YFP. The BMG Omega has smaller absolute values but they vary by a similar order of magnitude; ~$1.2x10^4$ for CFP and $1.3x10^3$ for YFP.

From the calibration but the lower limit for detection of cell fluorescence above background is approximately $1x10^5$ cells. Since $1x10^4$ is ~$2^{\wedge 13}$ and $1x10^5$ is ~$2^{\wedge 16\ 6}$,

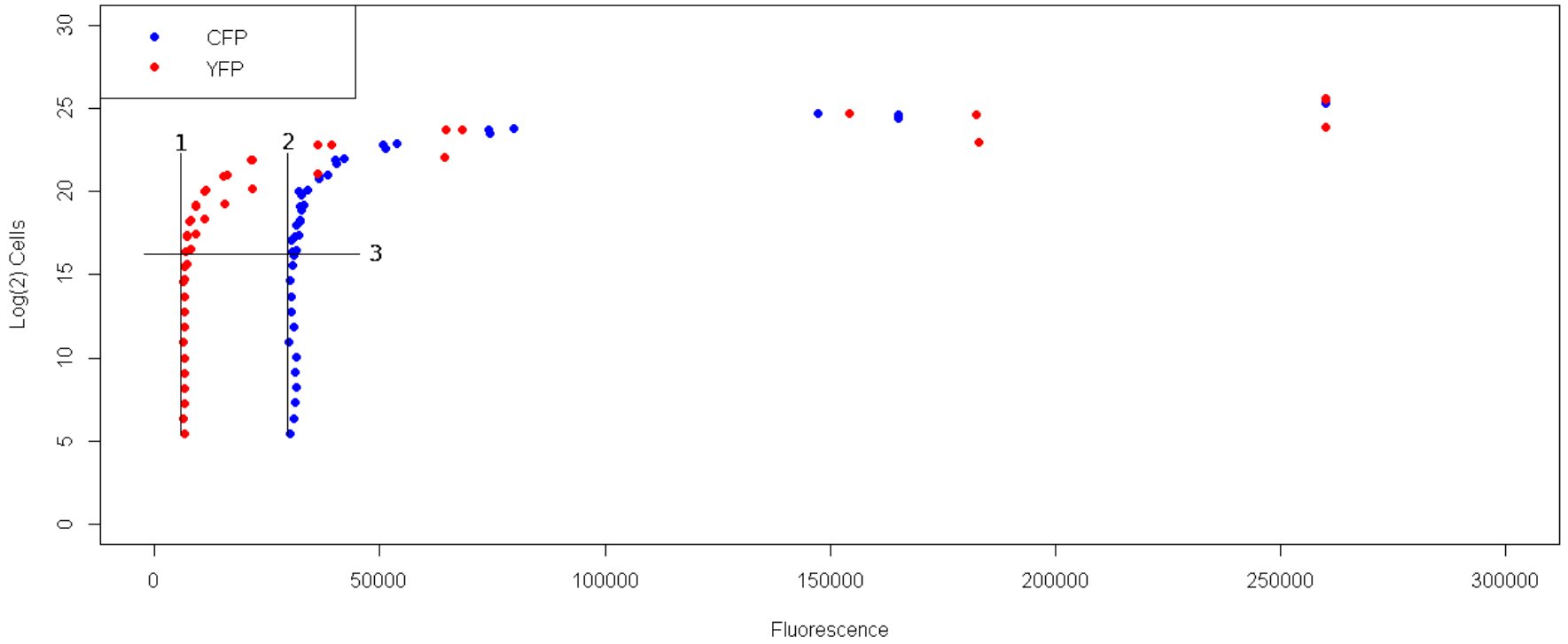


Figure 3.6 Calibration curves for log(2) cell count plotted against unblanked fluorescence for overnight cultures of ancestral CFP and YFP strains.
Cultures were counted in a Nexcelom Cellometer and serially diluted to give a range of CFP and YFP fluorescence values. Vertical line 1 shows the blank fluorescence in the YFP band and vertical line 2 shows the greater blank fluorescence in the CFP band. Horizontal line 3 shows the value at which both CFP and YFP are detectable above background. Otherwise the curves are similar in form and their minimum and maximum cell counts for accurate measurement.

The strain comparisons in chapter 5 used an initial inoculum of $1x10^4$ cells so there are three to four divisions before yeast fluorescence can be detected above the background due to growth medium and plate (see Figure 3.6). Meanwhile the background fluorescence is not constant; it falls during the period before growth is detectable, particularly in the first hour and more substantially in the CFP band of fluorescence than in that of the YFP (see Figure 3.7 below), the lower wavelength of CFP excitation may also bleach more effectively. There is an intensive sampling regime of fluorescence excitation every five minutes so this fall in background fluorescence is likely to be due to photo-bleaching (Shaner et al., 2005). Some photo-bleaching of the protein fluorophores is also likely but their structure confers resistance to bleaching (see Chapter 2, section 2.5.4). Bleaching of the protein fluorophores was not quantified.

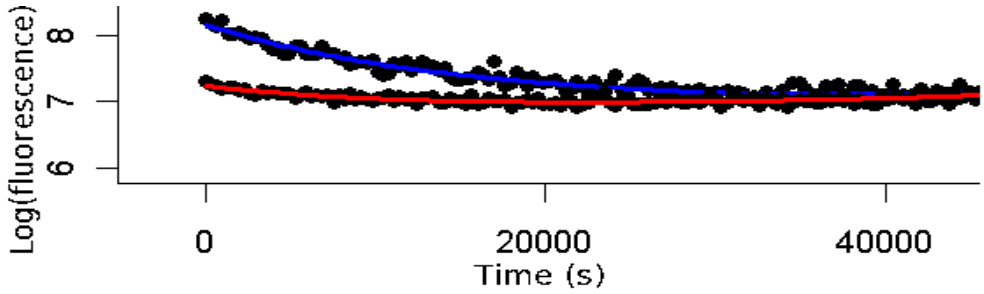


Figure 3.7 Part of a plot of the time course (time in seconds) of the natural log of the fluorescence data for the CFP strain Evo10_20_CFPe_F10_2 competing with the ancestral YFP strain.
A pspsline fit of the CFP fluorescence is in blue and the YFP fluorescence is in red. Note the minimum fluorescence for each strain has been blanked to a value of 1000 units (see Chapter 5, section 5.2.1).

3.5 Evolution protocols

All the plates for an evolution experiment came from the same batch (see section 3.3 for details of plates and plate readers). Lids were taped but the plates were not sealed. Whilst this allowed gaseous exchange, it also allowed the evaporation of water and ethanol from the medium. This was mitigated by placing blank wells in a ring on the periphery of each plate; for example see Figure 3.8, the layout for evolution 6. Substantial evaporation occurred in these blank wells, particularly at the plate corners. Evaporation limits the length of experiments to approximately 3 days and this puts an upper bound on the magnitude of the ethanol stress, particularly a fixed stress that is imposed from the outset.

The plates for evolutions 6 and 8 were read in the PolarStar Omega, and the FluoStar Omega plate reader was used for evolution 10 (see section 3.3 for details of

machines and plate reading scripts). The fluorescence calibrations were set to the same value throughout the evolution.

Wells on the plate layouts are identified by their row letter and column number. Figure 3.8 shows the layout used for evolution 6 (a 4.5% fixed ethanol stress). In addition to blanks around the outside of the plate, there are two columns of single strain cultures. Column 2 has the CFP strain and column 11 has the YFP strain. These single strain cultures allowed the growth curves of each strain to be compared, allowed the calculation of CFP fluorescence from the YFP strain, and provided some means of estimating potential cross contamination between wells. No contamination of the single culture wells by the other fluorophore strain was detected. It is not possible to conclude from this that no cross contamination occurred but no contaminating strain managed to grow beyond the background fluorescence. Evolution 10 used the same general layout as figure Figure 3.8 but the fixed ethanol stress, constant across all wells on the plate, differed between plates. The first two plates were grown with no ethanol, thereafter each plate had 0.5% more ethanol than the last up to a concentration of 9%.

	1	2	3	4	5	6	7	8	9	10	11	12
A	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank
B	YPD + 4.5% EtOH Blank	CFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	YFP + 4.5% EtOH	YPD + 4.5% EtOH Blank
C	YPD + 4.5% EtOH Blank	CFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	YFP + 4.5% EtOH	YPD + 4.5% EtOH Blank
D	YPD + 4.5% EtOH Blank	CFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	YFP + 4.5% EtOH	YPD + 4.5% EtOH Blank
E	YPD + 4.5% EtOH Blank	CFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	YFP + 4.5% EtOH	YPD + 4.5% EtOH Blank
F	YPD + 4.5% EtOH Blank	CFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	YFP + 4.5% EtOH	YPD + 4.5% EtOH Blank
G	YPD + 4.5% EtOH Blank	CFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	YFP + 4.5% EtOH	YPD + 4.5% EtOH Blank
H	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD + 4.5% EtOH Blank

Figure 3.8 Plate layout for evolution 6.
This evolution series had a fixed ethanol stress with 4.5% ethanol added to the culture medium of every well. The layout shows the peripheral location of blanks (in amber yellow), the location of the CFP/YFP competition wells in columns 3-10 of rows B-G (in red), and the single culture columns in rows B-G for CFP (column 2 in blue) and YFP (column 11 in yellow).

Evolution 8 had three ethanol treatments on one plate (see Figure 3.9) with separate blanks for each treatment.

	1	2	3	4	5	6	7	8	9	10	11	12
A	YPD Blank	YPD + 4.5% EtOH Blank	YPD + 6.5% EtOH Blank	YPD Blank	YPD + 4.5% EtOH Blank	YPD + 6.5% EtOH Blank	YPD Blank	YPD + 4.5% EtOH Blank	YPD + 6.5% EtOH Blank	YPD Blank	YPD + 4.5% EtOH Blank	YPD + 6.5% EtOH Blank
B	YPD + 6.5% EtOH Blank	CFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	YFP + YPD	YPD Blank
C	YPD + 4.5% EtOH Blank	CFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	YFP + 4.5% EtOH	YPD + 4.5% EtOH Blank
D	YPD Blank	CFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	YFP + 6.5% EtOH	YPD + 6.5% EtOH Blank
E	YPD + 6.5% EtOH Blank	CFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	CFP + YFP + YPD	YFP + YPD	YPD Blank
F	YPD + 4.5% EtOH Blank	CFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	CFP + YFP + 4.5% EtOH	YFP + 4.5% EtOH	YPD + 4.5% EtOH Blank
G	YPD Blank	CFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	CFP + YFP + 6.5% EtOH	YFP + 6.5% EtOH	YPD + 6.5% EtOH Blank
H	YPD + 6.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD Blank	YPD + 6.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD Blank	YPD + 6.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD Blank	YPD + 6.5% EtOH Blank	YPD + 4.5% EtOH Blank	YPD Blank

Figure 3.9 Plate layout for evolution 8.
This evolution series had fixed ethanol stresses of 4.5% and 6.5% and controls with no added ethanol. The layout shows the peripheral location of blanks (controls are white, 4.5% are coloured amber yellow and the 6.5% are puce), the location of the CFP/YFP competition wells in columns 3-10 of rows B-G (in green with shade darkening with increasing ethanol concentration), and the single culture columns in rows B-G for CFP (column 2 in blue, the shade darkening with increased ethanol concentration) and YFP (column 11 with shades ranging from pale cream to orange with increasing ethanol).

The OD_{600} was checked to ensure that all cultures on a plate had entered stationary phase. For evolution 8, which had three different ethanol treatments, this meant that the controls with no ethanol remained in stationary phase longer than the 4.5% or 6.5% ethanol treatment wells. Similarly, for evolution 10 this required the extension of the fermentation up to the three day limit for higher ethanol concentrations. Once the next plate had been set up, the old plate was refrigerated at 5°C for 24 hours so that the next plate could be monitored for contamination of the blanks and the plate restarted if required. If more than two blanks were contaminated the plate was restarted. Such contamination occurred in evolution 10 with plates 10 and 13. Once the risk of contamination was past the plates were frozen at -80°C after mixing with an equal volume of 50% glycerol.

The cell concentration of stationary phase plates was determined to calculate the dilution required for the next plate. It would have been desirable to have different dilution rates for each well however this was not possible in practice. For those plates with a single ethanol stress applied across the whole plate, the same dilution was applied to all the wells. However, for evolution 8 three different dilution regimes were applied for each of the ethanol treatments; the 0% ethanol control and the 4.5% and 6.5% ethanol treatments.

Two methods were used to calculate the bottleneck serial dilution. For evolution 6, 15µl from each well was diluted in 1815µl of growth medium to give an initial dilution of x122. Four of these diluted populations had their cell counts determined by Nexcelom Cellometer (see section 3.2 for details) and the cell count was used to determine a second dilution. The wells counted were: B02, which contained solely CFP, E07, and F04 (a CFP/YFP competition) and G11 (a YFP single culture).

Although the cell counts in the four wells appeared very similar, it was considered undesirable to base the dilution of the whole plate on a few cell count determinations. Accordingly for the bottleneck calculations for evolution 8 and 10 a different method was employed. A calibration curve was prepared for OD_{600} against cell count (see 3.4) and that was used to estimate the median cell content for each well and calculate the serial dilution required for the desired bottleneck.

Each evolution series used different bottleneck sizes. Evolution 6 used the smallest bottleneck of 2000 cells. Evolution 8 had a bottleneck of 10,000 cells and evolution 10 used 15,000 cells. Neither method of determining stationary phase cell count can distinguish between live and dead cells. If dead cells constitute a substantial proportion of the culture then the true bottleneck count will be smaller. Similarly, the true bottleneck size may differ from the desired value if flocs have not been completely broken up prior to serial dilution or the culture settles.

Each round of measurement in the plate readers produces a flat text file with times and values for readings. Perl scripts were used to concatenate these files into a single file. The data was then loaded into a MySQL database.

3.6 Isolation of strains

Populations were chosen which showed changes in their final fluorophore ratio curves indicating that one population had substantially increased at the expense of the other (see chapter 4). The chosen populations were spread on a YPD agar plate

and examined under a Leica MZ10 fluorescence stereo microscope (Leica Microsystems Ltd, Stereo and Macroscope Systems, CH 9435 Heerbrugg, Switzerland) capable of exciting fluorescence in either CFP or YFP. For evolution 6 and 8 four colonies with the prevailing fluorophore were picked (this was increased to seven for evolution 10) and grown up prior to competition with the ancestral strain with the opposite fluorophore. For strains isolated from evolution 6 or 8 with fixed added ethanol stresses this competition was carried out under the conditions of their experimental evolution. For evolution 10, in which the ethanol stress varied from 0 to 9%, the ancestral competitions were carried out in 4.5% ethanol. Some of the derived strains which outcompeted their ancestors were selected for a more comprehensive survey of their performance in competition with their ancestor across a range of added ethanol stresses (see chapter 5) and had their genomes sequenced (see chapter 6).

4. Analysis of experimentally evolving populations

4.1 Introduction

The core research aim was to isolate and characterise the phenotypes and genotypes of *Saccharomyces cerevisiae* strains with enhanced ethanol tolerance using many parallel experimentally evolved batch cultures (see Section 2.7). Batch culture can more easily be scaled up to run many experimental evolutions in parallel, is closer to the current industrial niches for *S. cerevisiae* in brewing and bioethanol production in which enhanced ethanol tolerance would be of greatest benefit and may be closer to the ecological niche in which *S. cerevisiae* evolved (see Section 2.1.2 “Ethanol secretion, competition and niche construction”). To isolate adaptively evolved strains it is necessary for the experimental evolution parameters to be within effective ranges. These parameters include the correct intensity of ethanol stress (whether a fixed ethanol stress or one that progressively increases); if the ethanol stress is too small then there may be no fitness benefit for ethanol tolerance (indeed there may be a trade-off in which enhanced ethanol tolerance mutations may decrease fitness without significant ethanol stress) but too large a stress may greatly decrease the viable population size. If the effective population, which is dominated by the size of the bottleneck population, is too small then the effects of neutral drift will vitiate selection and no adaptive strains will be found. However if the bottleneck population is too large then many generations will be required to before any adaptive clone can sweep and a complex competing (clonally interfering) mix of genetically distinct populations is likely to arise (see Section 2.4.4 “Mutation rate, population size and adaptive trajectories”).

Running many batch cultures in parallel can enable a greater diversity of mutations to be found, including rare mutations. However, that does require an effective method for identifying those cultures in which an adaptive evolution has taken place for strain isolation. It is also important to determine whether the experiment needs to run for more generations (too few adaptive changes have arisen) or whether it can be stopped (continuing the evolution is likely only to generate multiple clonally competing populations). Such a method should also be able to indicate whether effective values for the experimental parameters have been chosen. Analysis of fluorophore ratio changes was used to monitor each experimental evolution. A single adaptive evolution in one fluorophore is accompanied by a curving trajectory of

fluorophore ratio change which can be fitted to a Dirac equation model (see the following section for a detailed discussion of this approach). This model has parameters which estimate of time of origin (tau) and the effect size (alpha) for the adaptive change. However, if several mutational events occur then a discontinuous curve will result which cannot be fitted using the Dirac equation. Accordingly truncation of the dataset prior to the emergence of secondary changes was used to fit a number of fluorophore curves. However, such truncation may allow the fitting of changes in ratio due to neutral drift, particularly as the effective population sizes decrease. Therefore, a random sampling simulation able to model the effect of different bottleneck sizes was used to generate test fluorophore curves using neutral drift. These simulated datasets were analysed to determine whether they were distinguishable from the putative adaptive evolution curves using the data truncation and Dirac equation fitting routines and hence whether the fluorophore ratio changes observed in these truncated datasets are likely to represent genuine adaptive change.

Effective tools for analysing fluorescently tagged strains in competition can be utilised to explore adaptive change for other stresses or changed growth conditions (e.g. different carbon sources, culture temperatures etc.). Such approaches potentially enable a greatly increased number of parallel evolutions to be scrutinised.

4.1.1. Dirac Curve Fitting

Estimates were made of the fluorophore ratio change parameters using the nonlinear (weighted) least-squares (NLS) method in the base R distribution (version 2.15.2, http://cran.r-project.org/) fitting data to the Dirac equation used in Hegreness (Hegreness et al. 2006). This equation has parameters for relative fitness (α) and a time of origin (τ). Since the initial fluorophore ratio in each evolution was not exactly 1, we also fitted an offset for the starting ratio (c). The sign of the deviation (an implicit fourth parameter) was determined by the sign of the slope of a straight line fit to the data. Curves with increasing ratio (CFP increasing with respect to YFP) were fit using Equation 4.1:

$$log10\ \frac{CFP}{YFP} = c + log10(1 + 0.5 * e^{\propto\ t-\tau}\ )$$

Equation 4.1

And curves with decreasing ratio were fit to equation Equation 4.2:

$$log10\ \frac{CFP}{YFP}\ = c - log10(1 + 0.5 * e^{\propto\ t-\tau}\ )$$

Equation 4.2

The Dirac equation assumes that there was parity of fitness between the fluorophores prior to a change (which occurs before τ and may be prior to the start of the experiment) and that the size of the fitness increase in one fluorophore remains constant. This is most unlikely to be true for every adaptive change; irrespective of secondary mutations arising in the sweeping fitter strain, and adaptive changes in the competing fluorophore, frequency dependant selection may occur which favours an allele at low frequency but confers a smaller benefit or a detriment at higher frequency e.g. a cheater strain may very successfully exploit a community resource at low frequency but have lower fitness at higher frequency.

The starting strains were intended to be isogenic apart from the fluorophore (something largely, though not wholly, supported by sequencing, see chapter 6, section 6.3.3) and evolutions 6 and 8 have ethanol stresses and other environmental conditions that remain constant throughout the experimental evolution. Accordingly there is reason to suppose that the starting strains have equal fitness (variation is effectively zero) and any solitary beneficial mutations that arise in these experimental populations will have a consistent, improved fitness which enables their fluorophore curves to be fitted using Dirac equation. However, after plate 2 in evolution 10, 0.5% more ethanol is added with each plate. This both increases the stress due to ethanol toxicity, but also adds more carbon to each well as a potential respiratory substrate. It is an open question whether beneficial mutations at low ethanol concentrations have the same fitness magnitude (or even the same sign) at higher concentrations. Accordingly, beneficial mutations may arise in evolution 10 that cannot be fitted with the Dirac equation; those curves that can be fitted may have parameter estimates that are incorrect. Furthermore, discontinuities in the evolution 10 fluorophore curves that might be interpreted as clonal interference in evolutions 6 or 8 could be due to changes in the magnitude or sign of a mutation's fitness.

The "time of origin" tau ($\boldsymbol{\tau}$) is measured as a threshold departure from the initial fluorophore ratio (in this case when the log of the divergence = 0.5). A beneficial mutation that survives stochastic loss can spend a varying number of generations at low frequency subject to random drift before it any sweep due to its beneficial fitness

can occur. Beneficial mutations of small effect spend longer at low frequency, both the mean time to sweep (and the variance) are greater with lower effect size as is the risk of stochastic loss (Sniegowski & Gerrish 2010 “Beneficial mutations and the dynamics of adaptation in asexual populations”). Moreover, the rate of sweep to the threshold value for τ, is not independent of the size of the fitness benefit. Despite all these caveats tau (τ) provides a useful measure of the time of origin of mutations. Small (or negative) values for tau can indicate that a sweep is due to standing variation in a population rather than a new mutation. Tau (together with the fitness parameter alpha) allows a comparison of the emergence of beneficial mutations in experimental evolutions with differing selection pressures and bottleneck sizes. It can also be used to explore correlations between waiting time for beneficial mutations and effect sizes in different experiments.

If a beneficial mutation is present as a substantial fraction of the starting population (or arises very quickly) then there will be no initial period of constant fluorophore ratio in the data. This will result in large errors in the estimate of $\boldsymbol{\tau}$. Conversely if a fluorophore ratio curve has few data points diverging from the initial value (perhaps because it arises late in the evolution series) then the errors in the estimate of the fitness size ($\boldsymbol{\alpha}$) will be large.

The rate of supply of beneficial mutations relative to the rate they can sweep through the population will determine whether it is possible to determine significant features of the competition using curve fitting. If the competing populations are close to the adaptive peak or the effective population sizes are too small, then the rate of supply of beneficial mutations will be very low. Such ‘weak’ mutation is likely to result in the populations in many wells not substantially diverging from the initial fluorophore ratio.

If the mutation rate is high relative to the strength of selection then more than one mutation is likely to be sweeping at any one time, either within a single fluorophore population or in both of them, and clonal interference may prevent the interpretation of the curves produced using these methods (see chapter 2, section 2.4.4). In the extreme case, the fluorophore ratio will remain constant despite profound evolutionary change if beneficial mutations of identical fitness arise at the same time in both fluorophores. Some cases of clonal interference can be interpreted by truncating the data if there is a delay before the onset of subsequent mutations. If one beneficial mutation is succeeded by another of larger fitness in the same fluorophore then this may visibly steepen the curve. Conversely, if a beneficial

mutation in one fluorophore is succeeded after an interval by a competing mutation in the other fluorophore then then the curve will become shallower.

Small bottleneck sizes lead not only to the loss of beneficial mutations but large fluorophore ratio changes by neutral drift. Data truncation to analyse mutations that become subject to clonal interference increases the risk that short runs of drifting fluorophore ratios will be mistaken for the sweep of beneficial mutations.

Modelling was used to determine the minimum bottleneck sizes (see section 4.2.2) likely to minimise the time for beneficial mutations detectably to sweep without promoting large fluorophore ratio changes due to drift. Ultimately, the isolation of common strains with greater fitness than the ancestors would be unlikely from populations in which drift dominated beneficial mutation. A minimum range of ratio change was used to filter out those wells in which no detectable population change occurred.

Whilst modelling may determine the desired bottleneck, implementing that bottleneck for each well is not straightforward. It is important to monitor and change the dilution rate if the carrying capacity changes, either due to an increasing stress (as in evolution 10) or due to possible increases in fitness or a change in the proportion of respiration to fermentation during growth. If the final population size changes and the dilution remains fixed then the bottleneck size will change. The serial dilution required to achieve the desired bottleneck was calculated from analysis of populations from three wells counted using a cytometer in the usual manner (see section 3.2).

If beneficial mutations arise in a population and it grows better than the average of the sampled wells then it may become under-diluted, i.e. have a larger bottleneck than the target bottleneck value whilst wells in which no beneficial mutations arise may become over-diluted and hence possibly more prone to drift later in an evolution series. With the protocol used (see section 3), the cytometer cannot distinguish between live and dead cells. A protocol for live/dead staining is available for this cytometer but trials of this were unsatisfactory and it was not done. The fraction of dead cells may increase with stress making the true bottleneck smaller than the desired value. The ethanol stresses increased to the highest levels (9%) in evolution 10 and hence this source of error is likely to be largest in this series.

Care was taken (see section 3.5) to break up flocs and suspend cells but cell adhesion, both to the well and to other cells, is a problem for accurate counting and

serial dilution. Mutated strains which promote cell adhesion could become under- or over-represented in the suspension and it is particularly problematic if mutations promote non-random cell-cell adhesion i.e. there is co-operative formation of flocs which exclude wild-type cells, so called 'green beard' mutations (Smukalla et al. 2008). The *FLO* genes (*FLO1*, *FLO5*, *FLO8* and *FLO11*) have a high mutation rate with sequences frequently being deleted, amplified or recombined. This is probably due to their tandem sequence repeats (which make detecting sequence variation problematic) and their sub-telomeric locations (Sato et al. 2001; Sato et al. 2002). The repeats make short sequence reads hard to uniquely map so sequence changes to these genes are difficult to detect with next-generation genome sequencing.

This method, using two fluorophores and the fitting of a simple curve, is limited to the analysis of initial mutations. If a mutation of large effect has already swept through a population, leaving the fluorophore ratio markedly changed, or several mutations compete, then further analysis of that population is problematic. However, curve analysis can indicate in which populations in which rare mutations of large effect have arisen, allowing concentration of expensive and time consuming effort to isolate and sequence strains. It can also indicate when to terminate evolution experiments by analysis of fitness errors; the large fitness errors for the evolution 8 control would likely have been reduced by continuing the experiment for a few more plates. This would also have allowed a beneficial mutation to sweep to greater prevalence making it easier to isolate. Conversely there is no need to continue an experiment if all populations have had a ratio sweep that can be fitted with small alpha fitness errors.

The simple sweeps of fluorophore ratio that can be fitted with the Dirac equation are not seen when there is clonal interference. If the interference is minor then there may be minor departures from normality of the distribution of residuals and the errors may appear to be auto-correlated. Illingworth and Mustonen use a maximum likelihood approach to make inferences about multiple mutations from marker trajectories (Illingworth & Mustonen 2011).

4.2 Methods

4.2.1. NLS fitting of fluorophore ratios

Estimating the parameters of a non-linear model by non-linear (weighted) least squares (NLS) is not an exact science (Bates & Watts 1988). The fitting algorithm requires starting parameter estimates and sometimes two apparently very similar

curves (e.g. datasets that differ only by a single point) can fit, or fail to fit, with the same starting parameter estimates.

Evolution	Asymptote	Alpha	Tau	Min Range
6	0.27	0.427	5	0.12
8	0.27	0.427	5	0.10
10	0.10	0.500	12	0.15

Table 4.1 Dirac equation showing starting parameter estimates for NLS curve fitting for each evolution series and the minimum data range

The non-linear least squares package in the base R (2.15.2) installation was used. If the range of the data is above a minimum, the R script tries to fit each population fluorophore ratio in turn using the initial parameter estimates for each dataset (see Table 4.1). A minimum data range is critical since without it NLS will attempt to fit curves to insignificant measurement variation or drift. If NLS successfully converges on a fit, the R script then extracts the parameter estimates and their standard errors for the converged solution and stores them along with the errors of fit (with the mean squared error and root) before performing the Lilliefors test from the R package 'nortest' on the errors. Lilliefors is a modification of Kolmogorov-Smirnov test for departures from normality and it differs solely in the critical values used. Whichever test is used, normality tests on small numbers of points are not very powerful (Steinskog et al. 2007) but it provides an index of whether the fit is improving or degrading as the dataset is truncated. If there are more than five points in the dataset, the script then removes the last point and tries to fit the data again. If the dataset before the last truncation was successful, the script re-uses the parameter estimates of the previous model, rounded to two decimal places (fitting will fail if the initial estimate correctly guesses the converged parameters) as starting guesses. If this attempt fails, it will revert to the starting guesses and retry. This process is repeated until the dataset has been truncated to five points.

If the data is truncated so that it has few points that deviate from the initial ratio, an extreme curve with no biological relevance may be fitted that has very low mean squared error. Figure 4.1 shows a Dirac curve that has been fitted to an over-truncated dataset. There is a shallow slope in the fluorophore ratio but a large scatter in the points. Truncating the data to five points produces an extreme curve with a smaller mean squared error than extending the curve. Over-truncating the

data may also enhance the problem of distinguishing between a beneficial sweep and neutral drift.

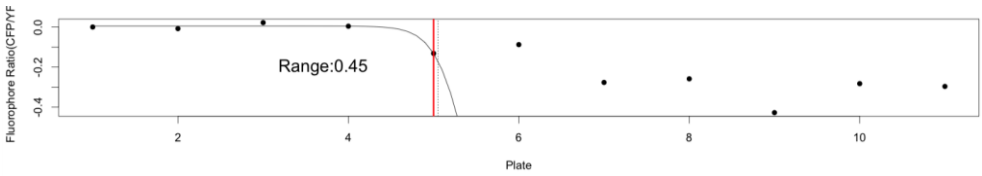


Figure 4.1 An example Dirac curve fitting to an over-truncated dataset.

Accordingly, deciding on the optimal truncation is a manual process, assisted by plots of the mean squared error, Lilliefors statistic and estimate of the alpha parameter to ensure the dataset is not under or over truncated. The optimal truncation removes the minimum number of points. If part of a dataset reflects a beneficial sweep that changes the fluorophore ratio following a Dirac curve but it is followed by clonal interference that changes the slope, then there are likely to be two transitions in the data during truncation. If a Dirac fit is possible, each removal of a point affected by clonal interference is likely to reduce the mean squared error. When the last point substantially affected by interference is removed, the Lilliefors statistic may improve as the distribution of errors for the fit is likely to be closer to normality. As points affected by large clonal interference are removed the fitness parameter (alpha) and onset parameter (tau) are likely to change but once the data is truncated to a set that fit the curve well then further truncation is likely to leave the estimates of fitness and onset little changed until the data becomes radically over-truncated and the curve merely fits an initial stable ratio and one or two points that deviate (as in Figure 4.1). Such over-truncation is likely to result in sharp transitions in the fitness and onset parameters.

Figure 4.2 shows a curve fitted to a dataset from evolution 10 well G08. The data is truncated to point 16.

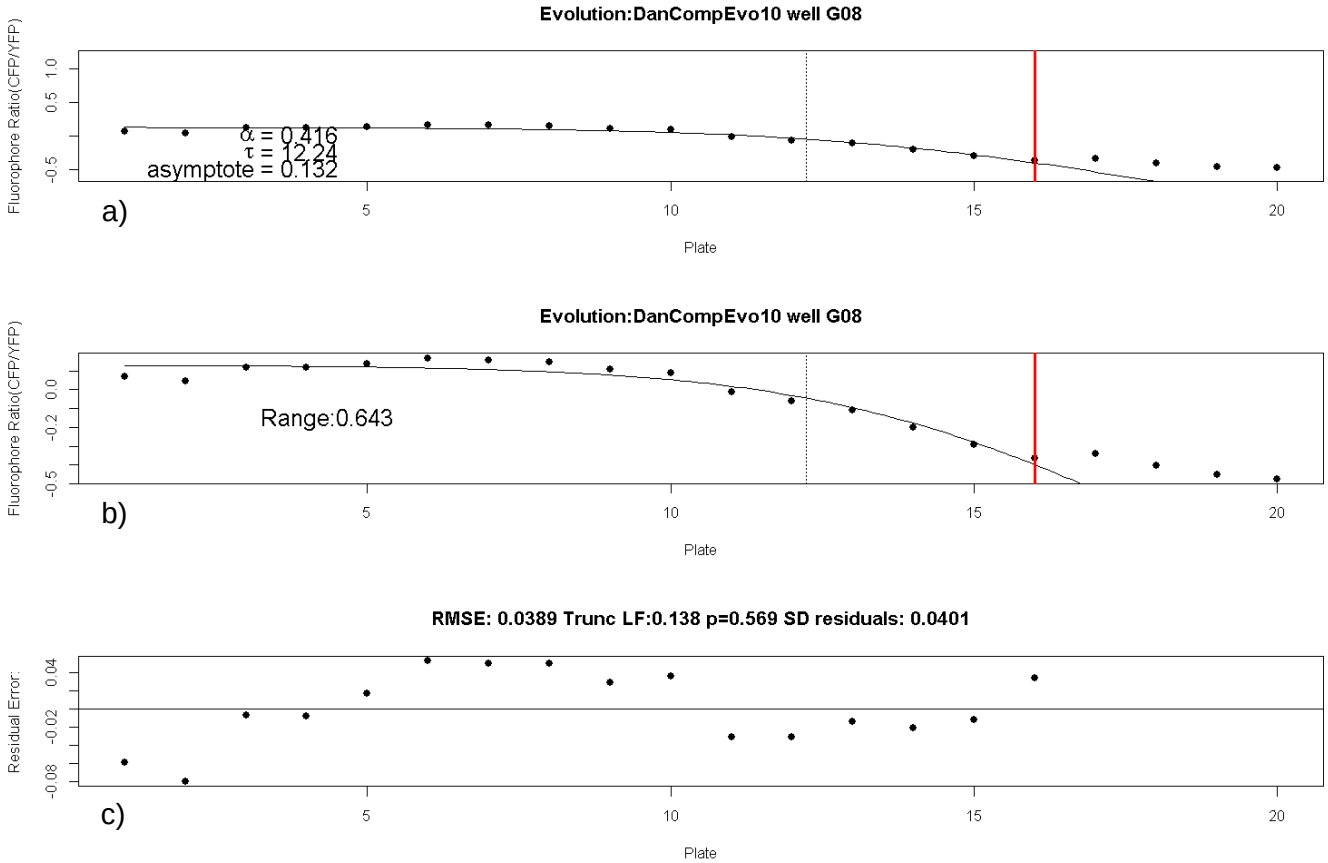


Figure 4.2 Fluorophore ratio data against plate for evolution 10, well G08:
a) shows the log fluorophore ratio data (CFP/YFP) with a curve fitted from points 1 to 16. The y axis is a standard range for curve comparison. b) shows the same data and fitted curve with the y axis range adapted to the data range. c) shows the residual error of the points in the truncated dataset to the fitted curve. The red vertical line in a) and b) shows the last point fitted in the truncated dataset.

Figure 4.3 shows the error and ancillary data used to decide on the proper truncation point for evolution 10 well G08: the mean squared error of the fit (Figure 4.3.a), the Lilliefors statistic (Figure 4.3.b) and the fitness (alpha) parameter (Figure 4.3.c). The reasoning for this truncation was as follows. Each deletion of a point from 20 to 16 reduces the mean squared error of the fit substantially; the 16 point dataset has about half the error of the complete dataset. There is only minor impact on the Lilliefors statistic by truncating from 20 to 16 (with a minor increase with the 17 point dataset) but truncating to 16 points gives the minimum Lilliefors value. The fitness parameter (alpha) increases steadily, with a small slope, when the data is truncated to sixteen points.

Truncating the dataset from 16 to 13 points has very little effect on the mean squared error, suggesting that all these points support the same fitted curve. However, at fewer than 15 points the Lilliefors statistic increases sharply.

The minimum mean squared error is found with a truncation to 11 points. However, an eleven point dataset produces the maximum value for Lilliefors test and a sharply bigger alpha value. Referring to Figure 4.2 part b) it can be seen that truncating to an eleven point dataset reduces the curve to fitting an early deviation from the initial ratio.

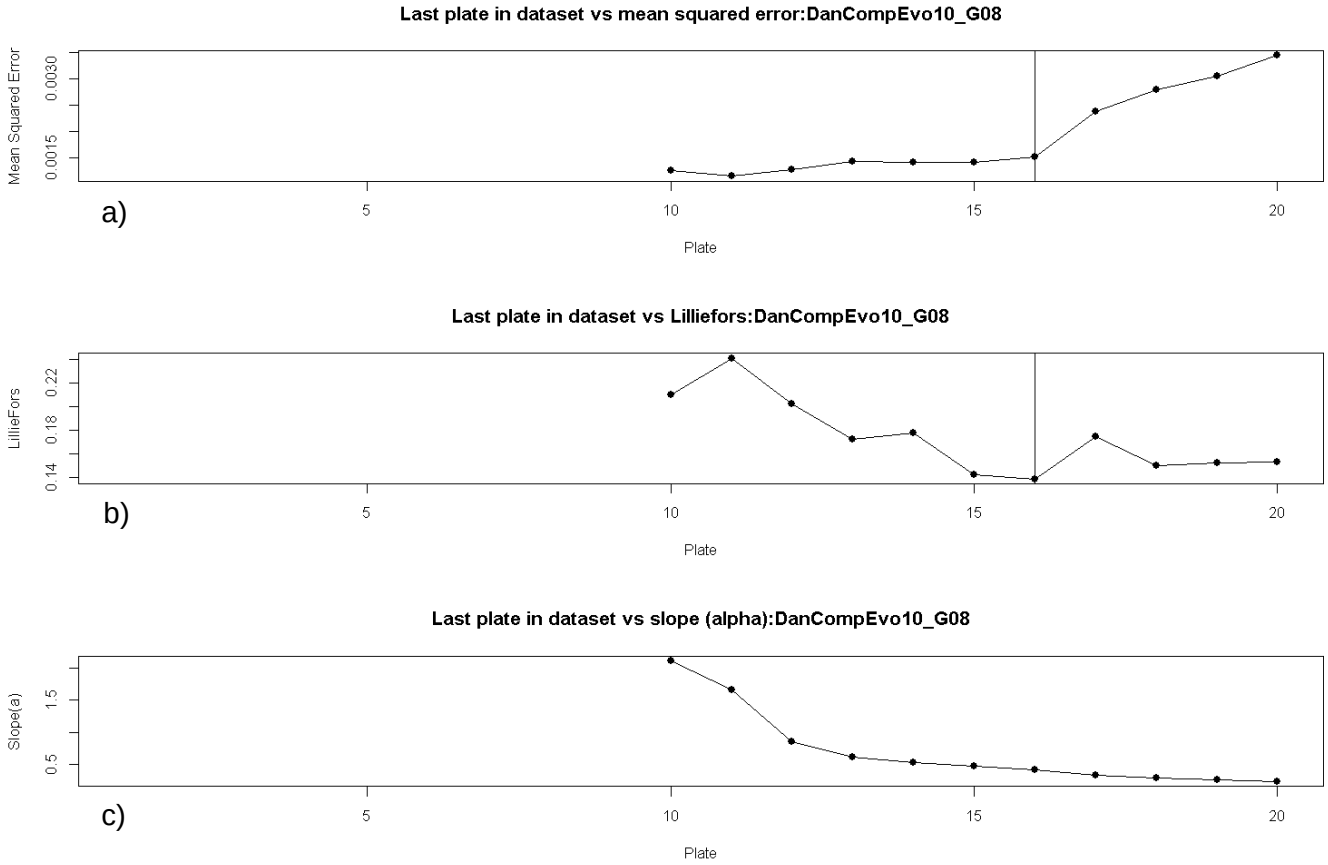
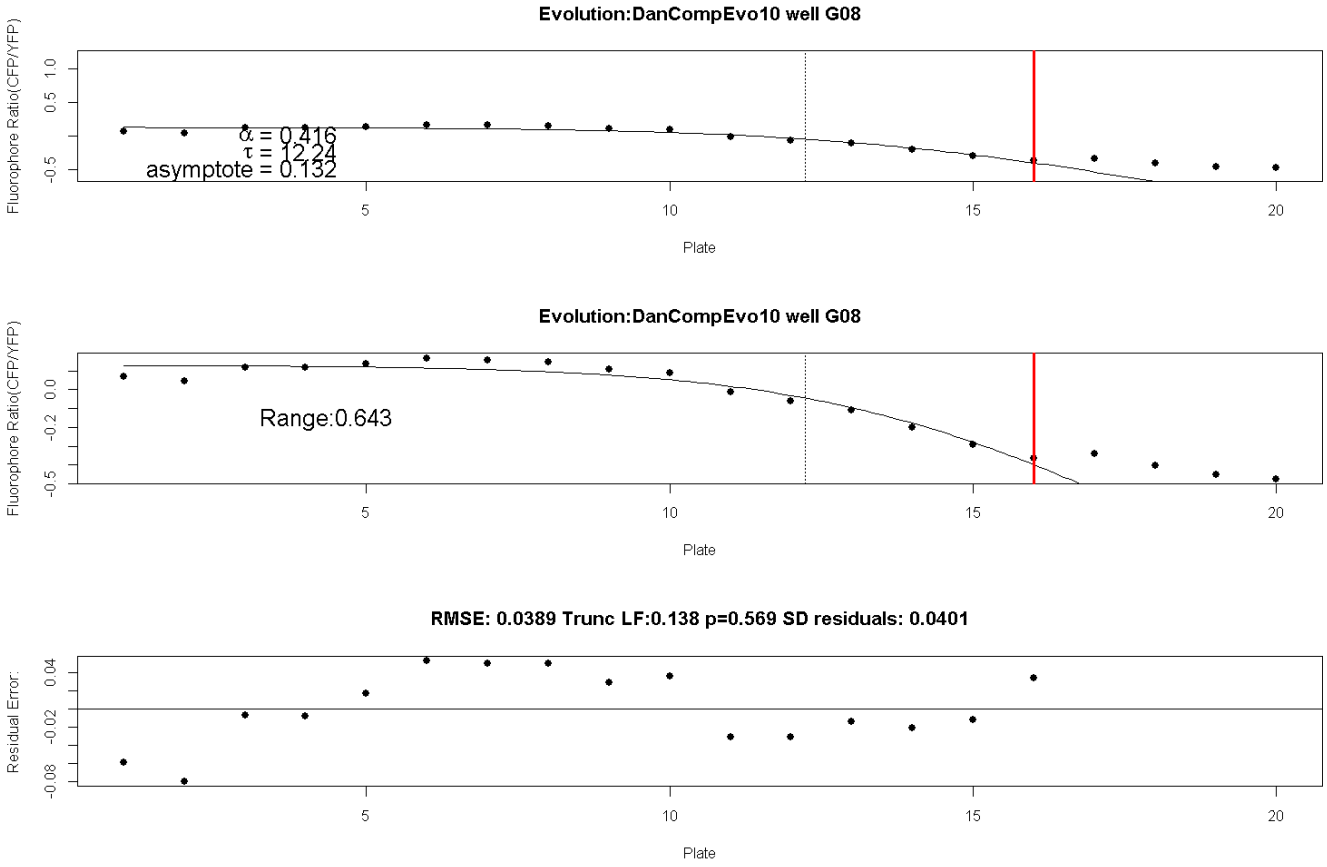


Figure 4.3 Plots of mean squared error, Lilliefors statistic and Dirac alpha (fitness) estimate for the fluorophore ratio dataset for evolution 10 well G08.
The data is derived from NLS curve fitting of the Dirac equation to the data truncated to that point. The larger the value of Lilliefors statistic the less likely the data is to be normally distributed. No fit was possible below 10 points. The vertical black line indicates the last point in the truncated dataset.

Whilst truncation of the dataset can be justified it is frequently not possible to completely resolve problems of data fitting. The fit to the sixteen point truncation of evolution 10, well G08 leaves residual errors that appear auto-correlated and the initial fluorophore ratio appears to rise before falling (see Figure 4.2).



Both of these features suggest the data prior to point 16 is more complex than a simple sweep favouring a YFP mutation and may be affected by neutral drift or clonal interference.

4.2.2. Neutral drift modelling (null model)

The effect of varying bottleneck size on fluorophore ratio was modelled using a simple R script (see appendix I). The script models change in fluorophore ratio in repeated experimental evolution trials. The script accepts a bottleneck size, a final carrying capacity, a probability that a cell will divide each 'generation', and a number of plates. The 'experiment' starts with equal numbers of 'cells' marked with one or two fluorophores (either 0 or 1) that make up a population of the bottleneck size. Each cell in the population divides, or not, according to a division probability. Once every cell has had a chance to divide the total number of cells and the fraction of each fluorophore is stored and another round of division starts. As soon as the population reaches the carrying capacity division stops and random sampling

reduces the population to the bottleneck size. The range measures the deviation in both directions and is likely to be larger than the final deviation. The model was run for fifty times for each bottleneck size.

4.3 Results

4.3.1. Evolution 6 Carrying capacity and cell size

Three sets of experimental evolution experiments were carried out. Each of the evolutions had 60 experimental wells and 36 peripheral blanks. There were competitions between fluorophores in 48 of the experimental wells and single-strain cultures in the remaining 12. The evolutions differed in the number of plate transfers; evolution 6 had 10 plate transfers, evolution 8 had 14 transfers and evolution 10 had 19 transfers. All the wells in experimental evolution 6 were subjected to a fixed ethanol stress of 4.5% while evolution 8 had three treatments with 0, 4.5% and 6.5% ethanol and evolution 10 had the same ramped ethanol stress applied to all wells, increasing from 0% at the start of the experiment to 9.5% at the end.

Evolution 6 had the smallest bottleneck size with 2000 cells and carrying capacity between 1.5-3x10^7 cells which gave around 13 doublings (approximately 140 generations in total for the evolution). Evolution 8 had a bottleneck size of 1x10^5 cells and evolution 10 had 1.5x10^5 cells which gave approximately 11 and 10 generations per plate, ~150 in total for evolution 8 and ~200 generations for evolution 10 (see General Methods 3.5 for evolution protocols).

For evolution 6, an estimate of the final cell counts in four wells was made using the Nexcelom Cellometer. The mean of these population estimates was used to calculate the serial dilution required for the desired bottleneck of 2000 cells and this dilution was applied to all the wells containing culture on the plate. Practical difficulties (the length of time required to count more wells and the cell culture volume required) limited the number of wells counted. Two of the counted wells contained single strain cultures; B02 contained CFP and G11 contained YFP, while the other two wells contained competitions between the strains (E07 and F04). Figure 4.4 shows the cell counts at stationary phase for these wells. The plot for each well shows similar values for population size. This is important since greatly varying population sizes might have suggested that the approach of applying a fixed serial dilution to all wells might have greatly under or over-diluted populations in each well. It is unclear why there is a significant drop in the maximum cell count between plate 1 to 2; the plate batch, machine calibration and the medium were all unchanged throughout the series. After plate 2 there is a moderate fall in cell count over the course of evolution 6 from a mean value of 2.26x10^7+/-3.34x10^6s.e. for plate 2 to 1.90x10^7+/- 1.43x10^6s.e. for plate 10 (the cell count of plate 11 was not measured).

This reduction in cell count is accompanied by an increase in cell size. Figure 4.5 shows the cell sizes of the four monitored wells; after the first two plates all show an increase in cell sizes with the largest increases occurring in the two competition wells, E07 and F04.

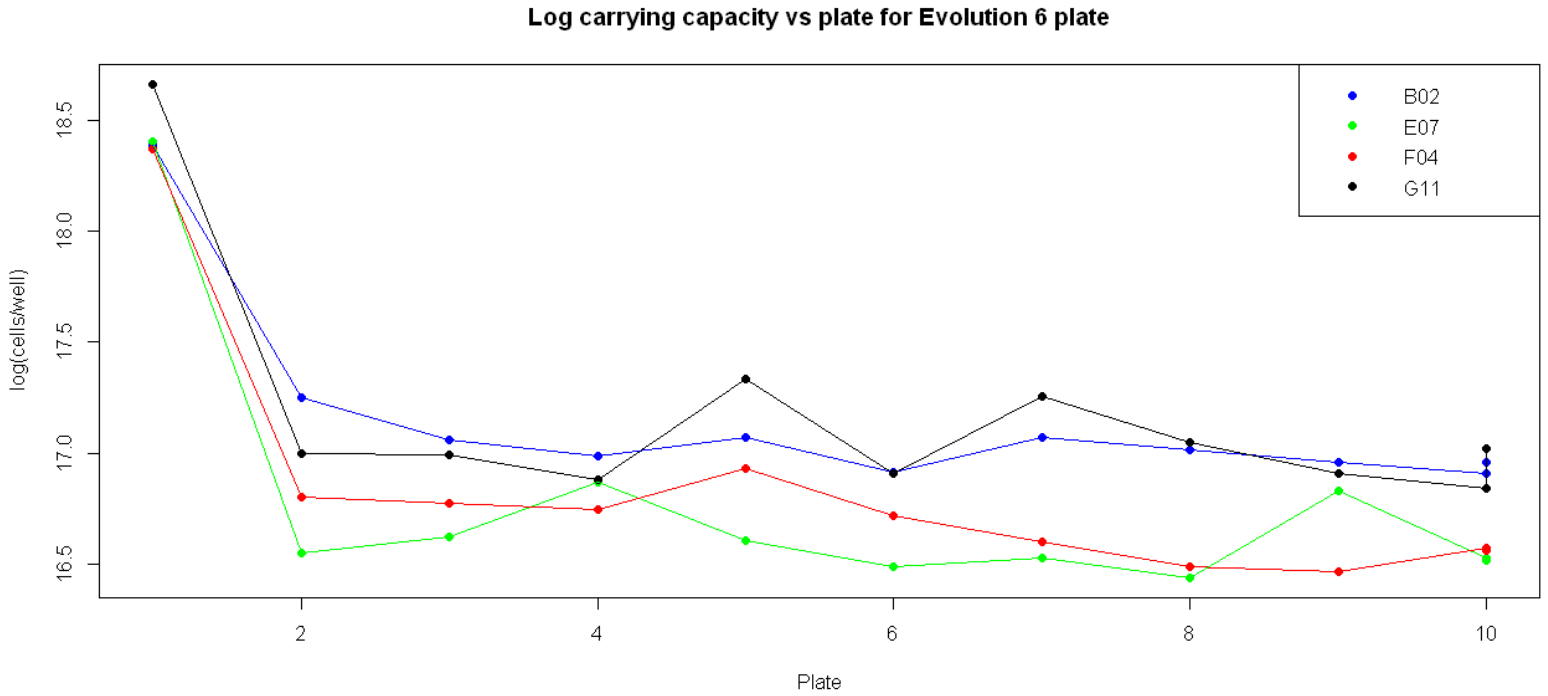


Figure 4.4 A plot of the log final cell count (cells/well) versus plate for evolution series 6.
Well B02 (CFP) and well G11 (YFP) contain a single fluorophore strain, whilst E07 and F04 have both in competition. Cell contents were determined using a Nexcelom Cellometer counting a 20µl sample of a x100 dilution of the sampled wells (see General Methods section **3.2**). All plates were post-diauxic slow growth phase.

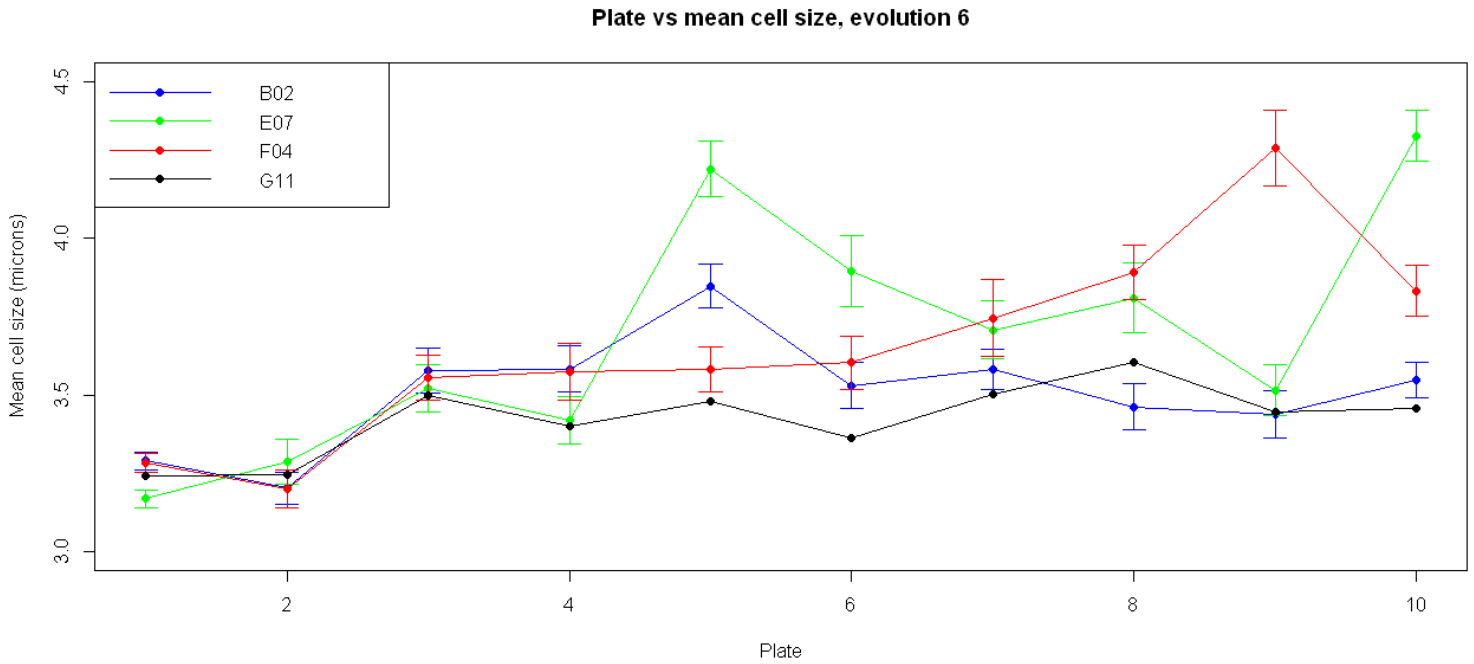


Figure 4.5 Mean cell size in microns against plate for four wells in evolution 6.
B02 (CFP only), G11 (YFP only), and E07 and F04 which are wells in which the CFP and YFP strains compete. The error bars are standard errors. Cell sizes were measured when the Nexcelom Cellometer counted cells (see General Methods section 3.2).

Both of the competition wells show changes in the ratio of the fluorophore signals; Figure 4.6 shows the fluorophore ratio change for competition well E07 and Figure 4.7 shows the change for F04. The fluorophore change for F04 has a small alpha fitness value (0.109) and negative tau value for onset (-9.152); see section 4.3.6 for a discussion of evolution 6 fluorophore ratio data.

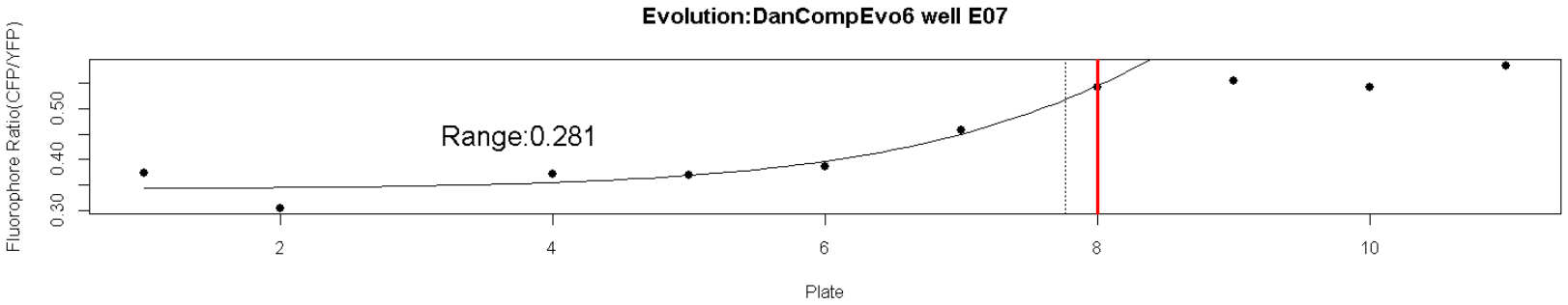


Figure 4.6 Fluorophore ratio change in evolution 6, well E07.
The x axis is the plate number in evolution 6 and the y axis is the log to base 10 of the ratio when the stationary phase raw CFP signal is divided by the raw YFP signal. A Dirac curve is fitted to the data truncated to plate 8 (alpha = 0.387, tau = 2.876); the red line indicates the last point fitted. The data range is 0.281 log fluorescence units.

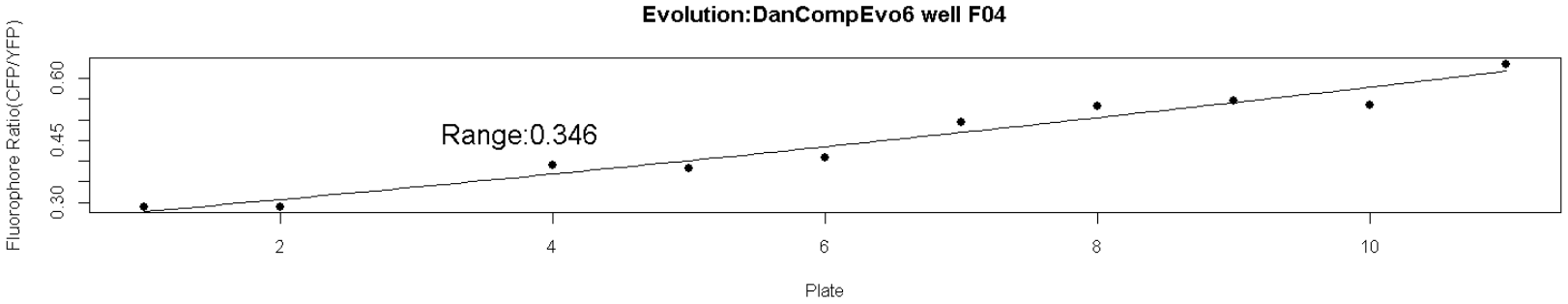


Figure 4.7 Fluorophore ratio change in evolution 6, well F04.
The x axis is the plate number in evolution 6 and the y axis is the log to base 10 of the ratio when the stationary phase raw CFP signal is divided by the raw YFP signal. A Dirac curve is fitted to the data (alpha = 0.109, tau = -9.152). The data range is 0.346 log fluorescence units.

4.3.2. Evolution 8 and 10 final optical densities

The serial dilutions required for the population bottleneck for experimental evolutions 8 and 10 were calculated, not from cell counts, but from estimates of the cells at stationary phase derived from the corrected OD_{600} using the calibration curve of cell density against the optical density at 600nm (described in Chapter 3, section 3.4.1: Optical density calibration). In general the Nexcelom Cellometer counts of diploid saturated cultures are smaller than haploid so it is likely that cell count calibration curves prepared using haploid ancestral cells overestimate cell numbers in diploid or polyploid cultures. Adapting this method to accurately assess counts in cultures that are mixtures of haploid diploid and polyploid is problematic.

Each of the ethanol treatments in evolution 8 showed an increase in the peak OD_{600} when compared with the starting populations (see Figure 4.8), this may be due to biomass increasing due to enhanced fitness. A research objective (see Section 2.7 “Research aims and goals”) was verifying that there had been adaptive change and increases in peak OD_{600} is one strand of evidence that this had occurred. Another objective (see Chapter 5) was the investigation of phenotypic change during experimental evolution. Fluorescence cytometry strongly supports ploidy increasing from haploidy to diploidy (and to triploidy in evolution 10) in all populations examined (see section 6.3.8) and in all isolated strains that out-competed their ancestors. This causes a predictable increase in cell size (see Figure 4.5 above and section 2.2.4: Physiology of ploidy change). It is unclear how the optical density of a cell suspension changes when cell sizes vary. These carrying capacities were not verified by plating out dilutions of the stationary populations and colony counting.

Figure 4.8 also shows that optical density increases in evolution 8 with added ethanol. This counterintuitive result could be explained by a greater proportion of more energetically efficient respiration in populations with added ethanol, either due to an effect of ethanol toxicity or because more carbon has been added to the growth medium in the form of ethanol. There is a substantial gap between the outliers of the 4.5% and 6.5% treatments in plate 1 but a good deal of overlap between them in plate 14, with the 4.5% treatment showing a larger variance between populations. This suggests that during the experimental evolution the populations in the 4.5% ethanol treatment have become more like that in 6.5%. Subsequent fluorescence cytometry of evolution 8 populations suggested that ethanol promoted a faster

transition to diploidy with increasing ethanol (see chapter 6, Figure 6.24) suggesting that ploidy is a key driver in optical density increase in ethanol treated populations.

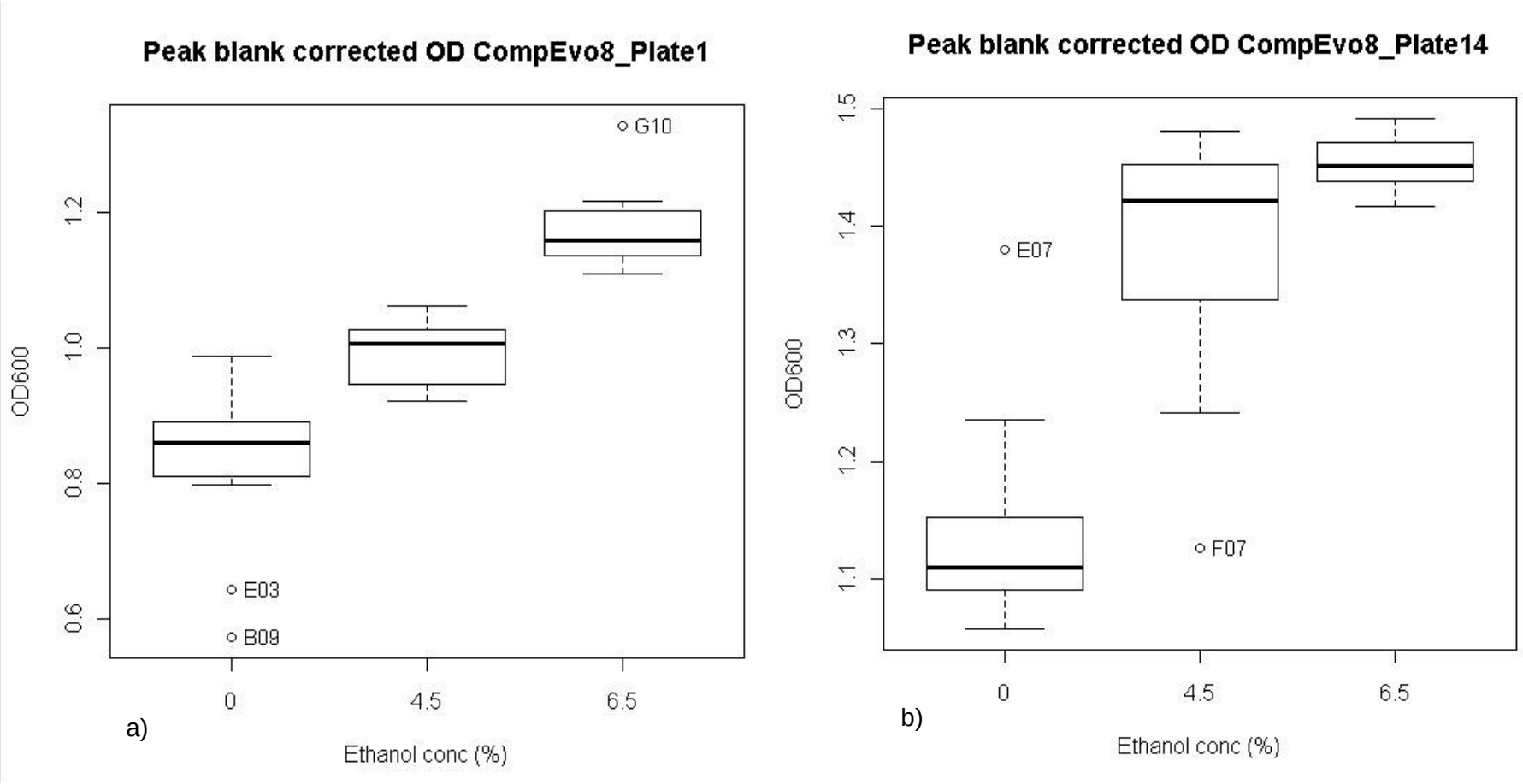


Figure 4.8 Box plots of the peak blank corrected OD_{600} for evolution 8 plate 1 (part a) and plate 14 (part b).
Please note differing y axes between a) and b). The black line in each box shows the median, the top and bottom of the box show the 25th and 75th percentiles. The whiskers above and below the box show 1.5 times the interquartile range (approximately two standard deviations) or the minimum/maximum values, whichever is the smaller. When outliers occur outside 1.5 times the interquartile range, they are named.

In evolution 10 there was a steady increase in peak optical density at 600nm up to plate 13, which contained 5.5% ethanol. Plate 14 (6% ethanol) had the highest OD_{600} values in the experiment. Subsequently the median OD_{600} values fall back to something close to the first plate without added ethanol. However above 6% ethanol there are more low outliers and plate 19 (8.5% ethanol) has a larger variance in the lowest quartiles. The population in well D09 is a low OD_{600} outlier from plate 7 to 17. Insofar as OD_{600} is an index of cell count, such low growth outliers are likely to be over diluted and have too small a bottleneck population. However, D09 does subsequently recover suggesting that the population was not irrevocably damaged by over-dilution. Subsequent analysis of the fluorophore ratio for D09 showed no significant change. It is possible that there is a physiological change in growing yeast above 6% ethanol that alters the relative fitness of different adaptations beneficial at lower ethanol levels. Alternatively, the rate of change in the ethanol stress (0.5% per plate) above 6% may be too fast for successful adaptation with this population size.

Figure 4.10 shows the fluorophore ratio curves for two populations in evolution 10, F10 and G10, chosen for further analysis and strain isolation. Both populations show strong fluorophore ratio change, favouring CFP in F10 (a) and YFP in G10 (b).

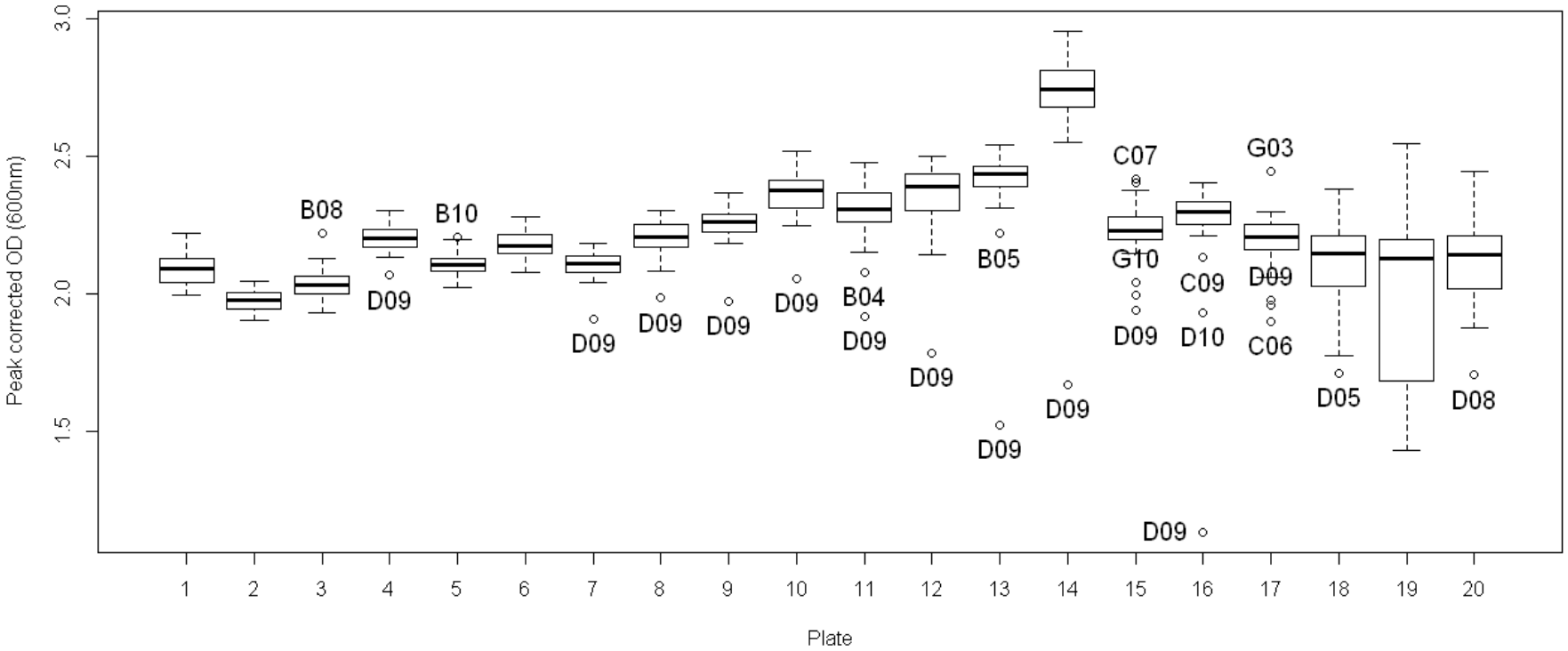


Figure 4.9 Box plots of the peak blank corrected OD_{600} for evolution 10.
The black line in each box shows the median, the top and bottom of the box show the 25th and 75th percentiles. The whiskers above and below the box show 1.5 times the interquartile range (approximately two standard deviations) or the minimum/maximum values, whichever is the smaller. Outliers that are outside 1.5 times the interquartile range are marked with points; some are named.

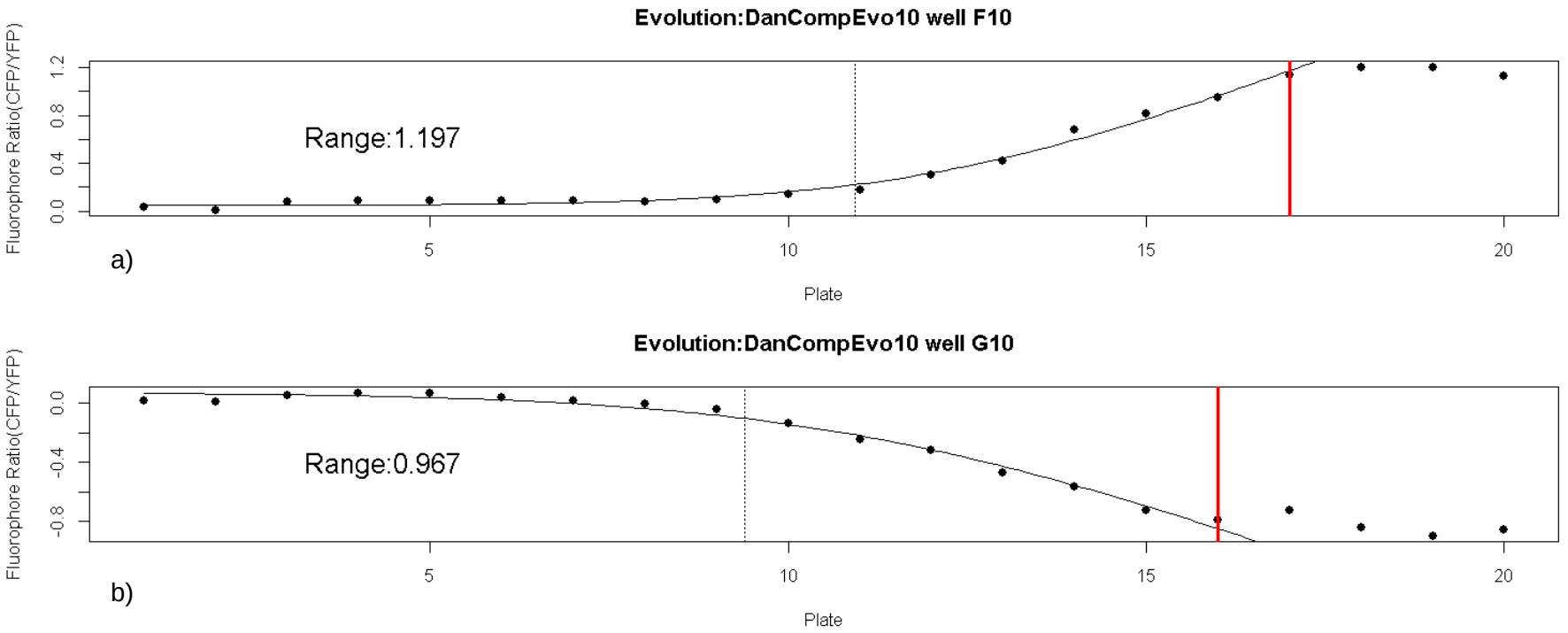


Figure 4.10 Plate versus fluorophore ratio curves, the log to the base 10 of the raw CFP count at stationary phase divided by the raw YFP count, for wells F10 and G10 in evolution 10.
Part a) is the curve for well F10 and part b) shows the curve for G10. Each curve shows a relatively large range of fluorophore ratio change (1.197 for F10 and 0.967 for G10) and good support for the fitted curve up to the red line, where the data in each case is truncated (point 17 for F10 and point 16 for G10)

Figure 4.11 shows the results of one ancestral competition for evolution 10 showing cultures grown from seven colonies competing against the ancestral strain with the complementary fluorophore. The derived CFP strains from population F10 have a much higher fluorophore ratio than the ancestors in competition with each other. This indicates a larger proportion of CFP. Conversely the derived YFP strains have a much lower ratio than the ancestral competition. All the derived strains have similar fluorophore ratios and none are close to that of the ancestors in competition. However, the second strains had marginally larger deviations from the ancestral ratio and they were chosen for further fitness analysis and whole genome sequencing with the strain designations Evo10_20_F10_CFPe_2 and Evo10_20_G10_YFPe_2.

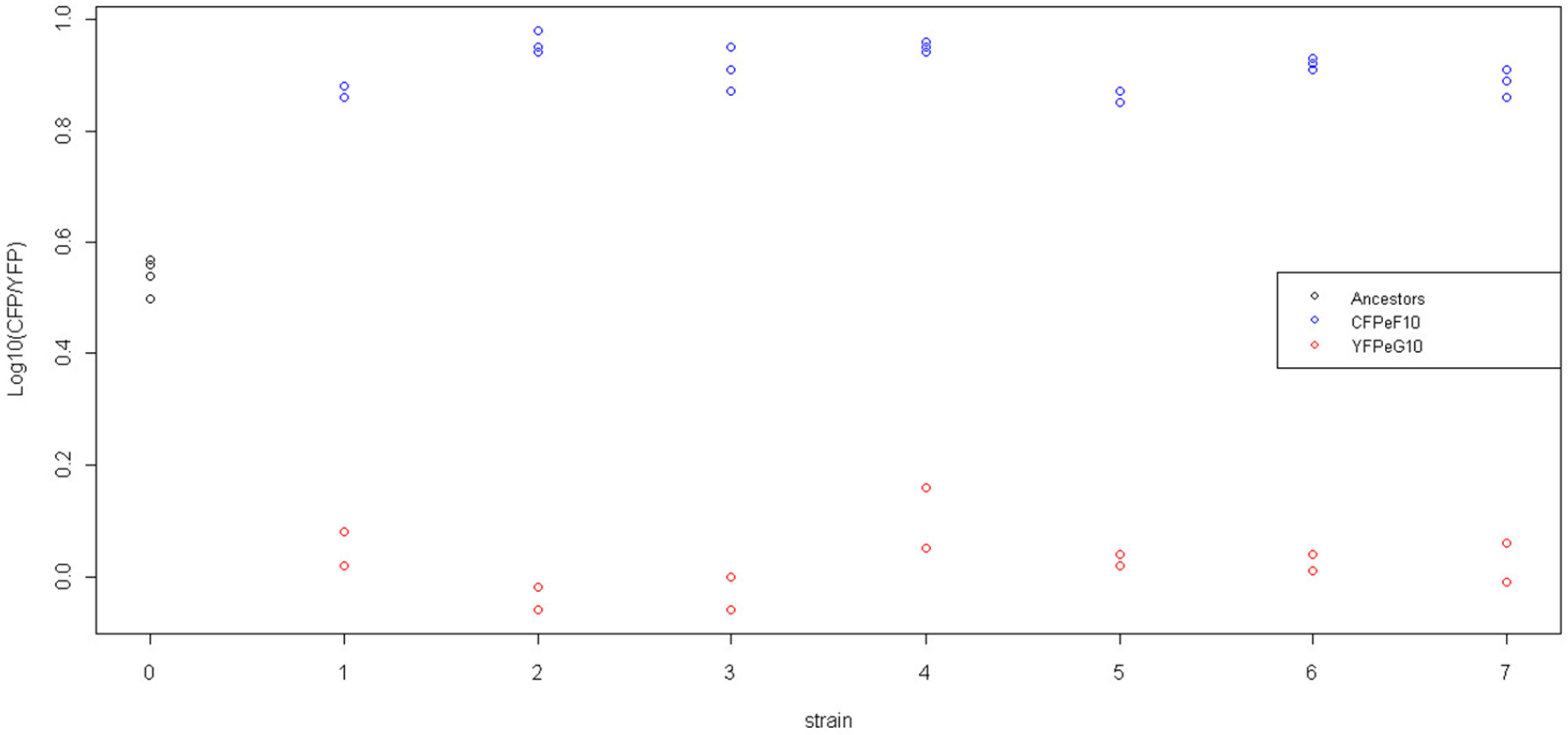


Figure 4.11 Log10 fluorophore (CFP/YFP) ratio for the ancestral competition of seven CFP strains isolated from F10 and seven YFP strains isolated from G10 in evolution 10.
Each strain was competed with the ancestor of complementary fluorophore. The F10 CFP strains are shown in blue and the G10 YFP strains are shown in red, each point is a replicate. The competition between the two ancestral strains is shown in black (strain 0)

4.3.3. Neutral drift modelling

The model suggests that fluorophore ratios can substantially change due to neutral drift in experimental evolution experiments with bottleneck sizes in the range 10^2 - 10^3. Figure 4.12 shows bottleneck size versus the range of population ratio change, the log to the base 10 of the ratio population1/population2, for modelling runs with eleven plates and a carrying capacity of $3x10^7$ cells (similar to that of evolution 6). At a bottleneck size of $5x10^3$ cells almost all the curves have a range of less than 0.1, as this reduces to $2x10^3$ cells, almost all have a range less than 0.2 and the ratio changes greatly increase as the bottleneck falls below 1000 ‘cells’.

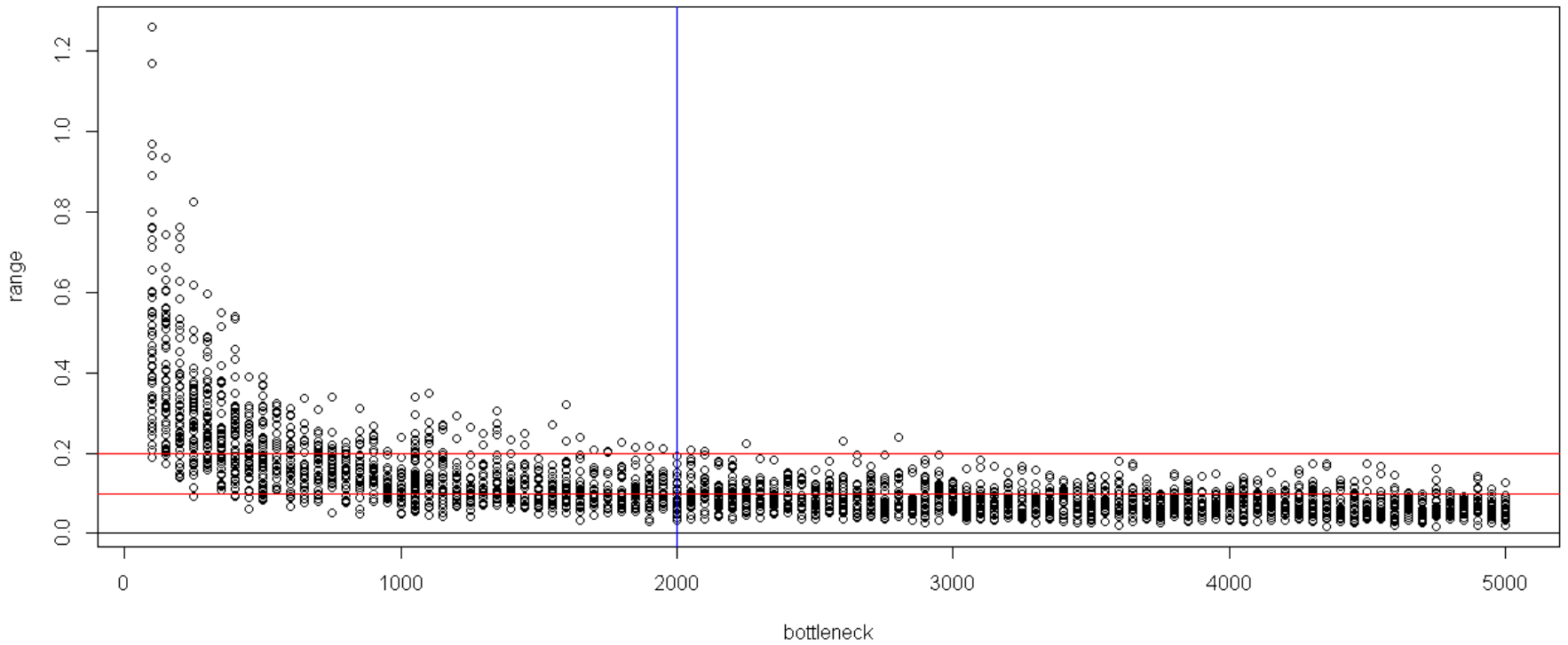


Figure 4.12 Experimental evolution modelling. A plot of bottleneck size versus the range of 'fluorophore' log(10) ratio change by random sampling for a simulated experimental evolution of eleven plates.
The carrying capacity was $3x10^7$, bottleneck sizes ranged from 100 to 5000 and the probability of 'division' each generation was 0.9, 25 trials at each bottleneck size. The blue line indicates the population bottleneck for evolution 6.

Table 4.2 shows the results of using non-linear least squares to fit the Dirac equation to fluorophore ratio curves generated by the random drift modelling script. Fifty sets of curves were produced with a carrying capacity of $3x10^7$ cells, a division probability 0.9 and bottlenecks varying from 100 to 15,000 cells. 86% of the model output with a bottleneck of 100 could be fit to the Dirac equation, the range of the fitted curves is substantial (0.5+/-0.038s.e.) and the mean fitness parameter, α, is large (0.73+/-0.135s.e.). This is larger than almost all the curves fitted to experimental data (see section 4.3.4). The fluorophore ratio data that could not be fitted also had a substantial range (0.362+/-0.06s.e.). Increasing the bottleneck size to 2000 cells approximately halves the number of fitted curves and reduces the range of fitted curves below 0.2 (mean 0.151+/-0.007s.e). At a bottleneck of 5000 or above, the maximum range of the ratio data falls below the 0.12 minimum for curve fitting.

Bottleneck size	Fitted Curves (% of 50 trials)	Mean range fitted curves +/- s.e.	Mean fitted alpha +/- s.e	Mean fitted tau +/- s.e	Mean range unfitted +/- s.e.
$1.0x10^2$	86	0.522 +/- 0.038	0.730 +/-0.135	5.11 +/- 0.845	0.362 +/- 0.06
$2.0x10^3$	42	0.151 +/- 0.007	0.433 +/- 0.9	7.21 +/- 1.732	0.078 +/- 0.004
$5.0x10^3$	0	NA	NA	NA	0.061 +/- 0.003
$1.0x10^4$	0	NA	NA	NA	0.049 +/- 0.003
$1.5x10^4$	0	NA	NA	NA	0.038 +/-0.002

Table 4.2 NLS Fitting of Dirac curves to the output of a neutral drift model with varying bottleneck sizes.
The carrying capacity of model $3.7x10^7$ cells, division probability per cycle of growth 0.9. Table shows percentage of fifty trials that could be fit, and the alpha and tau Dirac parameters as well the mean range of the fitted and unfitted curves.

Figure 4.13 shows the distribution and relationship of the Dirac curve parameters for those fluorophore ratio curves that could be fit from a neutral drift model run 50 times with a bottleneck size of 100 (carrying capacity $3x10^7$ and division probability 0.9). The distribution of tau is widely spread with many negative values. Negative tau values typically have large tau error bars and small values for alpha. In these cases it is difficult to establish the onset of a curve because there are no points representing a stable fluorophore ratio. Similarly, numerous points with large tau

values in Figure 4.13 have very large alpha error bars because only a few data points deviate from the initial ratio. Both of these extremes could be considered derived from short runs – potentiated by truncation. However, there are a number of fitted curves in the middle of the alpha/tau plot (with alpha ~0.5 and tau ~5) with small error bars for both alpha and tau. These curves (see Figure 4.14) have long runs of points that fit the Dirac equation well and the data has a substantial range; it is unclear how these could be distinguished from a sweeping beneficial mutation.

The alpha distribution generated by neutral drift in Figure 4.13 is positively skewed, with a long tail of rare larger values, similar to the distribution of beneficial mutations.

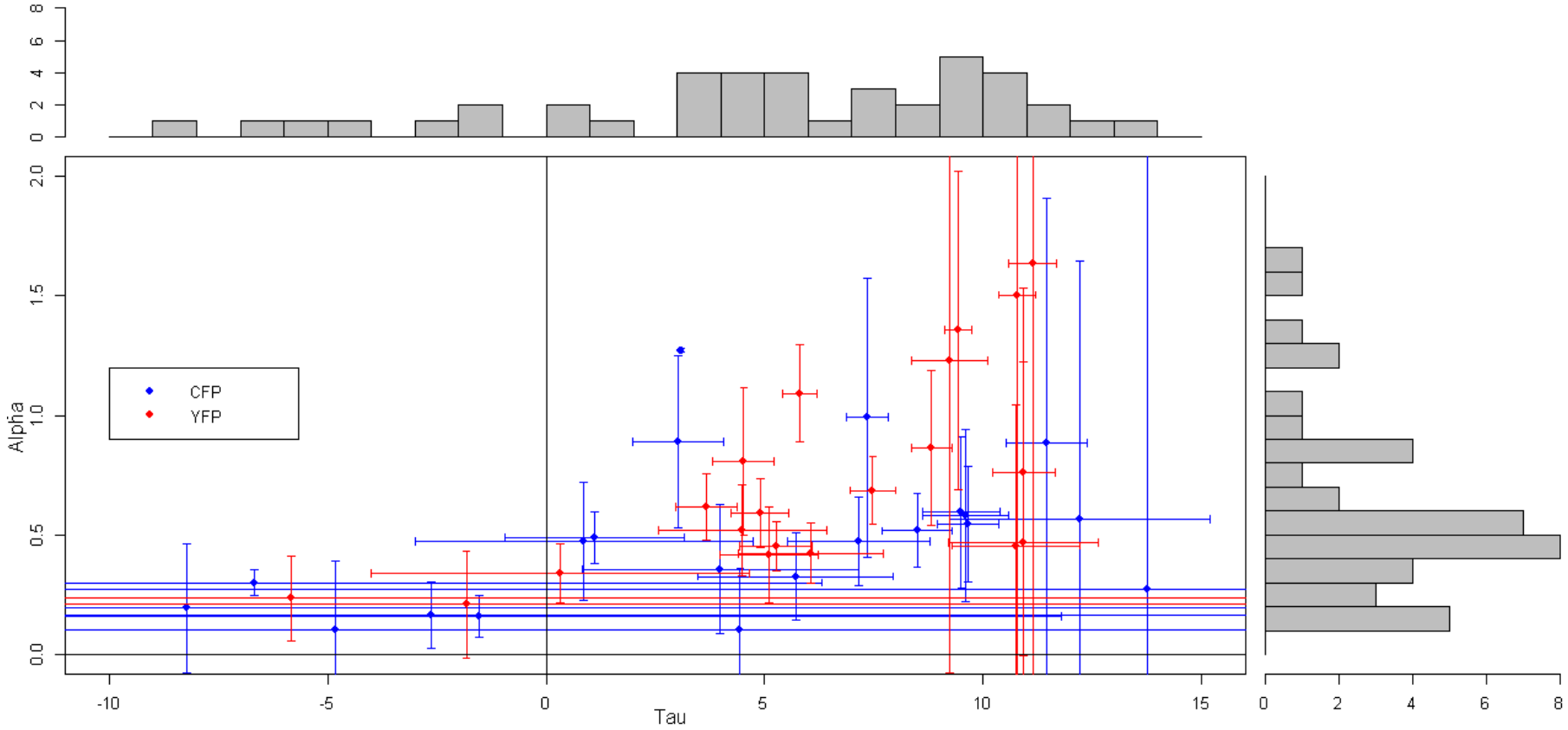


Figure 4.13 Distribution and relationship of the Dirac curve parameters for those fluorophore ratio curves that could be fit from a neutral drift model
The model was run 50 times with a bottleneck size of 100, carrying capacity $3x10^7$, and division probability 0.9.

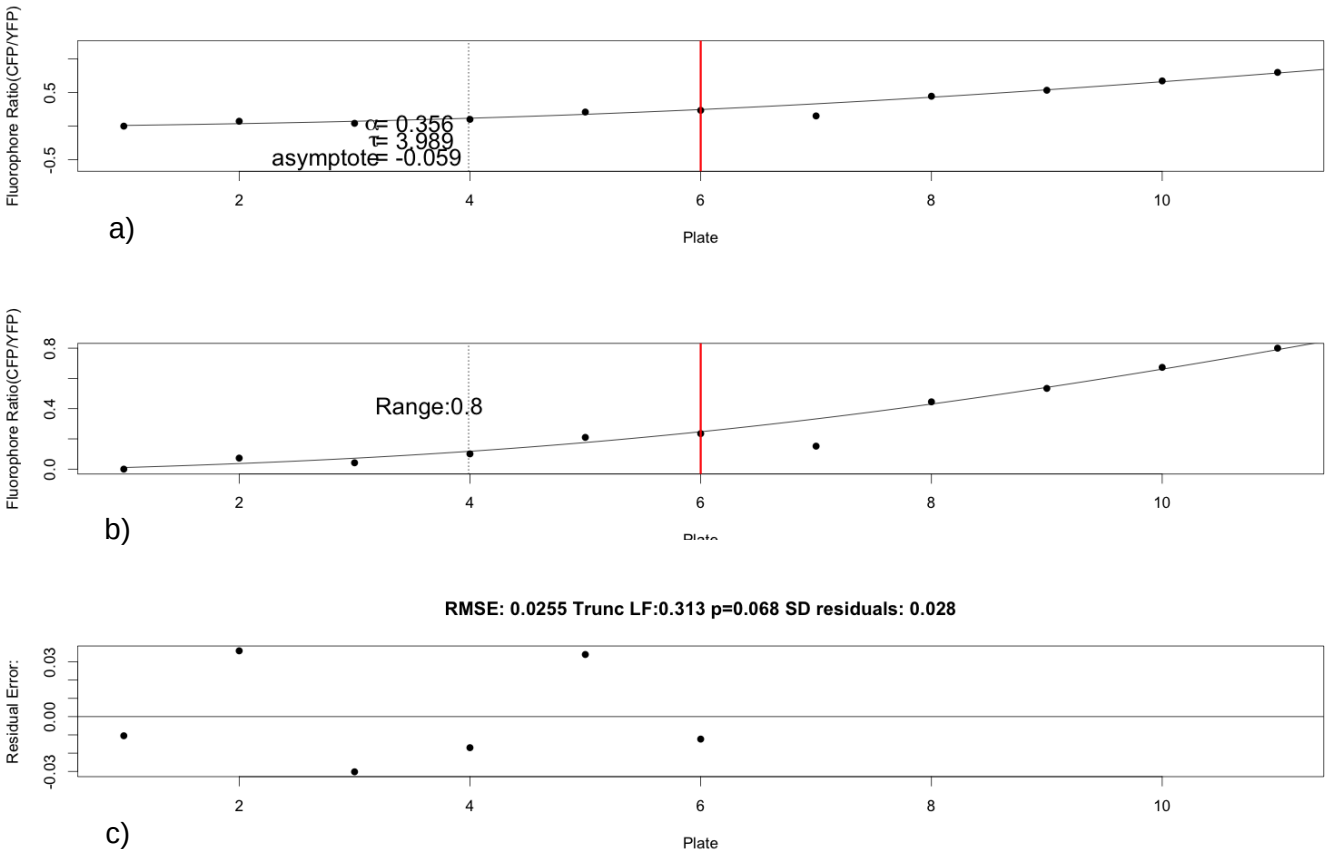


Figure 4.14 A simulated fluorophore ratio curve generated with a 100 cell bottleneck.
The carrying capacity was $3x10^7$ and a division probability 0.9. Part a) shows the fluorophore ratio with standard axes from -0.5 to 1 and the curve parameters, whilst b) magnifies the y axis to show minor variation of points with the data range. Part c) shows the distance of data points from the fitted curve.

Figure 4.15 shows the distribution and relationship of the Dirac fitness parameters with a bottleneck size of 2000 (carrying capacity $3x10^7$, division probability 0.9). The increase in bottleneck size has substantially reduced the range of the fluorophore data (see Figure 4.16 for a typical fitted curve). Relatively few of the models could be fit to the Dirac equation since most of the data has a range less than 0.12 (the maximum is 0.172). This suggests that increasing the minimum range for curve fitting would prevent drift being mistaken for a beneficial mutation where none had occurred. However, such drift may add significant error at this bottleneck size.

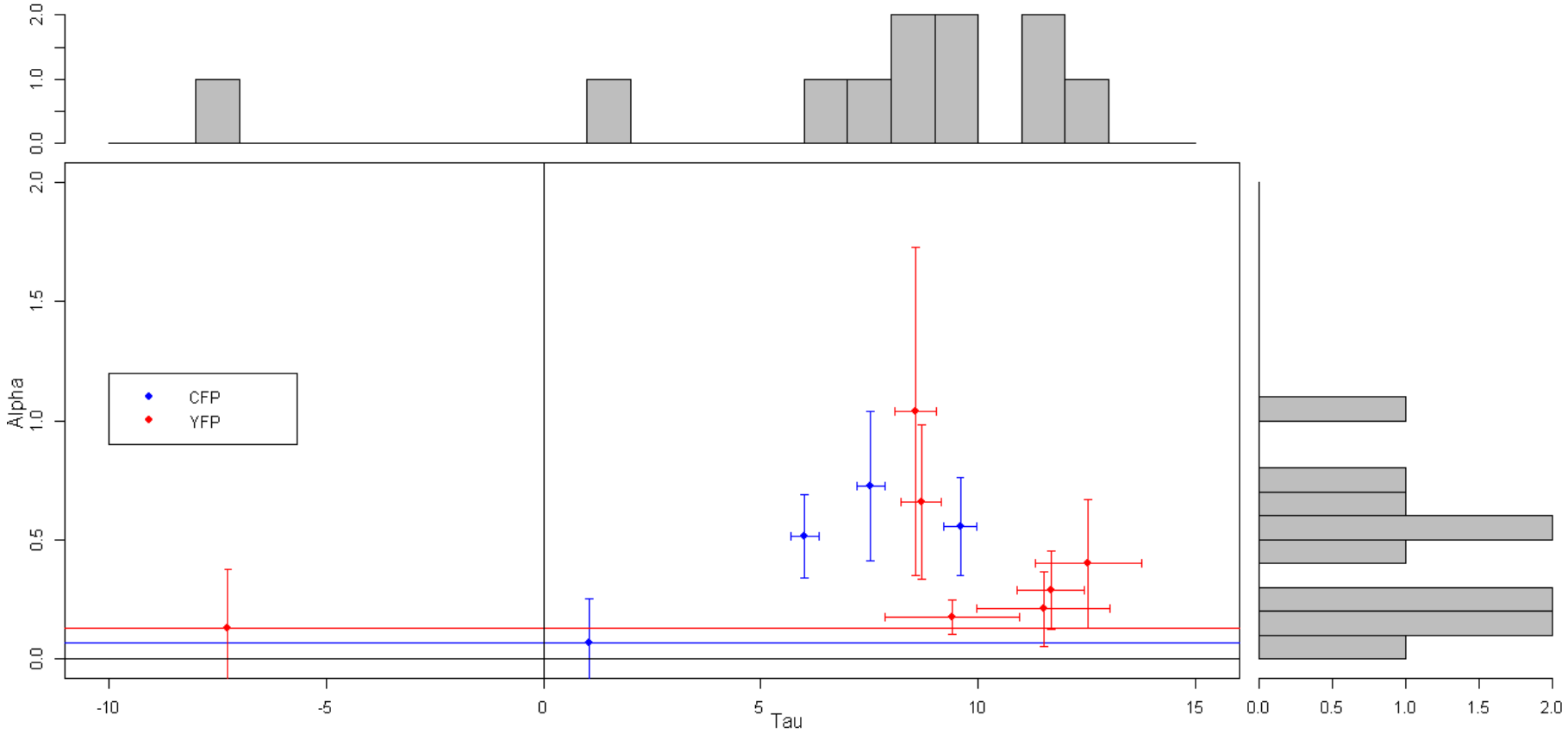


Figure 4.15 The Dirac curve parameters for fluorophore ratio curves that could be fit from a neutral drift model run 50 times with a bottleneck size of 2000
The carrying capacity was $3x10^7$ and division probability 0.9. The alpha fitness parameter plotted against the tau onset parameter.

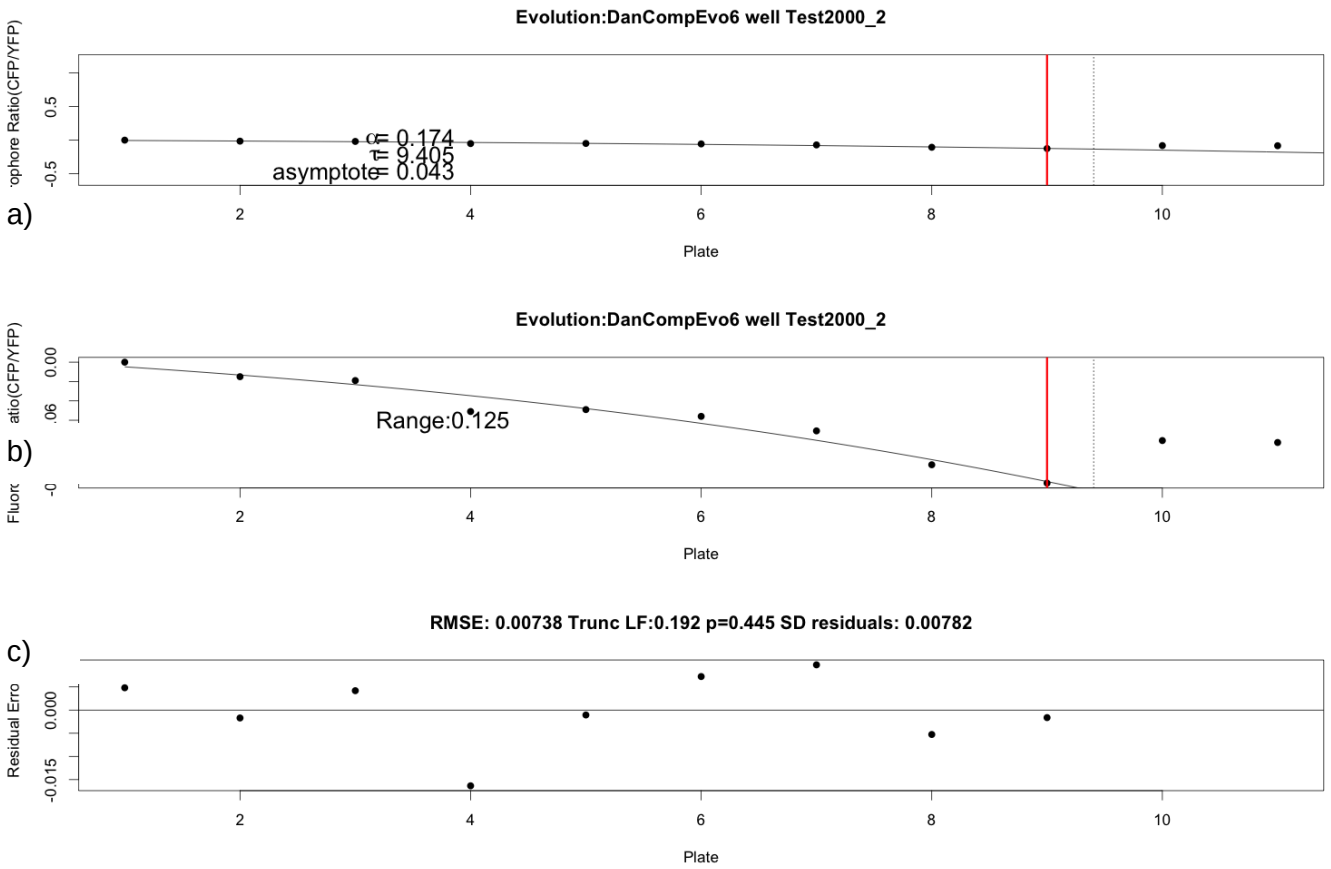


Figure 4.16 A simulated fluorophore ratio curve generated with a 2000 cell bottleneck.
The carrying capacity was $3x10^7$ and a division probability 0.9. Part a) shows the fluorophore ratio with standard axes from -0.5 to 1 and the curve parameters, whilst b) magnifies the y axis to show minor variation of points with the data range. Part c) shows the distance of data points from the fitted curve

Figure 4.17 shows typical fluorophore ratio data generated by neutral drift with a bottleneck size of 10,000 (carrying capacity $3x10^7$, division probability 0.9). Part a) on standard axes shows that there is no significant change in the fluorophore ratio. However, part b) shows that there is trivial variation in the fluorophore ratio which could be fit with data truncation without a minimum range. The model data output for a bottleneck size of 15,000 cells was very similar.

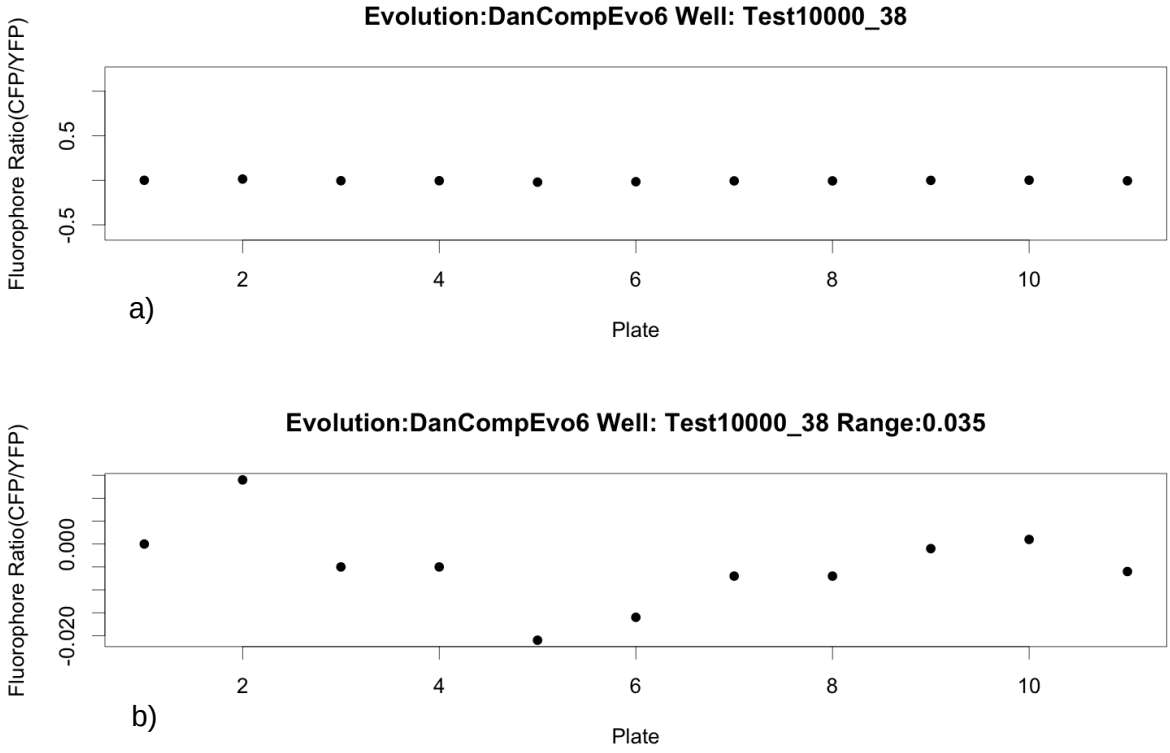


Figure 4.17 A simulated fluorophore ratio curve generated with a 10,000 cell bottleneck.
The carrying capacity was $3x10^7$ and division probability 0.9. a) shows the fluorophore ratio with standard axes from -0.5 to 1, whilst b) magnifies the y axis to show minor variation of points

4.3.4. Experimental data curve overview

If the starting strains are isogenic apart from their fluorophores which exert the same fitness cost, then it should be equally likely for a beneficial mutation to arise in either strain. Table 4.3 is a summary of the results of NLS fitting to the Dirac equation for each evolution series, showing the number of wells that were fitted and the direction of fluorophore ratio change. These counts were subjected to a two-tailed binomial test for departures from equal likelihood.

Evolution	Num. no fit	Num. Fitted	CFP+ fit	YFP+ fit	P value, binomial test
6	10	38	24	14	0.14
8 (0% Ethanol)	5	11	3	8	0.23
8 (4.5% Ethanol)	3	13	6	7	1.00
8 (6.5% Ethanol)	3	13	0	13	2.4×10^{-4}
10	9	40	17	22	0.52

Table 4.3 A summary of results for the NLS fitting of the Dirac equation to the fluorophore ratio changes in evolutions 6, 8 and 10.
The table shows the number of wells that could be fitted and the mean range of the curves, which fluorophore increased and whether a two-sided binomial test of the ratio of fluorophore increases showed a significant departure from the null model.

Only the wells in Evolution 8 that were subjected to 6.5% ethanol showed a significant departure from equality in prevailing fluorophore counts. All the fitted populations showed a decrease in CFP and an increase in YFP ($p=2.4 \times 10^{-4}$). The other treatments in evolution 8 with fitted curves did not show significant departures from chance ($p=0.23$ and 1.00 for the 0% ethanol and 4.5% ethanol treatments respectively). In the other evolutions, the ratio of fitted fluorophore increases in evolution 6 was closest to significance ($p=0.14$) but the treatment in this evolution, a 4.5% fixed ethanol stress, was the same as that in 13 fitted wells in evolution 8 in which $p=1$. Evolution 10 also showed no significant bias in fluorophore increase ($p=0.52$). Accordingly, there is no evidence for a bias favouring either fluorophore except in the evolution 8, 6.5% ethanol treatment. From this, it seems most likely that there was a beneficial mutation in the overnight YFP culture subjected to 6.5% ethanol, and this mutation had swept to a proportion of the population that was

unlikely to be stochastically lost when the overnight culture was diluted and divided into the wells of the evolution 8 starting plate.

4.3.5. Fitted and unfitted curves

The population fluorophore ratio datasets that could be fit to the Dirac equation had a larger range of fluorophore ratio change than the unfitted e.g. fitted evolution 6 populations had ranges with a mean of 0.55 (max=1.07, min=0.16) whilst the unfitted had ranges with a mean of 0.13 (max=0.289, min=0.064). Of the 38 wells in evolution 6 that were fitted, 14, just over one third, were truncated to improve the fit; 7 of those had one point removed and the other seven were truncated by two or three points.

Figure 4.18 a) and b) show the NLS fit to the Dirac equation for the evolution 6 well B07. Figure a) shows the curve with axes which encompass all variation on the plate whilst figure b) allows the axes to vary with the data. Both figures show points that closely follow the curve. Figure c) shows the residual error, the distance of the points from the fitted line; the deviations appear evenly distributed in sign and magnitude with little obvious sign of autocorrelation and the Lilliefors test (0.178,p=0.493) shows no significant departure from normality.

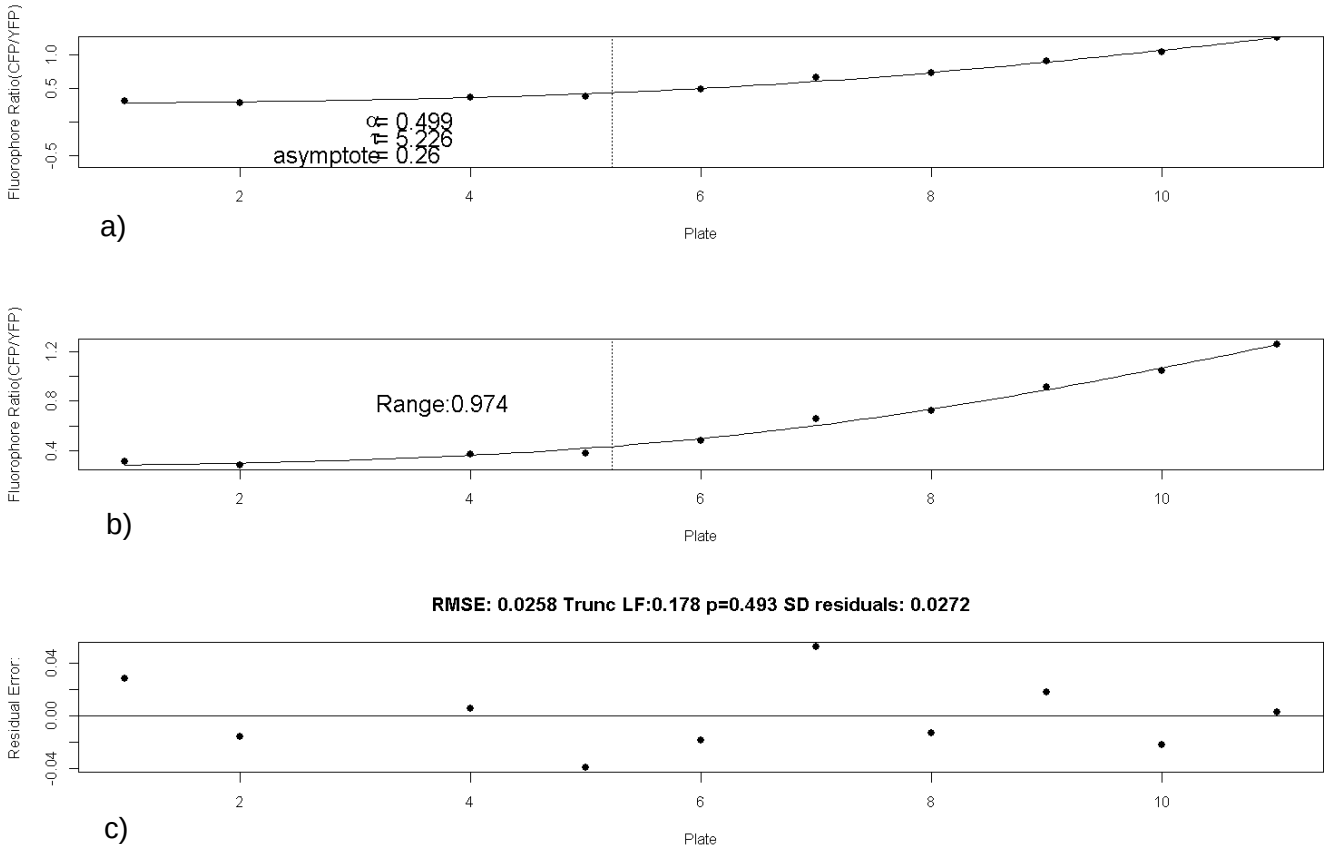


Figure 4.18 Fluorophore ratio curves (log10(CFP/YFP) against plate) for evolution 6 well B07.
Figure a) shows the curves with a y axis restricted to a range of -0.5 to 1 for comparison with populations in other wells. Figure b) shows the data with the same x-axis limits but with the y axis expanded to maximise detail. Figure c) shows the root mean squared error for the NLS fit to the Dirac equation. The range of fluorophore ratio change is 0.974 and the value of the Lilliefors test statistic is 0.178, with p=0.493.

Figure 4.19 shows the fluorophore ratio data for well B03 in evolution 6, a typical dataset that could not be fitted using NLS with the Dirac equation. The range is much smaller (0.149), Figure 4.19 a) shows data close to a straight line with no slope, while Figure 4.19 b) shows much of the range in the data due to rare discrete jumps rather than incremental increases which follow a curve. This data appears consistent with neutral drift rather than the sweep of a beneficial mutation.

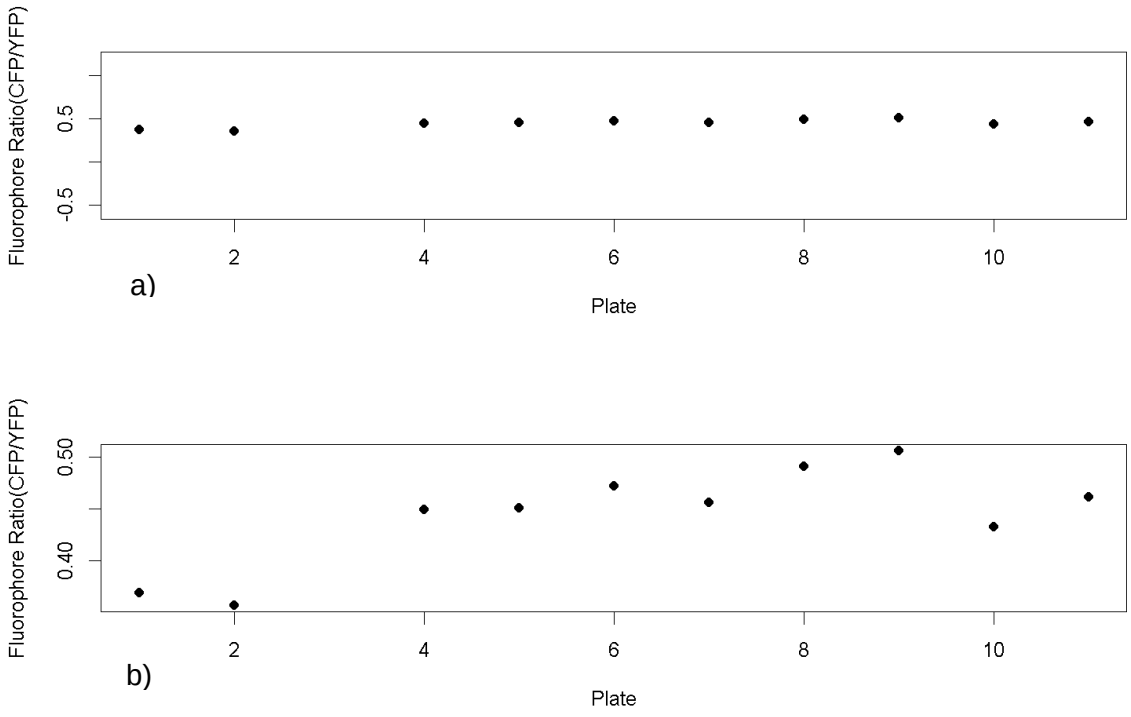


Figure 4.19 Fluorophore ratio curves (log10(CFP/YFP) against plate) for evolution 6 well B03.
Figure a) shows the curves with a y axis restricted to a range of -0.5 to 1 for comparison with populations in other wells. Figure b) shows the data with the y axis expanded. The range of the data is 0.149.

4.3.6. Evolution 6

The data for Plate 3 in evolution 6 was consistently extremely low across all wells. Close parity between the points in plates 2 and 4, particularly in those wells without significant ratio change, suggest that this was some machine calibration discontinuity (although the calibration and other settings remained unaltered) rather than a biological problem. Accordingly, these data points were omitted for all analysis of evolution 6. No similarly aberrant data was found in the other evolution series.

Figure 4.20 shows the distribution of estimates of alpha and tau for Dirac fitted populations in evolution 6. This figure clearly shows a feature of the curve estimates for every evolution; curves that diverge early in the evolution series tend to have small errors in the estimate of alpha but large errors in their estimates of tau, while late diverging curves have the converse, large alpha errors but small tau errors. If there is little waiting time before the fluorophore ratio diverges then it is difficult to estimate the time of divergence but there are lots of data points to analyse the changing slope of the curve to estimate alpha. Conversely if the waiting time is long then tau can be estimated more accurately but there are fewer data points to estimate alpha.

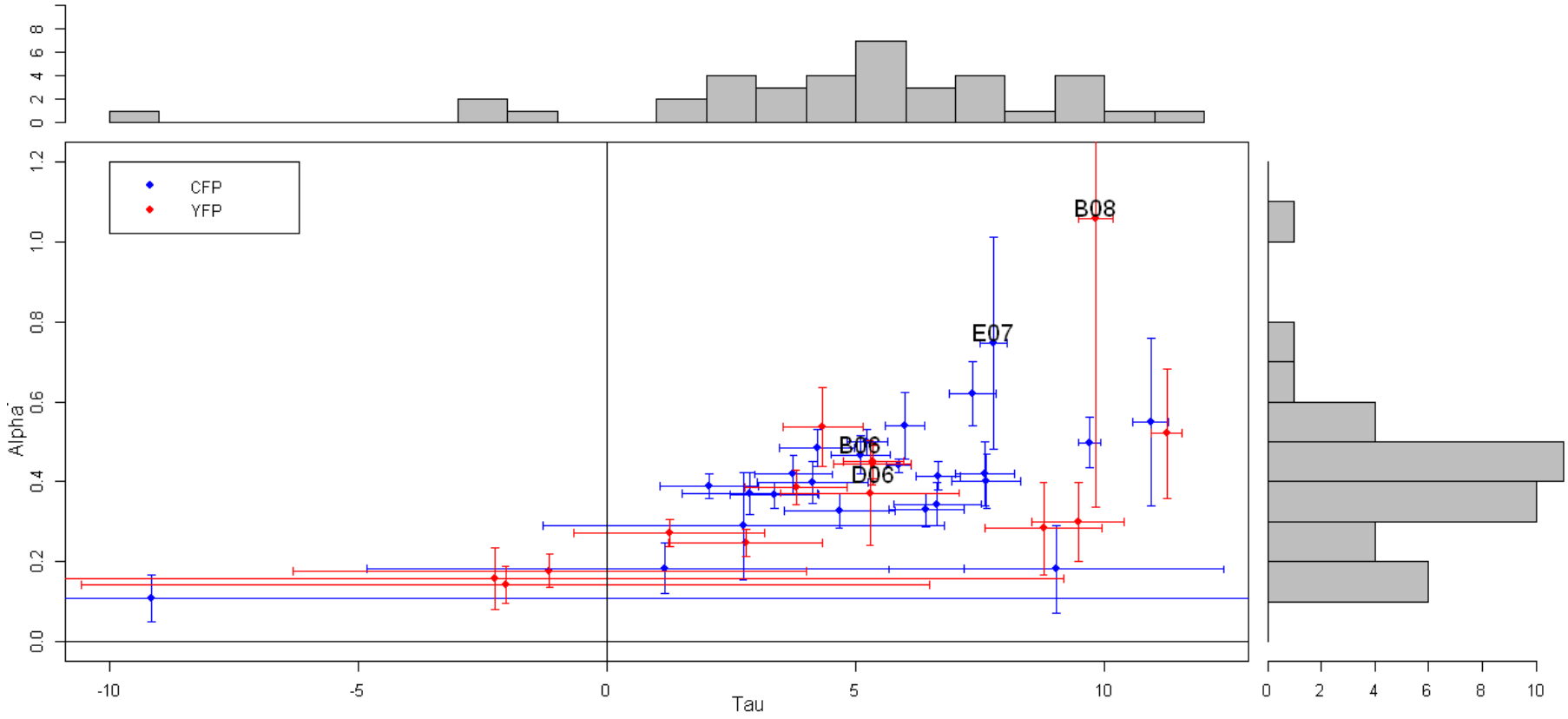


Figure 4.20 The alpha (fitness) parameter is plotted against tau (onset) for those well populations of Evolution 6 that were fit to the Dirac equation using NLS.
Populations with increasing CFP are shown in blue and those with increasing YFP are shown in red. Error bars are standard errors. The well identifiers of some populations are shown, with labels above all points, apart from population D06 which is labelled below. A colony was sequenced from the labelled populations in well B06 (CFP) and D06 (YFP). Populations E07 and B08 are mentioned in the text.

As is common for zero bounded parameters, the histogram of alpha values for evolution 6 is positively skewed (skew=1.22 using function ‘skewness’ type 3 in R package e1071 (Joanes & Gill 1998)). A positive skew was also observed in the alpha estimates due to neutral drift (see Figure 4.13). More significantly, the correlation of alpha and tau values is notably ‘wedge shaped’. Populations with values of alpha smaller than 0.4 (mean 0.398 +/- 0.029 s.e.) occur throughout the range of tau estimates, (see Figure 4.21 a) and several populations with alpha values less than 0.2 have negative tau estimates. However, those populations with alpha estimates larger than the mean do not have tau values less than 3.7 (see Figure 4.21 b). Large alpha values are rare and only appear later in the evolution series, though the largest, for wells E07 (alpha=0.747 +/- 0.266 s.e.) and B08 (alpha= 1.058 +/- 0.722 s.e.) appear so late in the experiment that their alpha values have large error values which make their true magnitude extremely uncertain.

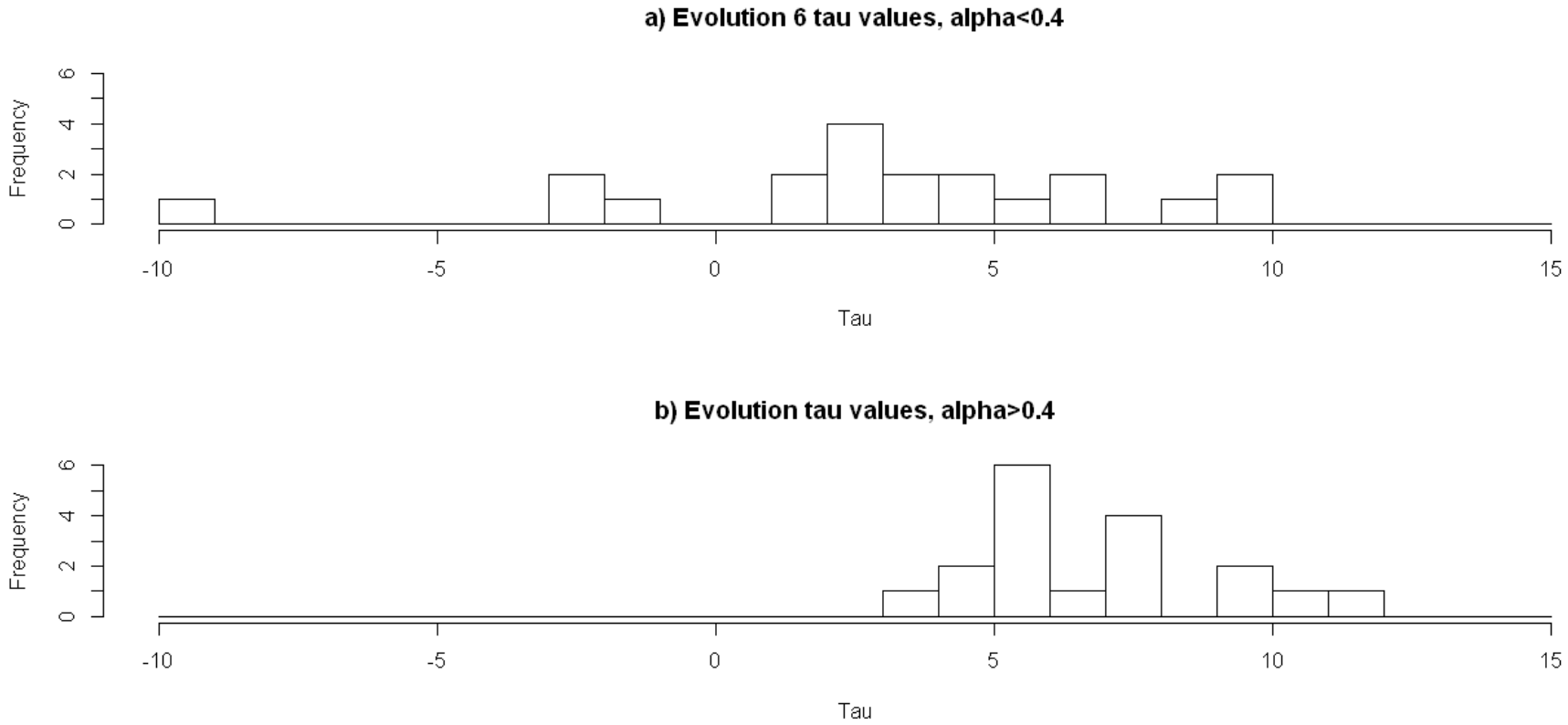


Figure 4.21 Histogram of estimates of tau for evolution 6 populations with Dirac fitted fluorophore curves.
Figure a) shows the tau distribution for populations with alpha values less than 0.4 and Figure b) shows the tau distribution for alpha values greater than 0.4.

According to the estimates of bottleneck and final population sizes, in each plate the population increases from $2x10^3$ to $3x10^7$ cells i.e. 2^{11} to 2^{25} cells which is approximately 14 generations. The mean tau of 4.835 (+/-0.656) suggests that typical beneficial mutations are sweeping though populations by ~5x14 or ~70 generations.

Neutral drift modelling with a 2000 cell bottleneck used in evolution 6 suggests that fitted curves with the same Dirac parameters as seen in evolution 6 can arise in a minority of wells due to drift (see Figure 4.15). The ranges of fluorophore ratio curves in evolution 6 were larger than the model output (no fitted curve had a range less than 0.2) and the diploid nature of the two sequenced strains suggests that at least two beneficial sweeps have occurred which suggests that both beneficial mutation and drift have given rise to the fluorophore ratio curves. Curves with alpha values less than 0.2, and tau values indicative of an event prior to the first plate are particularly likely to be due to drift. Although there might be some instances in which beneficial mutations of small effect present in the originating colonies or overnight cultures might sweep in a small population rather than a larger one.

Modelling further suggests (see Figure 4.13) that if the true value of the bottleneck in evolution 6 had been an order of magnitude less, at 100 cells or fewer, then the result would have been much greater fluorophore ratio deviations than those seen, with very few populations showing no deviation from the initial value. In contrast, in the experimental data, 10 wells show very little deviation from the initial ratio (see Figure 4.22, ratio data for well G04 which is typical).

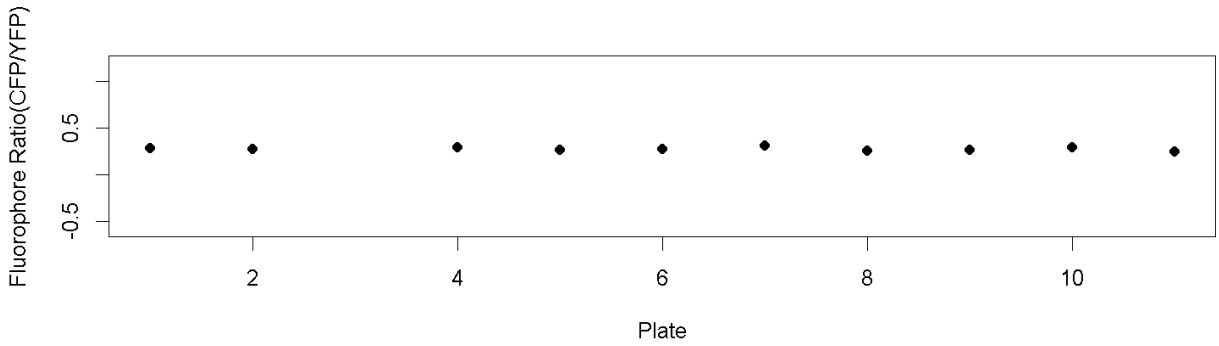


Figure 4.22 Plate versus fluorophore ratio (log10(CFP/YFP) against plate) for evolution 6 well G04
The standard y axis (-0.5 to 1) shows no change in fluorophore ratio that can be fitted with the Dirac equation.

A two-sided Kolmogorov-Smirnov test (D=0.3333, p=0.2235) shows no significant difference between the alpha values of fitted curves which deviate in favour of either CFP or YFP. Figure 4.23 is a histogram of alpha values broken down by the

fluorophore prevailing in the competition, which show very similar distributions of alpha values for each fluorophore.

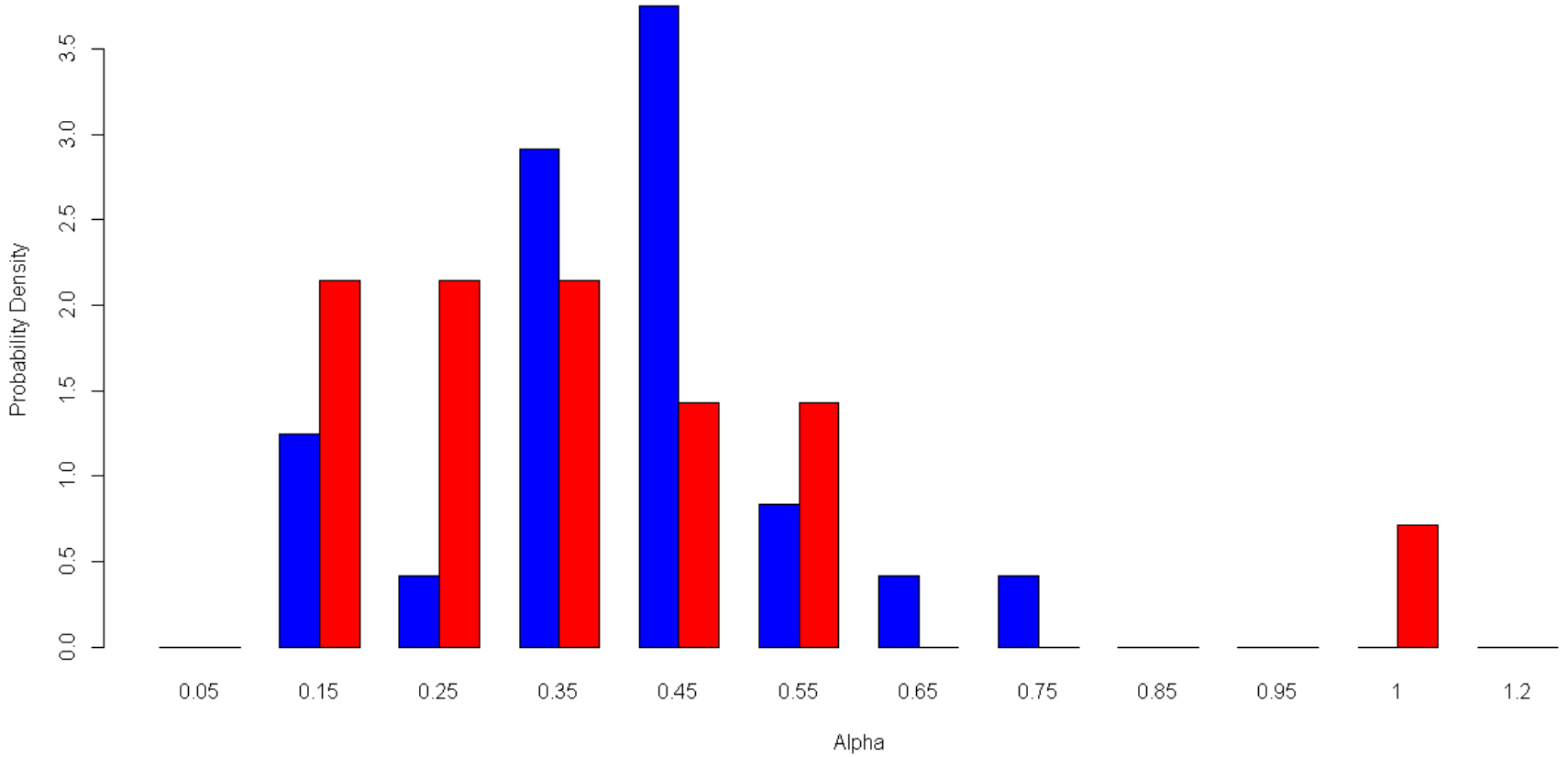


Figure 4.23 A histogram of alpha (fitness) estimates for fitted fluorophore ratio curves in evolution 6, split by the fluorophore prevailing in competition.
The fitness distributions of the two fluorophores overlap.

4.3.7. Evolution 6 populations with sequenced strains

A CFP strain was sequenced from well B06 (Evo6_11_B06_CFPe_4) and a YFP strain from well D06 (Evo6_11_D06_YFPe_4). The populations that these strains were isolated from show fluorophore ratio changes with typical alpha and tau values that are close to each other and the mean of alpha and tau estimates for all fitted evolution 6 populations (see Table 4.4).

Well	Fluorophore	Alpha (+/-s.e)	Tau (+/- s.e.)
Means	-	0.398 (+/-0.029)	4.835 (+/-0.656)
B06	CFP	0.467 (+/-0.049)	5.096 (+/-0.605)
D06	YFP	0.444 (+/-0.052)	5.333 (+/-0.771)

Table 4.4 Estimates of alpha and tau for populations in wells B06 and D06 in evolution 6
The table shows the mean estimates of alpha and tau with standard errors for the fitted curves in evolution 6.The population in well B06 showed a deflection in favour of increasing CFP while D06 showed increasing YFP.

Figure 4.24 shows a plot of the fitted fluorophore curve for population B06 and Figure 4.25 for population D06.

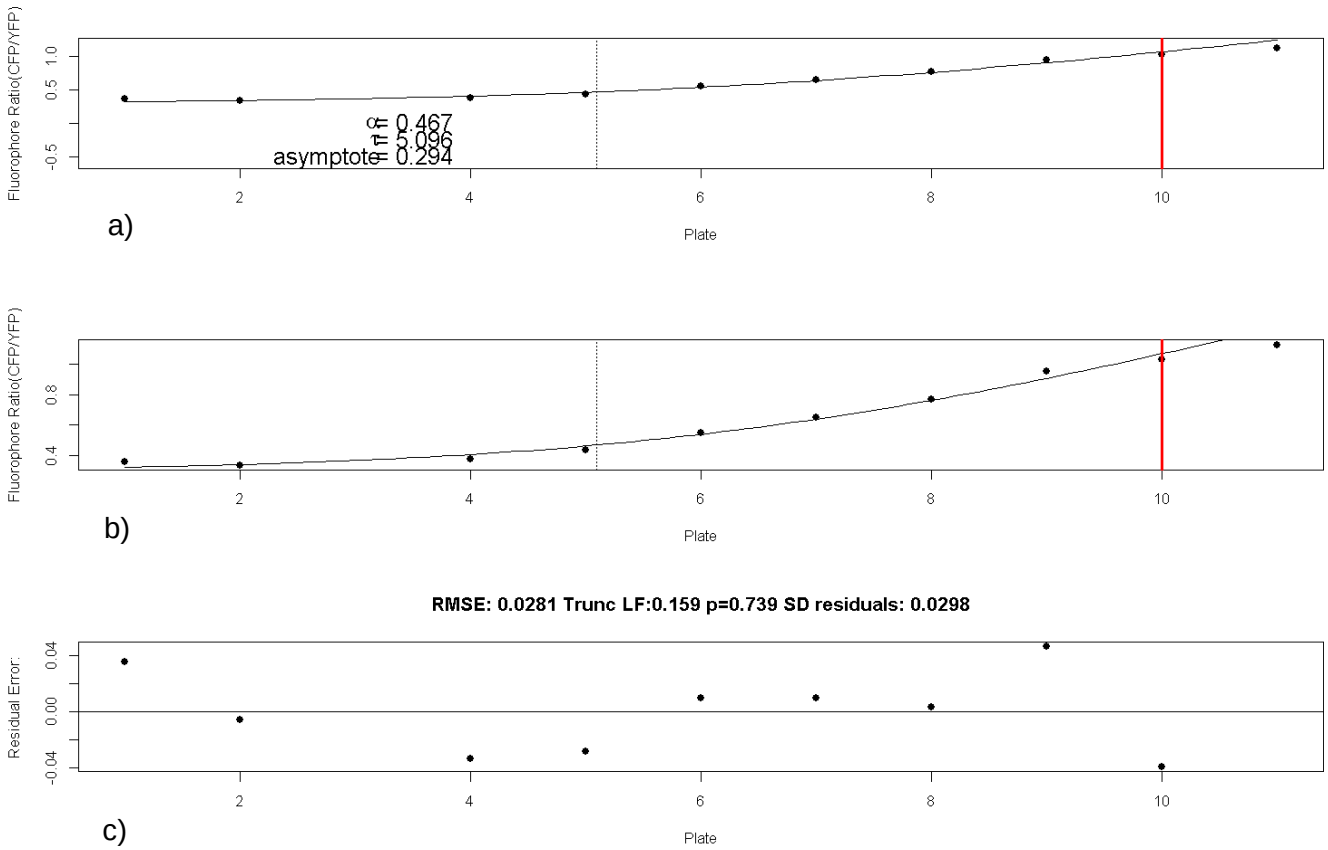


Figure 4.24 A plot of the log fluorophore ratio of Evolution 6 well B06 showing the NLS fit to the Dirac equation.
Figure a) shows the curves with a y axis restricted to a range of -0.5 to 1 for comparison with populations in other wells. Figure b) shows the data with the same x-axis limits but with the y axis expanded to maximise detail. Figure c) shows the root mean squared error for the NLS fit to the Dirac equation. The range of ratio change is 0.79, the α (fitness) estimate is 0.467(+/-0.049s.e.), tau (onset) 5.096 (+/-0.605s.e.), and asymptote = 0.294.

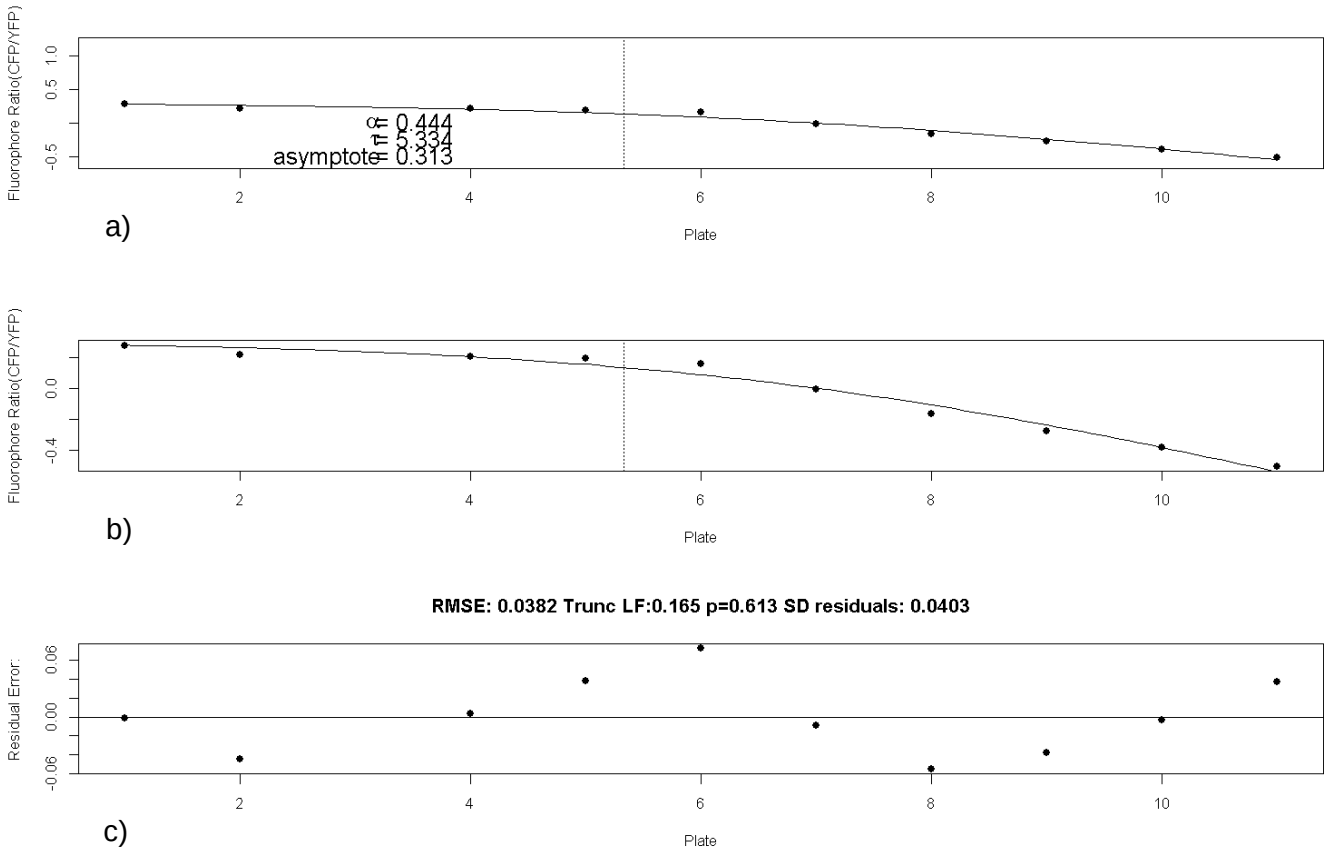


Figure 4.25 A plot of the log fluorophore ratio of Evolution 6 well D06 showing the nls fit to the Dirac equation.
Figure a) shows the curves with a y axis restricted to a range of -0.5 to 1 for comparison with populations in other wells. Figure b) shows the data with the same x-axis limits but with the y axis expanded to maximise detail. Figure c) shows the root mean squared error for the NLS fit to the Dirac equation. The range of ratio change is 0.788, the α (fitness) estimate is 0.444 (+/-0.052s.e.), tau (onset) 5.333 (+/-0.771s.e.) and asymptote = 0.313. c) shows that the errors are not well distributed.

No changes in coding sequence relative to their ancestors were detected in either of the sequenced strains in this evolution (see 6.3.4) though copy number variation, cell size measurement, and fluorescence cytometry, strongly suggest that they were both diploid.

4.3.8. Evolution 8

Figure 4.26 shows the alpha (fitness) and tau (origin) parameters for fluorophore curves fitted to the Dirac equation in evolution 8.

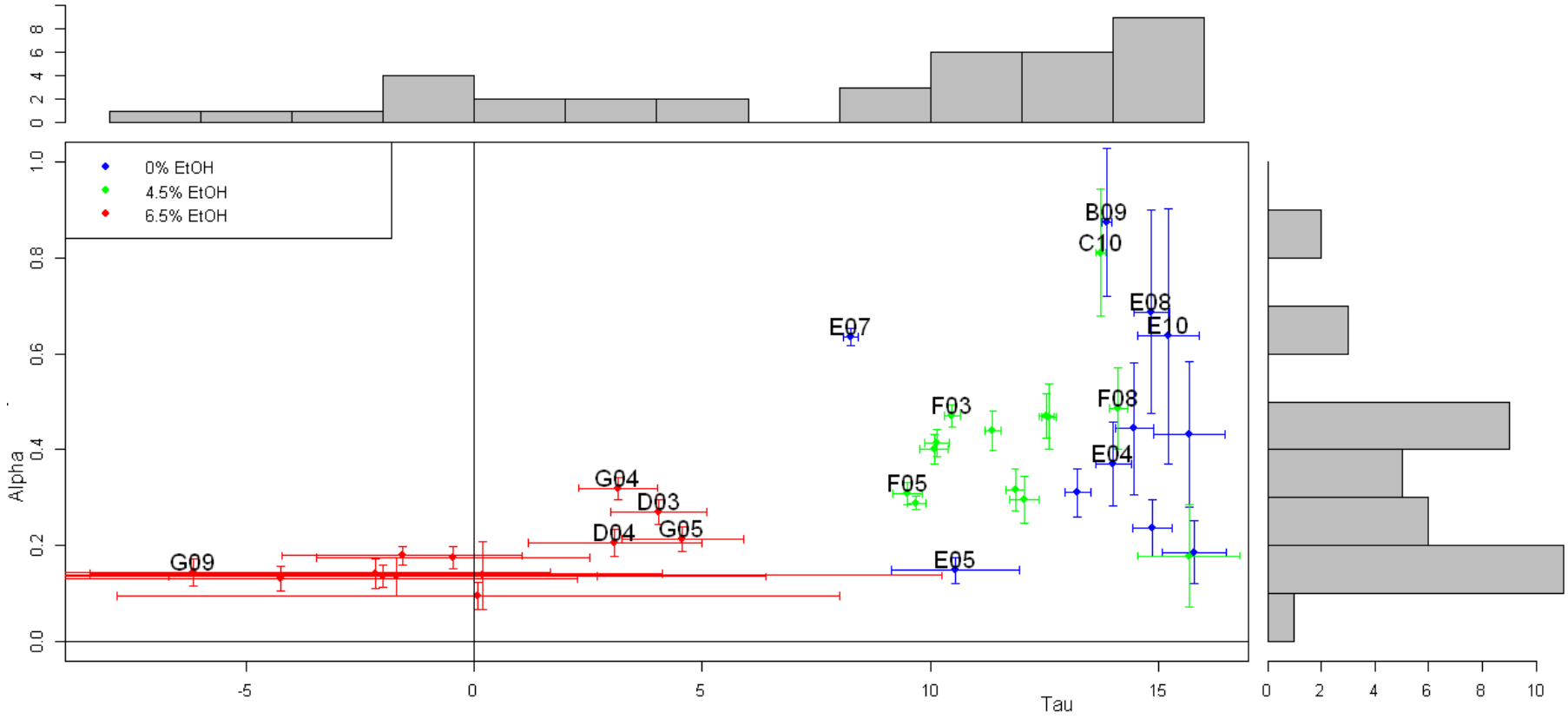


Figure 4.26 Alpha vs tau parameters for Dirac fitted fluorophore ratio change in evolution 8.
Points discussed in the text have labels above them. Error bars are 95% confidence intervals. Points and error bars in red in the central scatter plot are for wells with a 6.5% fixed ethanol stress, points in green a 4.5% fixed stress and points in blue were had no added ethanol.

Table 4.5 shows the means and standard errors for the alpha and tau estimates for evolution 8.

Treatment	Mean Alpha (s.e.)	Mean Tau (s.e.)
0	0.451 (+/- 0.070)	13.706 (+/- 0.699)
4.5%	0.412 (+/- 0.042)	11.833 (+/- 0.523)
6.5%	0.176 (+/- 0.018)	-0.245 (+/- 0.898)

Table 4.5 Means and standard errors for estimates of alpha and tau from NLS Dirac fitted fluorophore ratio curves in evolution 8.
The table splits the data by ethanol treatment in three classes: 0% (control), 4.5% and 6.5%.

The points in red on the central scatterplot of Figure 4.26 are for the 6.5% fixed ethanol treatment. Most of the populations in this treatment have an alpha fitness value less than 0.2 (mean =0.176 +/- 0.018, see Table 4.5) and an estimate of tau that is very close to zero or less with large error bars (mean -0.235 +/- 0.898). The minor exceptions are four populations in wells D03, D04, G04 and G05 which have slightly higher alpha values and positive, though early, estimates for tau (the D03 and G04 populations were used to isolate sequenced YFP strains, see section 5.3.2).

Modelling of the 10,000 cell bottleneck size for evolution 8 suggests that no fluorophore ratio curves would have arisen due to drift that could have been fitted with the Dirac equation (see Table 4.2). This is also true for a 5,000 cell bottleneck, giving a reasonable margin for error due to serial dilution, cell mortality and changing ploidy partitioning biomass into fewer cells. However, the 6.5% ethanol treatment might be expected to have the largest cell mortality during stationary phase due to the ethanol stress imposed, and fluorescence cytometry suggests the 6.5% ethanol treatment populations appear to make a faster transition to diploidy (see chapter 6, Figure 6.24). Null modelling suggests that curves due to drift can be fitted if bottleneck falls to ~2,000 cells, a five-fold reduction from the target value. However, 13 out of the 16 of the 6.5% ethanol treatment populations were capable of being fit with the Dirac equation, all with YFP proliferating and most with very similar alpha fitness values less than 0.2. Furthermore, the sequenced strains from two different wells in this treatment (Evo8_14_D03_YFPe_1 and Evo8_14_G04_YFPe_1) were diploid and had the same mitochondrial SNP (at M.41413, a transversion from T to G). This suggests that a population with the mitochondrial SNP had arisen in the 6.5% YFP overnight culture, or the colony it was derived from, that was then distributed to all the wells in this treatment and swept from rare in at least two wells.

The points and error bars in green in Figure 4.26 are Dirac parameter estimates for the evolution 8 populations subjected to a fixed ethanol stress of 4.5%. For this treatment, the alpha (mean 0.412 +/- 0.042) and tau (mean 11.833 +/-0.523s.e.) estimates are larger than the 6.5% treatment. Two-sided Kolmogorov-Smirnov (KS) tests of the parameters from the 4.5% and 6.5% treatment show both the alpha (D=0.8462, p=6.25e-05) and tau (D=1, p=1.923e-07) values are significantly different between the two treatments. The mean tau values for evolution 6 (4.835 +/-0.656 s.e.) are significantly smaller than that of the populations sharing its 4.5% ethanol populations in evolution 8 (11.833 +/- 0.523 s.e.) which is consistent with the larger bottleneck size for evolution 8. Both evolutions were run on the same machine but there are many other confounding factors that might influence each experiment, not least the potential for neutral drift in evolution 6 (see section 4.3.6).

The points in blue in Figure 4.26 are Dirac parameter estimates for the evolution 8 populations that were not subjected to an added ethanol stress. Both the alpha (mean 0.451 +/- 0.070 s.e.) and tau (mean 13.706 +/- 0.699 s.e.) estimates for this control are slightly larger than 4.5% treatment. KS tests show that while the alpha estimates for the control and 4.5% treatment do not significantly differ (D=0.287, p=0.607), the tau distributions do (D=0.587, p=0.020); despite the two early outliers: E05 and E07, the fluorophore ratios of the control diverge later than the 4.5% treatment.

4.3.9. Evolution 8 populations with sequenced strains

A strain from plate 14 of the 0% ethanol control, well E07 was sequenced; designated Evo8_14 E07_CFPe_4. This proved to be the only sequenced strain that had gained the ability to grow in minimal media deficient in histidine (Bharat Rash, personal communication). The *MATα* starting strains in the evolution experiments in this thesis were *his3Δ* which was conditionally complemented by a fusion between *STE2*, which codes for the alpha factor mating receptor fused to the *S. pombe HIS5* gene (which encodes an enzyme catalysing the same reaction as the *S. cerevisiae HIS3*, see Introduction section 2.6). Since the alpha factor mating receptor is strongly repressed in *MATα* strains, the acquisition of this phenotype strongly suggests that this diploid strain is likely to have arisen from a mating type switch followed by mating (see Discussion, section 4.4.1). The fluorophore ratio change of the E07 population had a large range with a good fit to the Dirac curve, though with likely autocorrelation of small scope in the residuals (see Figure 4.27). From the 4.5% ethanol treatment, a

strain from well E07 was sequenced, designated Evo8_14 F03_CFPe_2. For the fluorophore ratio change of the F03 population see Figure 4.28.

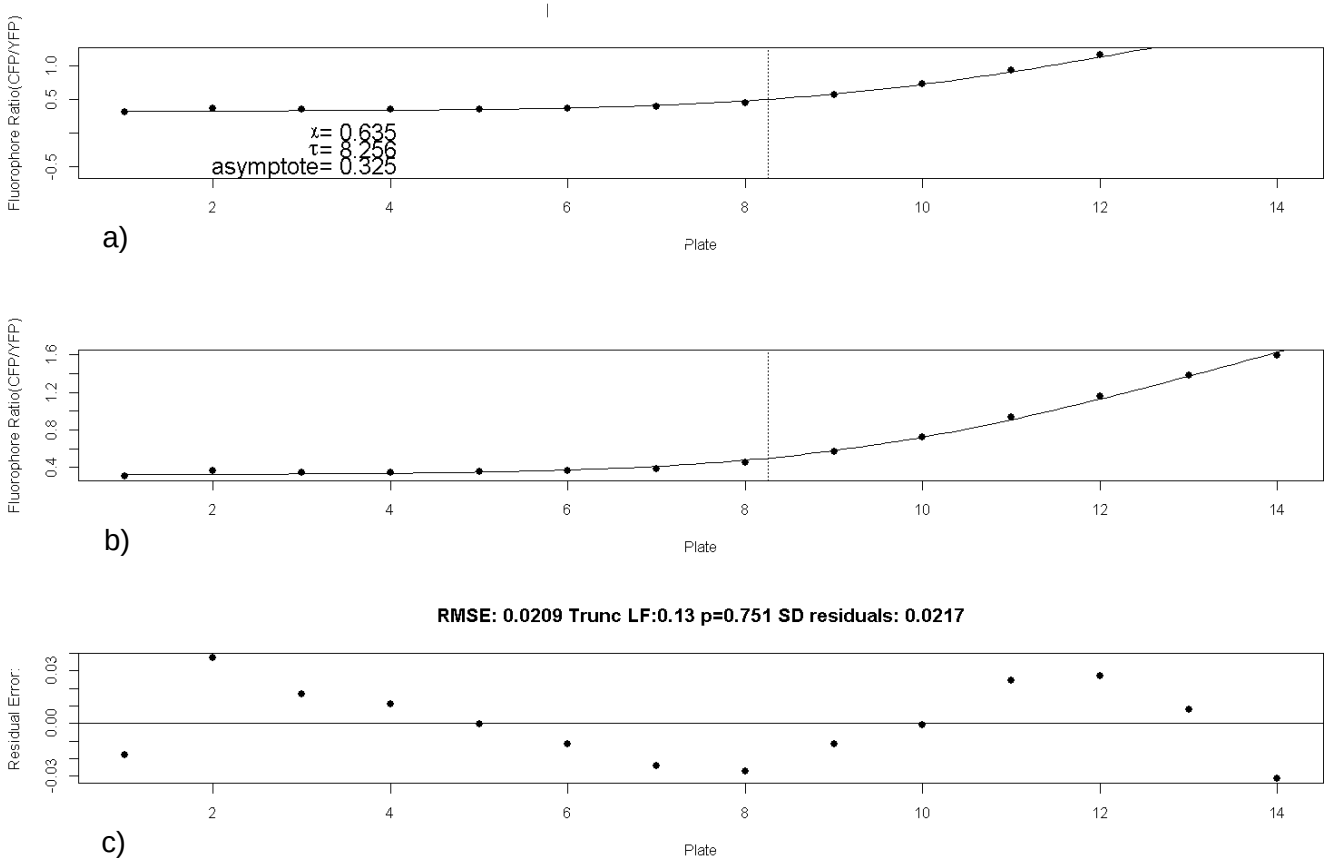


Figure 4.27 Plate versus fluorophore ratio curve log 10(CFP/YFP) for the evolution 8 well E07 population.
Part a) shows the fitted curve with standard y axis (-0.5 to 1), two points are out of range, b) shows the curve with y axis fitting points, c) shows the residual error of the fit. The alpha fitness value of the Dirac curve is 0.635, tau onset is 8.256. The data range is 1.29. The standard deviation of the residuals is 0.0217, the root mean squared error is 0.0209. The Lilliefors statistic is 0.13 (p=0.13) showing no significant departure from normality. However, Plot c) shows likely autocorrelation of the residuals.

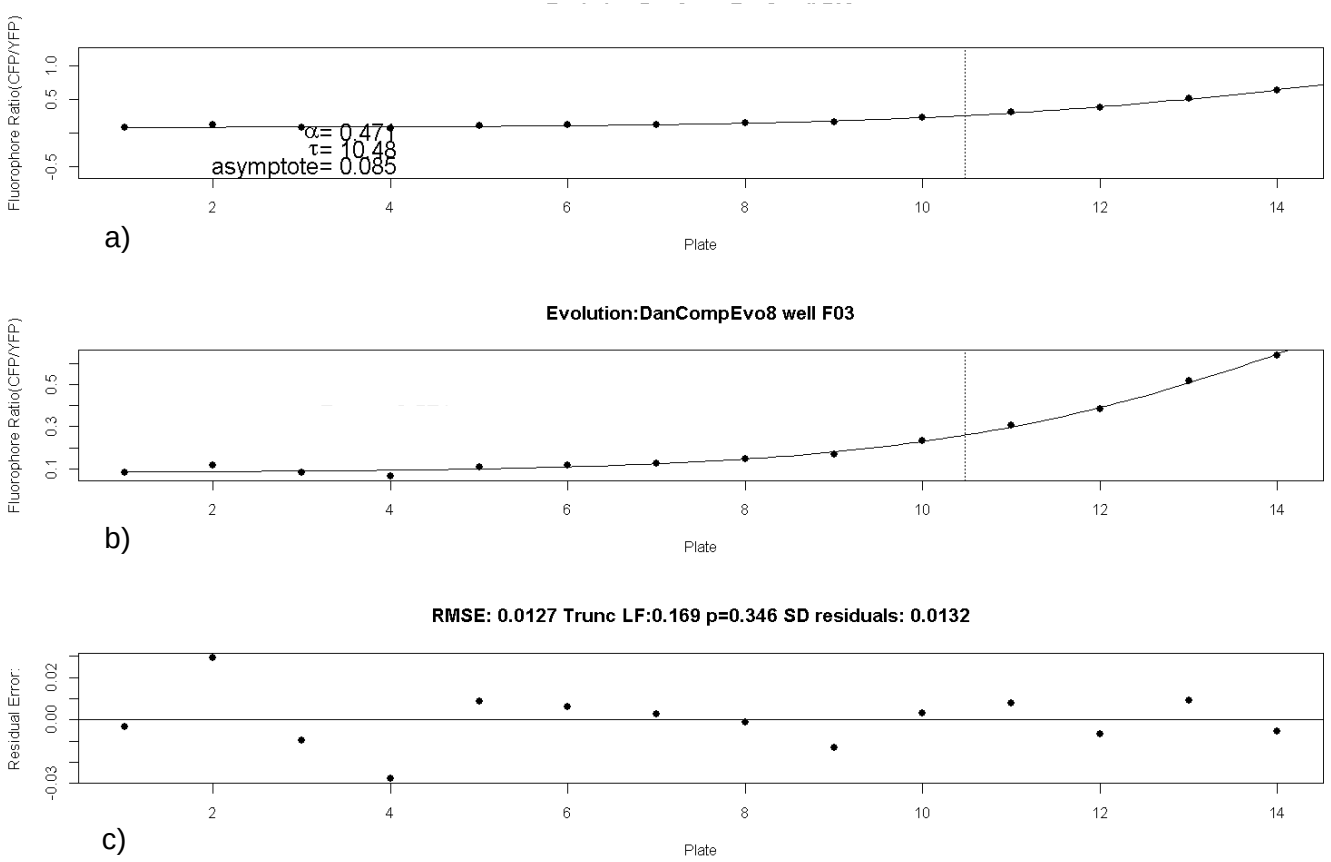


Figure 4.28 Plate versus fluorophore ratio curve (log10[final raw CFP fluorescence/final raw YFP fluorescence] for the evolution 8 well F03 population.
Part a) shows the fitted curve with standard y axis (-0.5 to 1), b) shows the curve with y axis fitting points, c) shows the residual error of the fit. The alpha fitness value of the Dirac curve is 0.471, tau onset is 10.48, asymptote=0.085. The data range is 0.571. The standard deviation of the residuals is 0.0132, the root mean squared error is 0.0127. The Lilliefors statistic is 0.169 (p=0.34) showing no significant departure from normality.

From the 6.5% ethanol treatment, two strains were sequenced, one from well D03, designated Evo8_14_D03_YFPe_1, and the other from G04 (Evo8_14_G04_YFPe_1). The populations from these wells showed very similar fluorophore ratio curves with YFP proliferating at the expense of CFP in both (see Figure 4.29 and Figure 4.30). Both curves were truncated by two points for the best Dirac curve fit and have similar low alpha fitness values: 0.27+/-0.026 s.e (D03), and 0.32+/-0.023 s.e. (G04) and similar values for tau 4.039+/-1.059 s.e. (D03) and 3.152+/-0.867 s.e.(G04). The similarity of the population curves is consistent with the identical mitochondrial SNP found in the sequenced strains (see section 6.3.2).

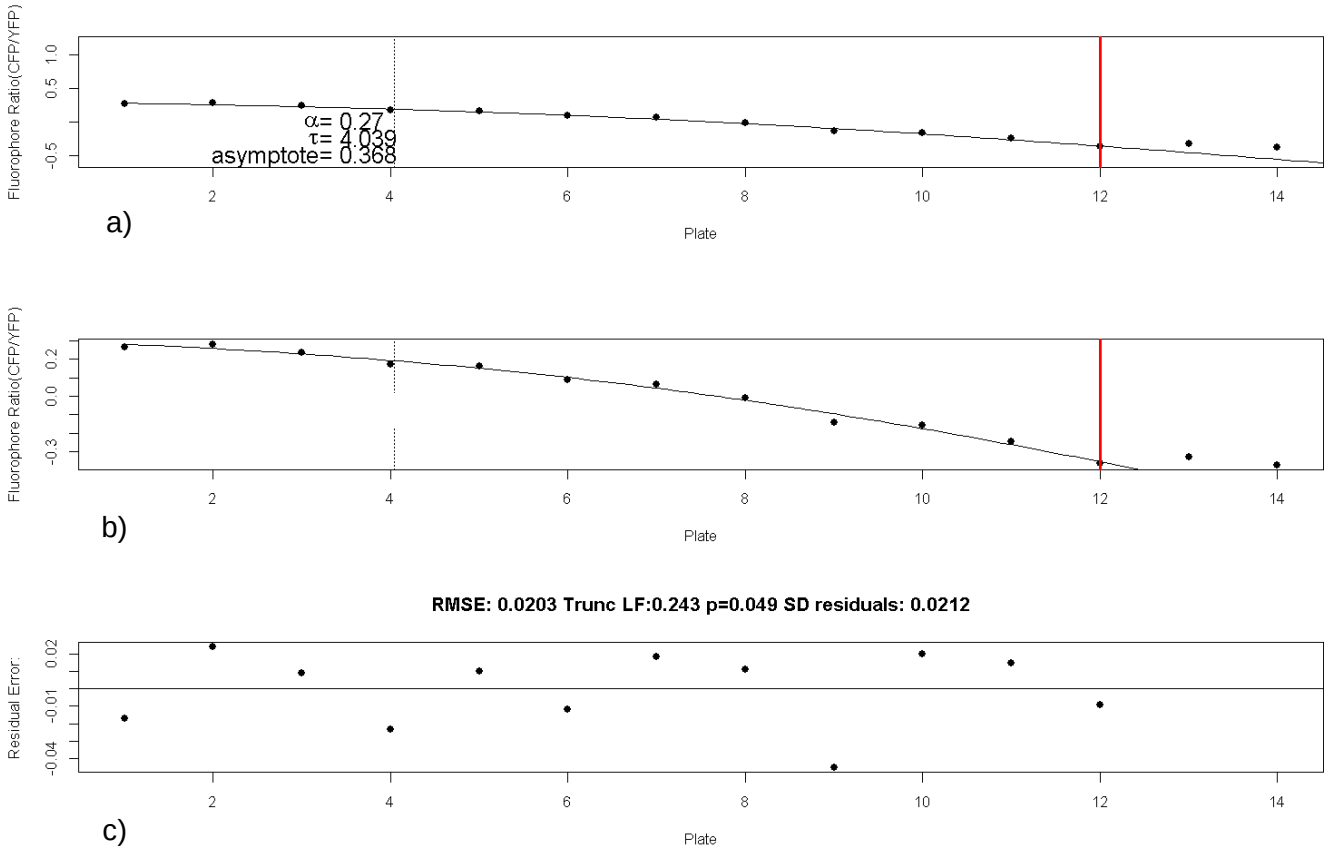


Figure 4.29 Plate versus fluorophore ratio curve (log10[final raw CFP fluorescence/final raw YFP fluorescence] for the evolution 8 well D03 population.
Part a) shows the fitted curve with standard y axis (-0.5 to 1), b) shows the curve with y axis fitting points, c) shows the residual error of the fit. The data was truncated by two points to best fit the curve. The alpha fitness value of the Dirac curve is 0.27, tau onset is 4.039 and asymptote is 0.368. The data range is 0.655. The standard deviation of the residuals is 0.0212; the root mean squared error is 0.0203. The Lilliefors statistic for the truncated data is 0.243 (p=0.049) indicating a departure from normality that is just significant.

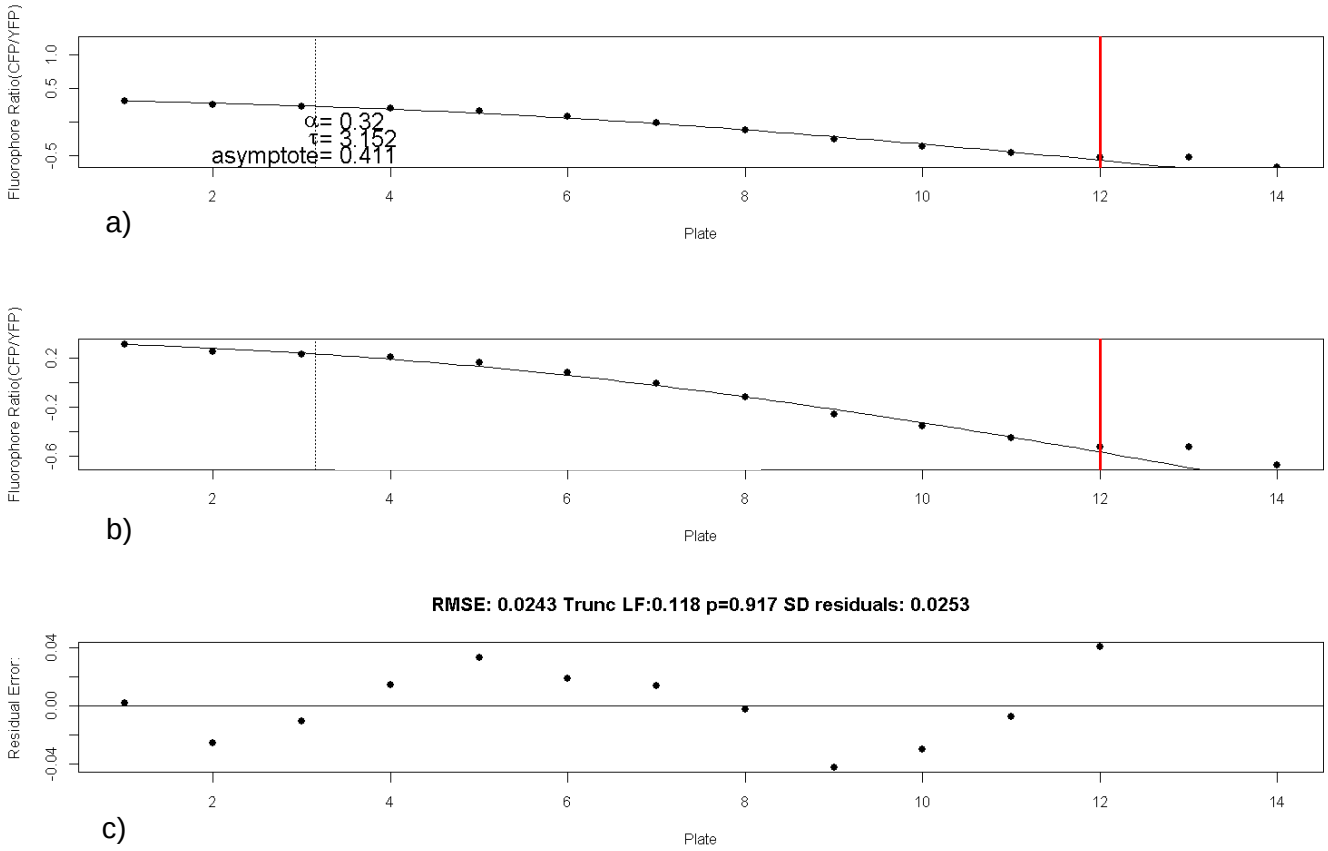


Figure 4.30 Plate versus fluorophore ratio curve (log10[final raw CFP fluorescence/final raw YFP fluorescence] for the evolution 8 well G04 population.
Part a) shows the fitted curve with standard y axis (-0.5 to 1), b) shows the curve with y axis fitting points, c) shows the residual error of the fit. The data was truncated by two points to best fit the curve. The alpha fitness value of the Dirac curve is 0.32, tau onset is 3.152, and asymptote is 0.411. The data range is 0.988. The standard deviation of the residuals is 0.0253, the root mean squared error is 0.0243. The Lilliefors statistic for the truncated data is 0.118 (p=0.917) indicating no departure from normality. However, Plot c) shows likely autocorrelation of the residuals.

4.3.10. Evolution 10 analysis

The Dirac equation assumes that conditions are constant, but the ethanol stress in evolution 10 increases. The fitness effects of beneficial mutations may alter in magnitude, or even sign, with varying ethanol and this would cause deviations of the data from the Dirac curve that may resemble (and indeed may be indistinguishable from) clonal interference. Out of 48 competition wells only 9 showed no fluorophore ratio change that could be fitted to the Dirac equation.

The evolution 10 bottleneck was 15,000 cells, a value that null modelling suggests is an order of magnitude larger than the value required for drift to produce curves with the range capable of being erroneously fitted to the Dirac equation. There is a corresponding increase in the estimate for tau, the time of onset of population fitness change, the mean value is 12.9 and only E05 has a tau estimate less than 9 (see Figure 4.12).

The alpha histogram in Figure 4.31 appears unimodal with a peak in the distribution of alpha values in the region of 0.3-0.4 and a positive skewness towards rarer higher α values.

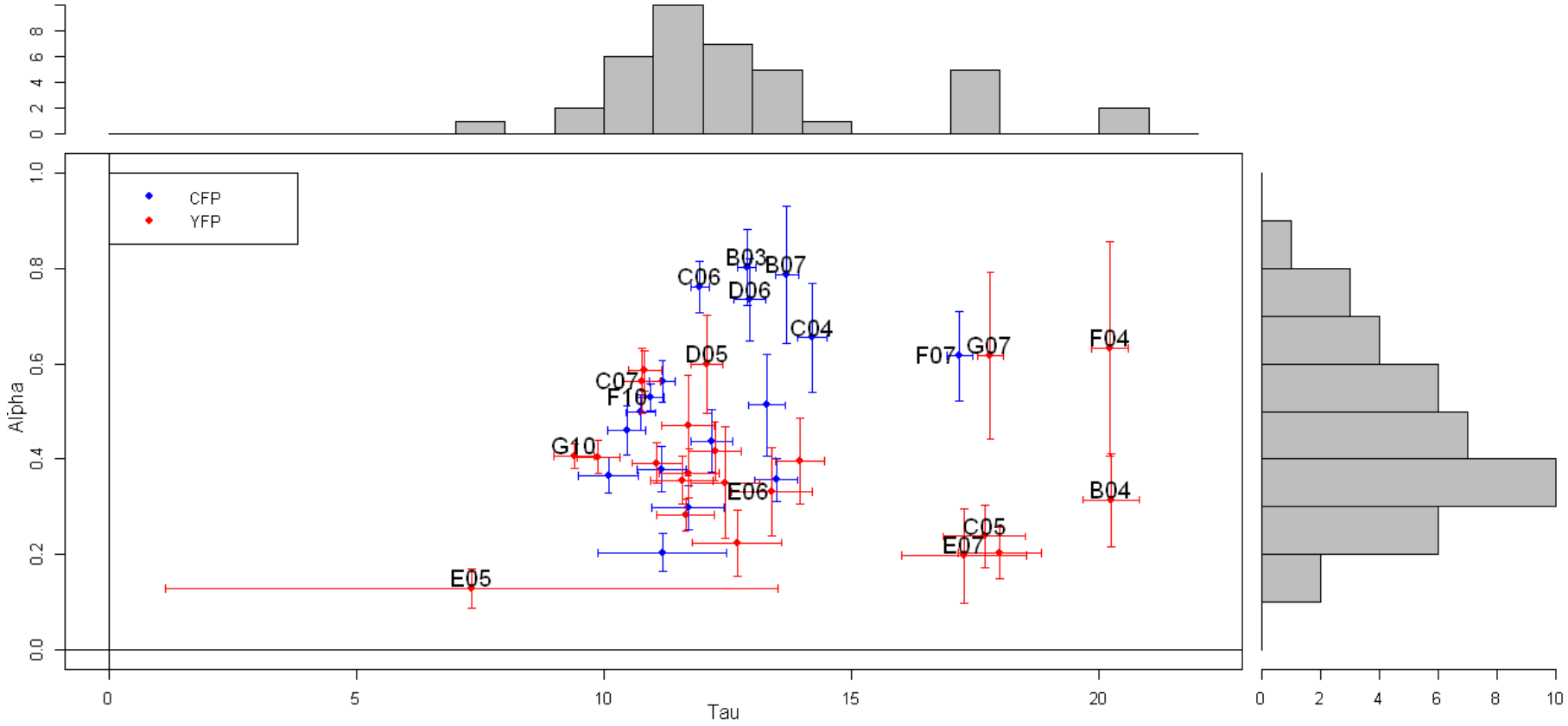


Figure 4.31 Dirac tau onset estimates versus alpha fitness estimates for evolution 10.
Populations showing an increase in CFP fluorescence are shown in blue, populations with an increase in YFP are shown in red. Error bars indicate standard errors.

B03 is the population with the highest alpha value estimate in evolution 10 (0.803+/-0.080 s.e.) and a triploid strain (Evo10_20_B03_CFPe_5) isolated from this population proved to have the highest fitness in competition (see section 5.3.4), substantially higher than the other (diploid) evolved strains. Fluorescence cytometry of the population in this well from plate 15 showed that it has a substantial polyploid component (see Figure 4.32, and section 6.3.8 for further discussion).

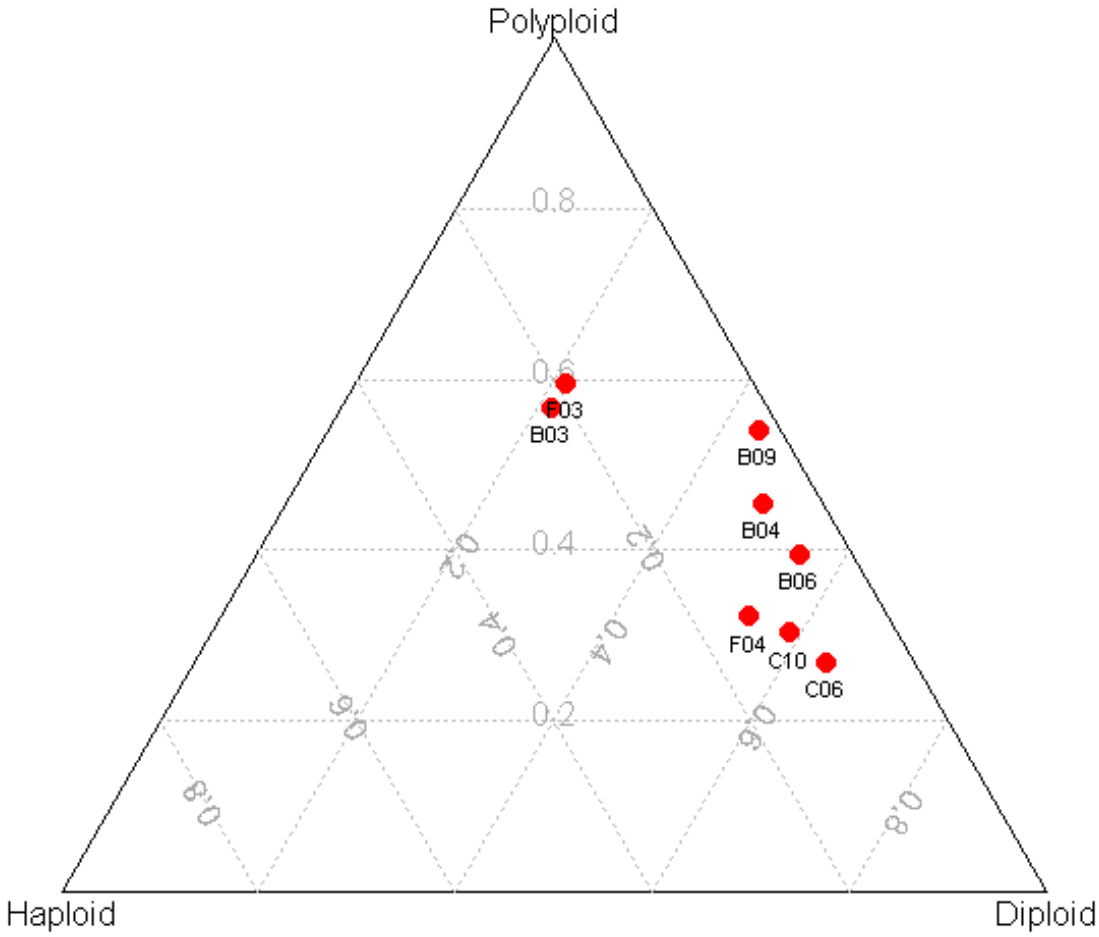


Figure 4.32 Ternary plot of population cell counts partitioned into ploidy categories for some evolution 10 populations sampled at plate 15.
A version of this figure appears subsequently as Figure 6.24d, some points are omitted here for clarity.

However, population C06 which also had a high alpha estimate (0.760+/-0.054 s.e.) did not have a substantial polyploid component in plate 15 despite a tau value (11.937+/-0.192s.e) indicating that prior population change should have occurred. The ploidy of the other populations with high alpha fitness, D06 and B07, was not tested. Similarly, not all of the populations identified in the plate 15 ternary plot (Figure 4.32) as having a substantial polyploid component have high alpha estimates; the alpha estimates for the populations in wells B04, B09 and F03 are unremarkable: 0.315+/-0.097s.e.; 0.405+/-0.034s.e.; 0.586+/-0.043s.e., respectively.

A high proportion of population polyploidy is apparently not necessary or sufficient for high alpha fitness values. The Evo10_20_B03_CFPe_5 strain was not sequenced and may therefore have coding changes which contributed to the high fitness of the population containing it.

Most of the populations fitted to the Dirac curve have tau estimates less than 15 and there is a wait before the final population ratio changes emerge. There was a drop in peak optical density at plate 15 (see Figure 4.9), suggesting that the 6.5% ethanol added might be crossing a physiological threshold for yeast and beneficial mutations that arise after 6.5% may be different from those arising before. However, by plate 15 there are relatively few wells that did not show fluorophore ratio change earlier and this may be a random pattern. No strains isolated from populations showing late changes were sequenced.

4.3.11. Evolution 10 populations with sequenced strains.

Figure 4.33 shows the fluorophore ratio curve for evolution 10 well F10. The curve has a good range (1.197). The dataset required the truncation of three points for an optimal fit but there are nevertheless signs of autocorrelation and non-normality in the errors. The alpha fitness value (0.530+/-0.028s.e.) is close to the mean alpha of 0.447, and the tau onset estimate of 10.931+/-0.288s.e is among the earliest in evolution 10. The diploid strain Evo10_20_F10_CFPe_2 was isolated from this population; it was sequenced (see 6.3.2) and its fitness assessed (see 5.3.4). No sequence changes from the ancestral genome were detected

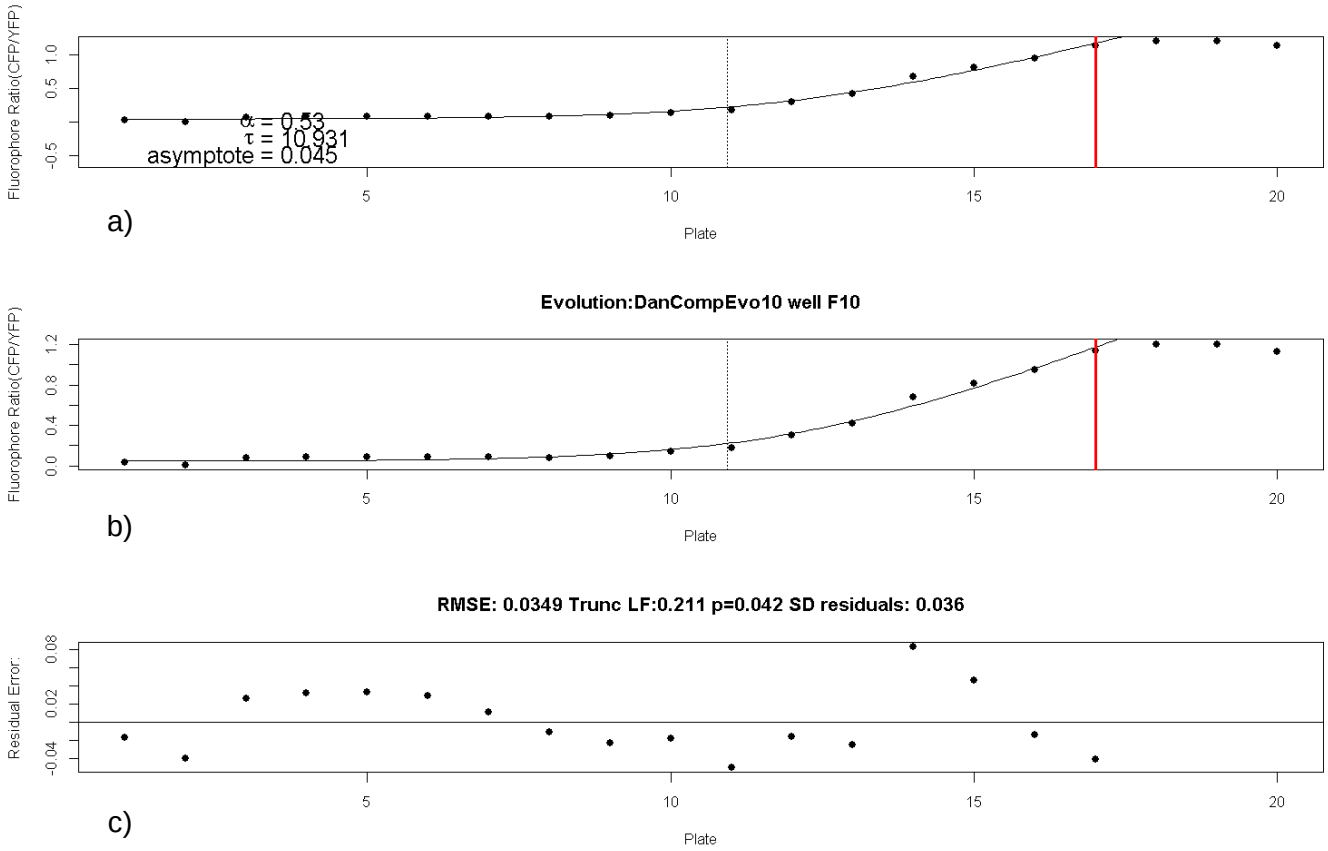


Figure 4.33 Plate versus fluorophore ratio curve (log10 [final raw CFP fluorescence/final raw YFP fluorescence] for the evolution 10 well F10 population.
Part a) shows the fitted curve with standard y axis (-0.5 to 1), b) shows the curve with y axis fitting points, c) shows the residual error of the fit. The data was truncated by three points to best fit the curve. The alpha fitness value of the Dirac curve is 0.530+/-0.028s.e., tau onset is 10.931+/-0.288s.e and asymptote is 0.045. The data range is 1.197. The standard deviation of the residuals is 0.036; the root mean squared error is 0.0349. The Lilliefors statistic for the truncated data is 0.211 (p=0.042) indicating a departure from normality that is just significant.

Figure 4.34 shows the fluorophore ratio curve for evolution 10 well G10. The curve is very similar to F10 with an early tau estimate (9.397+/-0.417s.e), a fitness value (0.406+/-0.0264s.e.) that is slightly lower than F10 and a not wholly satisfactory curve fit (four points were truncated from the dataset and there is some autocorrelation in the errors). The strain Evo10_20_G10_YFPe_2 was isolated from this population, as above see section 6.3.2 for the sequencing of this strain and section 5.3.4 for its fitness assessment. Sequencing also indicated a homozygous SNP in the glycogen de-branching enzyme *GDB1*. Copy number variation strongly indicated that Evo10_20_G10_YFPe_2 has a 50kb heterozygous deletion in a gene-rich subtelomeric region of chromosome V.

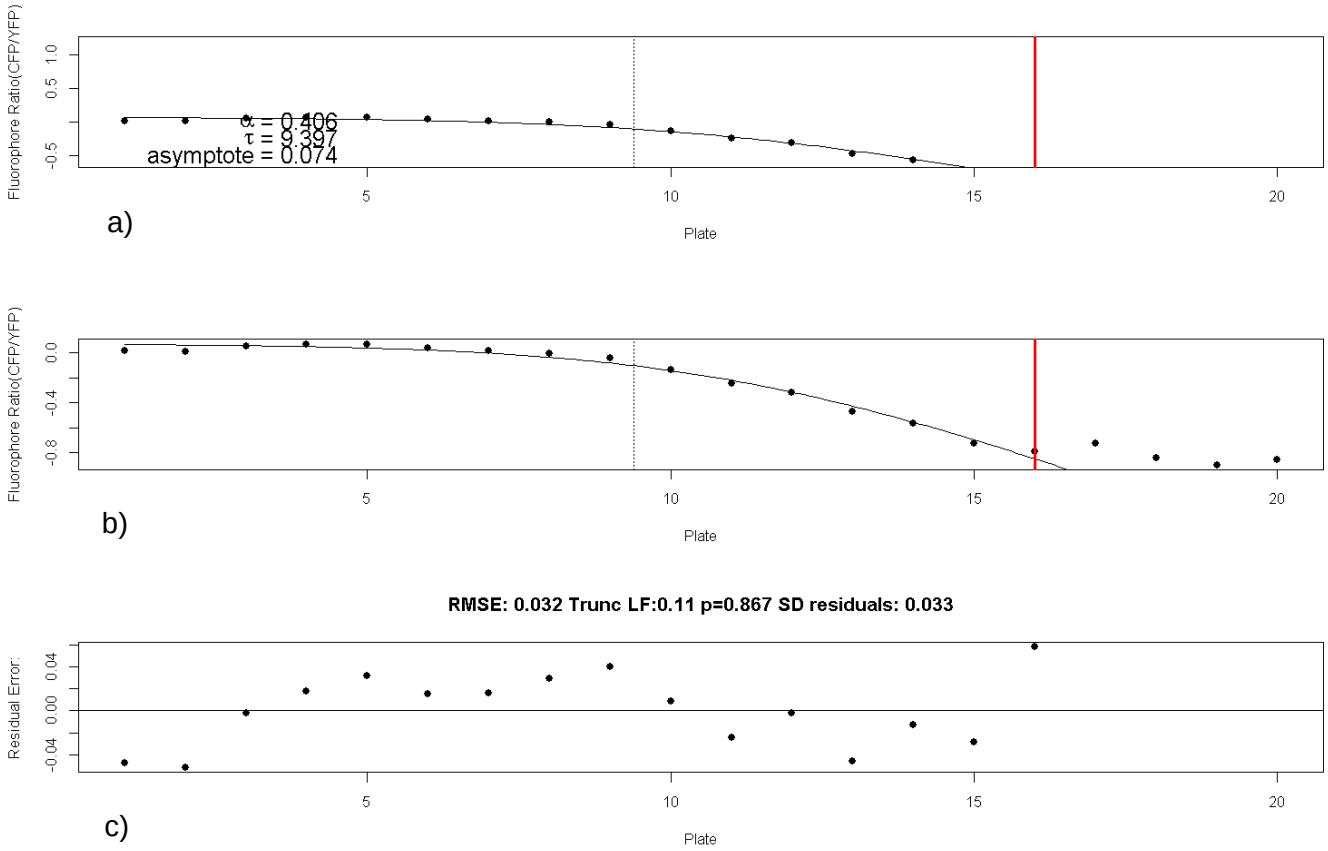


Figure 4.34 Plate versus fluorophore ratio curve (log10 [final raw CFP fluorescence/final raw YFP fluorescence] for the evolution 10 well G10 population.
Part a) shows the fitted curve with standard y axis (-0.5 to 1), b) shows the curve with y axis fitting points, c) shows the residual error of the fit. The data was truncated by three points to best fit the curve. The alpha fitness value of the Dirac curve is 0.406+/-0.0264s.e., tau onset is 9.397+/-0.417s.e and asymptote is 0.074. The data range is 0.967. The standard deviation of the residuals is 0.033; the root mean squared error is 0.032. The Lilliefors statistic for the truncated data is 0.11 (p=0.867) indicating no signficant departure from normality.

Despite the similarities between the two evolution 10 populations, Evo10_20_F10_CFPe_2 grew well in a series of organic acids in Biolog GenIII plates but Evo10_20_G10_YFPe_2 did not (see section 5.3.9 and Figure 5.70 and the figures following it).

4.4 Discussion

4.4.1. Summary

In this chapter the results demonstrated that several of the primary research goals were attained. A series of experiments were carried out that used experimental evolution selection pressures, bottleneck sizes and evolution durations that produced changes in fluorophore ratios within a matter of weeks (less than 200 generations). Population genetic modelling of the bottleneck sizes suggest that these fluorophore ratio changes were not due to neutral drift but due to adaptive evolution (see Section 2.7 “Research Aims and Goals”). Fitting of the Dirac equation to these fluorophore ratio changes enabled the identification not only of experiments in which adaptive evolution had taken place but also to estimate parameters representing the fitness size (α) and the time of origin (τ). In general it was possible to achieve good fits to the Dirac equation for a majority of populations (albeit often with apparent minor autocorrelation in the residuals). Heuristics were devised to enable fitting of truncated fluorophore ratio datasets without over-fitting to the model.

Experimental evolutions characterised by strong mutation and weak selection might be expected to have population fluorophore ratio curves that are very complex and prone to frequent discontinuities due to clonal interference. This would suggest that this evolutionary system is characterised by ‘strong selection’ and relatively ‘weak mutation’; one beneficial mutation often has time to sweep before the next arises, the ‘periodic selection’ of Attwood et al (Atwood et al. 1951). This is somewhat at odds with estimates of tau which suggest that beneficial mutations are starting to sweep in ~70-80 generations within populations picked from a single colony. This suggests that beneficial mutations arise relatively frequently.

This apparent paradox could be resolved if a frequent beneficial ‘mutation’ of substantial fitness size in this system is the conversion of haploid starting strains to diploids, either by endoreduplication or, less frequently, by mating-type switching and mating. This is consistent with the clustering of fitness values for many populations around a constant value. It is also consistent with the absence of detectable sequence changes in both strains isolated from evolution 6 (Evo6_11_B06_CFPe_4

and Evo6_11_D06_YFPe_4) and that isolated from the evolution 8 control (Evo8_14_E07_CFPe_4) and one of the evolution 10 isolated strains (Evo10_20_F10_CFPe_2), see section 6.3.2, despite all these strains being able to outcompete their ancestors (see section 5.3). Among the populations that appear to have higher fitness than the (likely diploid) norm, at least one, evolution 10 well B03, has triploid character. Fluorescence cytometry of cells stained for DNA content shows that the isolated strain with the highest fitness in competition with its ancestors, Evo10_20_B03_CFPe_5 (see section 5.3.4) is triploid and all populations examined by DNA content fluorescence cytometry show a transition from haploid to at least partial diploidy.

In competitive experimental evolutions it is important that the starting strains should have the same genotype aside from the (ideally neutral) marker ensuring comparable evolvability and display the same growth patterns enabling selection for the same traits in each fluorophore lineage. Accordingly, it is problematic that sequencing of the starting fluorophore strains in these experiments revealed that they are not isogenic (see Chapter 6, section 6.4.2), the CFP lineage having an *EGT2* mutation and the ancestral strain competitions showed an earlier lag and diauxy in the YFP lineage. However, there is no initial change in the final fluorophore ratios of the initial plates to suggest a significant difference in initial fitness. Neither is there a significant bias in the strain prevailing when the fluorophore ratio significantly deviates from the initial ratio, suggesting there is no difference in evolvability between the two strains.

4.4.2. Evolution 6

It is problematic to speculate too far on failure to detect genetic change beyond an increase in ploidy from haploid to diploid. Null modelling suggests that if the true bottleneck in evolution 6 fell much below the target of 2000 cells then drift could have produced a fluorophore ratio change that might be mistaken for a selective sweep, particularly in a truncated dataset. Though the range of fluorophore ratio change in evolution 6 was substantially larger than predicted by modelling, that could be explained by the presence of many dead cells in the counted stationary phase cultures.

Only a few strains were isolated from each population (4 colonies for evolution 6). Such small numbers might not retrieve samples of the beneficial mutation causing the fluorophore change.

4.4.3. Evolution 8

The linear ratio changes in the static 6.5% treatments are consistent with a YFP population of higher fitness than the starting CFP being present in the first plate in the evolution series. It is possible that the ancestral YFP strain is intrinsically fitter than the CFP at a static ethanol stress of 6.5%. However, no such disparity between the starting strains is apparent with a 4.5% stress in evolutions 6 or 8 and no general sweep in favour of YFP is seen in evolution 10 as the ethanol concentration reaches 6.5%. However, the most likely cause is that a beneficial mutation has swept from rare through the YFP population within the overnight 6.5% YFP culture and then been transferred to each well with that treatment in the first plate of the evolution. This is consistent with the shared mitochondrial SNP shared by the two sequenced strains from this treatment.

In evolution 8, the fluorophore ratio curves for the 4.5% treatment 8 diverged substantially earlier than the 0% ethanol controls (see Figure 4.26). That is to say that the Dirac fits of the 0% ethanol controls (blue points in Figure 4.26) had larger τ values than the 4.5% ethanol populations. This is consistent with adaptive mutations arising within fewer generations (i.e. arising with greater frequency) or a faster sweep of those mutations with increasing ethanol stress or a combination of the two. Thus in 0% ethanol the capacity of the system for adaptive evolution is smaller and the competing strains can be considered to have lower evolvability in this environment. It is possible that added ethanol stress changes the genotype-phenotype map, making the stressed phenotype less robust to genetic variation, whilst populations without added ethanol display evolutionary capacitance, with their genotypic variation buffered and not visible to selection in the phenotype. One aspect of this effect of stress is that it may induce haploinsufficiency (see Section 2.6 “Experimental evolution starting strains”) which may be ameliorated by ploidy increase (see Section 6.4.2 “Ploidy”).

Ethanol itself may be weakly mutagenic, particularly in the mitochondrial genome (see Section 2.1.3 “Mitochondrial DNA damage”) thereby increasing the supply of mutations. There is less evidence for stress-induced mutagenesis in yeast than in bacteria (Galhardo et al. 2007; Ram & Hadany 2012). An increase in mutational supply due to ethanol stress could increase evolvability.

It is possible that the adaptive landscape for experimental evolution with added ethanol is completely different from that without. However, this need not be the case because ethanol is endogenously produced by fermentation. Some ethanol

tolerance mutations may be antagonistically pleiotropic. This is discussed further when the experimentally evolved strains from the evolution 10 are considered (see Section 5.3.4). These strains are less fit than their ancestors without added ethanol but their relative fitness improves over their ancestor with increasing ethanol stress.

It is conceivable that the earlier divergence of fluorophore ratio might be due to increasing ethanol stress killing a larger proportion of the population and thereby enhancing neutral drift by reducing the true bottleneck population size. However, this seems unlikely since modelling suggests that the 10,000 cell population bottleneck for evolution 8 would have required a five-fold reduction in living cells to show significant neutral drift (see section 4.3.3).

4.4.4. Evolution 10

There was no a priori expectation that the Dirac equation would fit the fluorophore ratio change curves since the fitness sizes of beneficial mutations could have substantially changed with the increasing ethanol stresses in evolution 10. However, only 9 out of 48 populations could not be fit. This does not demonstrate that fitness sizes remained constant since extensive truncation was used to derive most fits. However, whilst inflections in the fluorophore ratio in evolutions 6 and 8 are likely due to clonal interference, in evolution 10 they may also be due to changes in the fitness size of a beneficial mutation with increased added ethanol stress.

4.4.1. MAT switching and histidine auxotrophy

Bharat Rash (personal communication) determined the mating type of the ancestral and evolved strains using the PCR method described by Gerstein (Gerstein et al. 2006). He showed that all the sequenced diploid strains are homozygous alpha at the MAT locus apart from Evo8_14_E07_CFPe_4 which is heterozygous MATa/Matα. This strongly suggests that, despite the HO locus deletion in these strains, the diploid Evo8_14_E07_CFPe_4 arose by mating type switching from *MATα* to *MATa*, followed by mating, whereas the other sequenced strains arose by endoreduplication of the original *MATα*.

The route to diploidy may matter since at least some capacity to synthesise histidine is conferred if they switch mating type from *MATα* to *MATa* (see chapter 2, section 2.6). The *MATα* ancestral strains have a complement of the deleted HIS3 gene fused to the *STE2* promoter of the α-receptor. In *MATα* haploid and diploid cells this is strongly repressed since simultaneous production of both mating pheromones and receptors would arrest cell division. However, it appears that part of the *STE2*

repression is post-transcriptional and some expression of the *STE2* promoter does occur in heterozygous MATa/MATα diploids but not in homozygous *MATα*/*MATα* diploids (Di Segni et al. 2011) conferring some histidine synthetic ability in these strains. Bharat Rash confirmed that on histidine deficient minimal medium, the *MATα* parental strains and the sequenced diploid strains are unable to grow, apart from Evo8_14_E07_CFPe_4 which nevertheless has a long lag phase. It is unclear what, if any, fitness benefit the histidine prototrophic (his+) phenotype might confer in rich media. It is possible that histidine synthesis might only be advantageous in the slow growth phase when the medium is close to exhaustion.

Table 4.6 lists the wells in which some growth was detected in histidine deficient minimal media in the final plate in of each evolution from fluorescence measured in a plate reader. The table also gives the direction of change in the fluorophore ratio (if any occurred) and the Dirac parameters of the curve if the data could be fitted.

Evolution	Well	Ethanol (%)	Fluorophore of His+ cells	Fluorophore of ratio curve	Population alpha (+/- s.e.)	Population tau (+/- s.e.)
6	B08	4.5	CFP & YFP	YFP	1.058 (+/-0.722)	9.830 (+/-0.352)
6	E04	4.5	CFP	CFP	0.289 (+/-0.133)	2.735 (+/-4.032)
6	E06	4.5	YFP	CFP	0.371 (+/-0.052)	2.876 (+/-1.359)
6	E07	4.5	CFP	CFP	0.747 (+/-0.266)	7.761 (+/-0.271)
6	E10	4.5	CFP	CFP	0.401 (+/-0.069)	7.631 (+/-0.693)
6	F09	4.5	YFP	YFP	0.247 (+/-0.033)	2.791 (+/-1.547)
8	B10	0	CFP	-	-	-
8	E06	0	YFP	-	-	-
8	E07	0	CFP	CFP	0.635 (+/-0.019)	8.257 (+/-0.154)
8	E08	0	YFP	CFP	0.687 (+/-0.212)	14.845 (+/-0.396)
8	F06	4.5%	CFP	-	-	-
8	F09	4.5%	YFP	-	-	-
8	F10	4.5%	YFP	YFP	0.400 (+/-0.031)	10.084 (+/-0.314)
10	C07	9% (ramped)	YFP	YFP	0.564 (+/-0.069)	10.758 (+/-0.371)
10	C08	9% (ramped)	CFP	-	-	-
10	E06	9% (ramped)	YFP	YFP	0.331 (+/-0.092)	13.390 (+/-0.808)

Table 4.6 Wells showing growth in his- minimal media in evolutions 6 (plate 11), 8 (plate 14) and 10 (plate 20) and the fluorophore ratio change in those populations. Histidine data from Bharat Rash (pers.comm.)

The populations indicated showed an increase in either CFP,YFP fluorescence or both when a sample frozen in 50% glycerol at -80°C from the final plates in an evolution was grown in minimal media lacking histidine. The table shows the fluorophore of his+ growth together with the prevailing fluorophore and Dirac curve parameter estimates from the fluorophore ratio curve of the population. Wells in which no fluorophore curve was fitted are shown as '-'.

Table 4.6 shows that there is no strong correlation between a population having the ability to grow in histidine deficient media and the fluorophore ratio change it displays, his+ cells arise in populations with alpha (effect size) and tau (onset) across their ranges. The Evolution 6, B08 population has both CFP and YFP colonies that are his+; if histidine biosynthesis is beneficial in this environment then this represents clonal interference. Evolution 6, well E06 has YFP his+ colonies but shows a

fluorophore ratio change in favour of increasing CFP. Five wells showed histidine synthesis without a fluorophore ratio change that could be fitted.

5. Evolved strain phenotypic analysis

5.1 Introduction

This chapter concerns the analysis of growth under ethanol stress of strains derived from populations with fluorophore ratio changes characteristic of adaptive evolution (see chapter 4). A core part of the research aim was to phenotypically characterise strains isolated from these populations and to demonstrate that they outcompeted their ancestors and how their growth profiles and capabilities differed from the ancestral pattern (see Section 2.7 “Research Aims and Goals”).

The competitive performance of these strains was analysed across a range of added ethanol stresses and this performance was broken down by phases of growth: emergence from lag, maximum growth rate, duration of active growth and total fluorescence.

It became apparent in the process of characterising evolved strains that they were larger than their ancestors and this was investigated further. The size change of an ancestral and an evolved strain (Evo10_20_F10_CFPe_2) during the transition from lag phase to logarithmic growth was investigated as well as the cell sizes in of post-diauxic populations grown under varying ethanol stress. In such post-diauxic cultures the glucose concentration in the medium has fallen below 0.2% and the cells have entered the slow-growth phase (See Introduction, section 2.5.1).

A physiological advantage of larger cell size may be, at least part of, the explanation for the emergence of diploids under ethanol stress in these experimental evolutions. A phenotypic screen was also carried out using Biolog GenIII plates which have a range of nutrients and chemical stressors dried into each well base. Evolved strains were grown on these plates to look for differences in phenotype between the evolved strains and also between evolved strains and their ancestors.

5.2 Methods

5.2.1. Evolved strain growth analysis

Overnight cultures were made from a single colony grown in YPD grown at 30°C with 200rpm shaking. 100µl of this culture was split into YPD media and YPD containing added ethanol (concentrations calculated as percentage by volume). These cultures were grown as before to the end of log phase (as measured by OD_{600}), approximately 48 hours of growth. The duration of incubation was determined by the

cultures with the highest ethanol concentrations. These cultures were diluted x100 and the cell concentration and sizes determined using the standard Nexcelom Cellometer method (see General Methods section 3.2). The counted x100 dilution cultures were then further diluted to give a desired cell count of 10,000 cells in 120µl; 5,000 (live and dead) cells per strain in competition. 10,000 cells is the population bottleneck size for evolution 8. This volume was added to the wells in a standard sterile 96-well plate (see Figure 5.1 for a typical layout) and read in a plate reader (see General Methods section 3.3). This plate layout has replicates of the competition in YPD and 2, 4, 6 and 8% ethanol as well as blanks and single strain growth trials in these ethanol concentrations. Unlike subsequent fitness trials, the first three strains Evo8_14_C06_CFPe_2, Evo8_14_E07_CFPe_4, and Evo8_14_F03_CFPe_2 were run without an ancestral competition for comparison.

	1	2	3	4	5	6	7	8	9	10	11	12
A	YPD Blank	YPDE 2% Blank	YPDE 4% Blank	YPDE 6% Blank	YPDE 8% Blank	YPD Blank	YPDE 2% Blank	YPDE 4% Blank	YPDE 6% Blank	YPDE 8% Blank	YPD Blank	YPDE 2% Blank
B	YPD Blank	CFPa (YPD)	CFPeF10:2 + YFPa (YPD)	CFPeF10:2 + YFPa (YPDE 2%)	CFPeF10:2 + YFPa (YPDE 4%)	CFPeF10:2 + YFPa (YPDE 6%)	CFPeF10:2 + YFPa (YPDE 8%)	CFPeF10:2 + YFPa (YPD)	CFPeF10:2 + YFPa (YPDE 2%)	CFPeF10:2 + YFPa (YPDE 4%)	YFPa (YPD)	YPDE 4% Blank
C	YPDE 8% Blank	CFPa (YPD)	CFPeF10:2 + YFPa (YPDE 6%)	CFPeF10:2 + YFPa (YPDE 8%)	CFPeF10:2 + YFPa (YPD)	CFPeF10:2 + YFPa (YPDE 2%)	CFPeF10:2 + YFPa (YPDE 4%)	CFPeF10:2 + YFPa (YPDE 6%)	CFPeF10:2 + YFPa (YPDE 8%)	CFPeF10:2 + YFPa (YPD)	YFPa (YPD)	YPDE 6% Blank
D	YPDE 6% Blank	CFPa (YPD)	CFPeF10:2 (YPD)	CFPeF10:2 (YPDE 2%)	CFPeF10:2 (YPDE 4%)	CFPeF10:2 (YPDE 6%)	CFPeF10:2 (YPDE 8%)	CFPeF10:2 (YPD)	CFPeF10:2 (YPDE 2%)	CFPeF10:2 (YPDE 4%)	YFPa (YPD)	YPDE 8% Blank
E	YPDE 4% Blank	CFPa (YPD)	YFPa (YPDE 2%)	YFPa (YPDE 4%)	YFPa (YPDE 6%)	YFPa (YPDE 8%)	YFPa (YPDE 2%)	YFPa (YPDE 4%)	YFPa (YPDE 6%)	YFPa (YPDE 8%)	YFPa (YPD)	YPD Blank
F	YPDE 2% Blank	CFPa (YPD)	CFPa + YFPa (YPD)	CFPa + YFPa (YPDE 2%)	CFPa + YFPa (YPDE 4%)	CFPa + YFPa (YPDE 6%)	CFPa + YFPa (YPDE 8%)	CFPa + YFPa (YPD)	CFPa + YFPa (YPDE 2%)	CFPa + YFPa (YPDE 4%)	YFPa (YPD)	YPDE 2% Blank
G	YPD Blank	CFPa (YPD)	CFPeF10:2 + YFPa (YPDE 2%)	CFPeF10:2 + YFPa (YPDE 4%)	CFPeF10:2 + YFPa (YPDE 6%)	CFPeF10:2 + YFPa (YPDE 8%)	CFPeF10:2 + YFPa (YPD)	CFPeF10:2 + YFPa (YPDE 2%)	CFPeF10:2 + YFPa (YPDE 4%)	CFPeF10:2 + YFPa (YPDE 6%)	YFPa (YPD)	YPDE 4% Blank
H	YPDE 8% Blank	YPDE 6% Blank	YPDE 4% Blank	YPDE 2% Blank	YPD Blank	YPDE 8% Blank	YPDE 6% Blank	YPDE 4% Blank	YPDE 2% Blank	YPD Blank	YPDE 8% Blank	YPDE 6% Blank

Figure 5.1 The plate layout used to assess the fitness in ancestral competition for the experimentally evolved strain Evo10_20_F10_CFPe_2 with varying amounts of added ethanol.

The plate has blanks with 0,2%,4%,6% or 8% ethanol around the outside and: wells containing CFP only in YPD medium (column 2, row B to G) and CFP with varying ethanol stress (row D, column 3 to 10); wells containing YFP only in YPD medium (column 11, row B to G) and YFP with varying ethanol stress (row E, column 3 to 10); a competition between ancestral strains in varying ethanol (row F, column 3 to 10); competitions between the evolved strain and the YFP ancestor in varying ethanol concentrations (rows B, C and G, columns 3 to 10).

Unless otherwise stated, all time-courses considered will be the natural logs of the fluorescence in arbitrary units against the elapsed time in seconds.

Correcting for the CFP fluorescence of the YFP fluorophore

Whilst the YFP fluorescence by CFP strains is negligible, there is a small but significant CFP fluorescence from the YFP fluorophore. Accordingly, all the plate layouts had wells containing YFP alone in column G (see Figure 5.1). These single strain populations were used to measure the CFP signal produced by the YFP fluorophore at different concentrations, enabling the CFP signal from CFP/YFP competitions to be corrected.

Figure 5.2 shows a plot of CFP against YFP for the YFP strain. A three-parameter asymptotic curve (see Equation 5.1) was fit to the data using SSasymp a self-starting non-linear regression modelling package in the base R statistics package (Pinheiro & Douglas Bates n.d.).

$$y = a - be^{-cx}$$

Equation 5.1 The three parameter asymptotic exponential equation fitted by the self-starting non-linear least squares package SSasymp.
The parameter a is the asymptote, the maximum fluorescence, b is the response when the fluorescence is zero and c is the rate constant.

This model was then used to calculate the aberrant CFP fluorescence due to the YFP fluorophore in mixed strain competitions. This value was then subtracted from the CFP fluorescence in those competitions. The blue points in Figure 5.2 show the effect of applying the correction to the raw CFP fluorescence. Several curves show increases in CFP without corresponding YFP increase which may be due to quenching of the YFP fluorescence by falling pH at diauxy while growth continues to increase the amount of YFP fluorophore, and hence its CFP fluorescence.

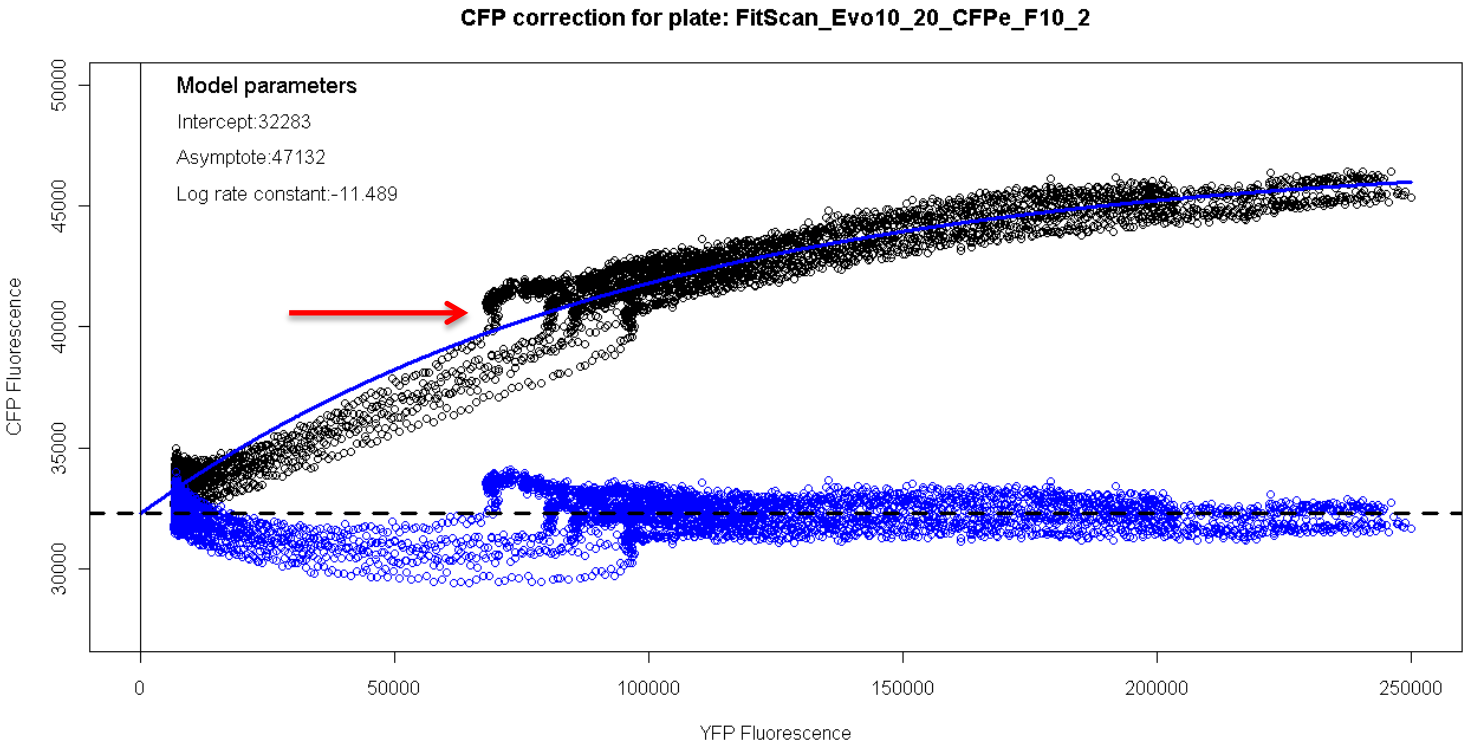


Figure 5.2 The raw (no blank correction) CFP fluorescence (black points) of the YFP ancestral strain plotted against its YFP fluorescence when the YFP strain is grown alone as part of the fitness analysis of Evo10_20_CFPe_F10_2.
These six wells (column 11 row B to G) contain YPD with no added ethanol. The blue curve is a three parameter asymptotic fit to the raw data with intercept 32283, asymptote 47132 and log rate constant -11.489. The blue points show the corrected CFP once the curve has been subtracted from the raw CFP. The red arrow marks one example of a discontinuity where the CFP value increases without corresponding YFP increase. This discontinuity may be due to quenching of the YFP fluorescence by a pH fall after diauxy while growth continues to produce more YFP fluorophore (and hence more CFP fluorescence).

The corrected CFP data was stored separately in MySQL from the raw data to retain an audit trail. Although this CFP fluorescence by the YFP fluorophore is an order of magnitude smaller than its YFP fluorescence, it can shift the appearance of significant events in the growth curve. In Figure 5.3 the log raw CFP fluorescence is in blue and the corrected CFP fluorescence in green. The timing of the exit from the lag and the maximum rate of growth are substantially changed.

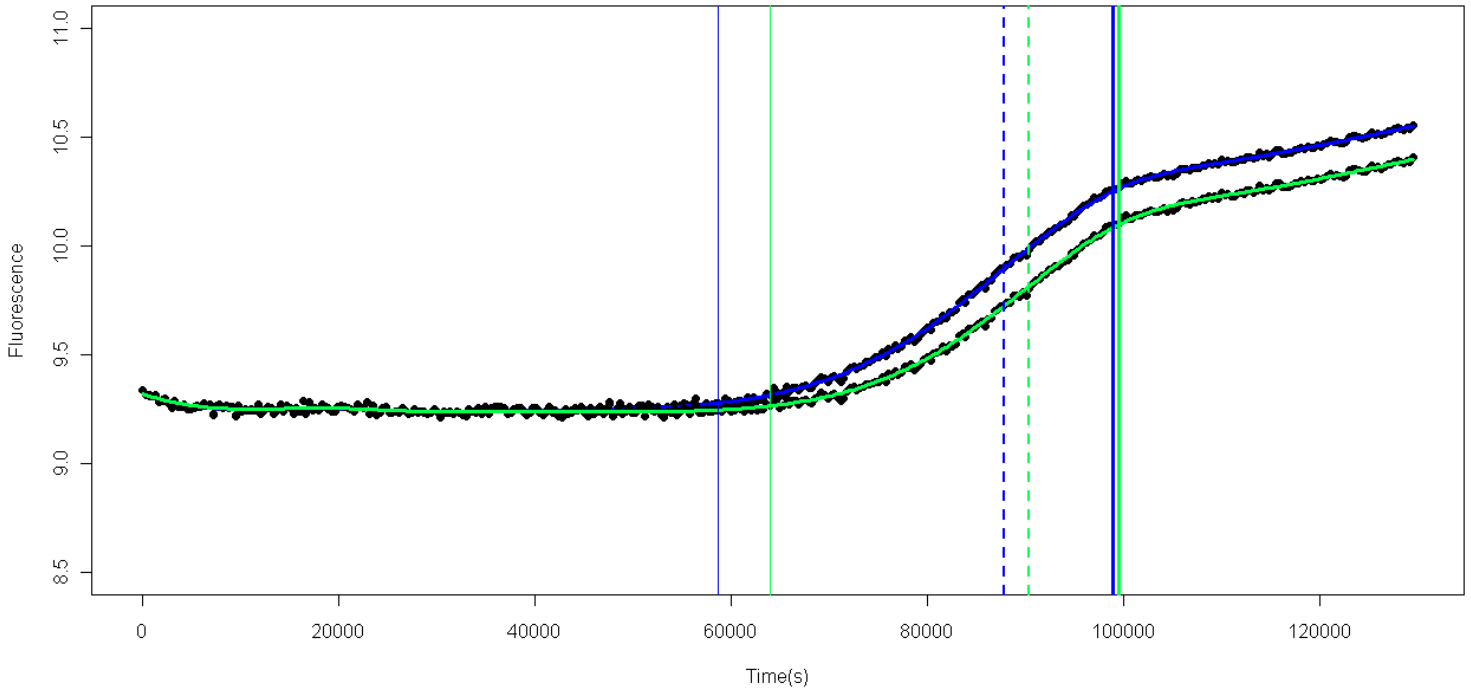


Figure 5.3 Changes in CFP fluorescence against time for Evo10_20_CFPe_F10_2 competing against YFP in 2% ethanol due to correction for fluoresce in the CFP band by the YFP fluorophore.

A Pspline smoothed curved for the raw CFP fluorescence (blanked so its minimum value is 1000) is in blue and the corrected curve is in green. The vertical lines indicate significant points in the growth curve; moving left to right, the fine solid lines show the exit from lag, the dotted lines show the maximum growth rate and the heavier solid lines the onset of the slow growth phase, blue lines show the uncorrected CFP data and green lines show the corrected data. The calibration curve shows that the time when the end of lag is called is substantially later for the corrected data (since fluorescence is reduced by correction and the end of lag is called when fluorescence exceeds a threshold), the maximum rate for the corrected data is a little later and there is almost no effect on the onset of slow growth

Blanking the data

Taking the log of the fluorescence allows the rate of growth to be independent of population size. It is critical in a log plot that the baseline for comparing fluorescence is the same otherwise the log transformation will magnify the increase in fluorescence of the lower plot and since the fluorescence baselines differ they must be blanked for comparison. However, blanking the initial values to zero will not only prevent the taking of logs but also magnify the noise in the baseline data compared to the signal of growth. The minimum fluorescence in each dataset was set to an arbitrary value of 1000 fluorescence units before taking the natural log. This means that there is a consistent baseline for detecting the emerging signal of growth that discounts the effects of photo bleaching.

Data smoothing

The blanked (and corrected) CFP and the YFP fluorescence was smoothed using Pspline (version 1.0-14), a penalized smoothing spline R Package (Heckman & Ramsay 1998; Ramsey & Ripley 2009). The intensive measurement of fluorescence at five minute intervals enabled the Pspline smoothing package to produce the first and second differentials of the data providing growth rate and acceleration of growth. Figure 5.4 shows the time course of a competition with Pspline smoothed curves, while Figure 5.5 shows the rate of growth of this competition and Figure 5.6 shows the second differential (acceleration).

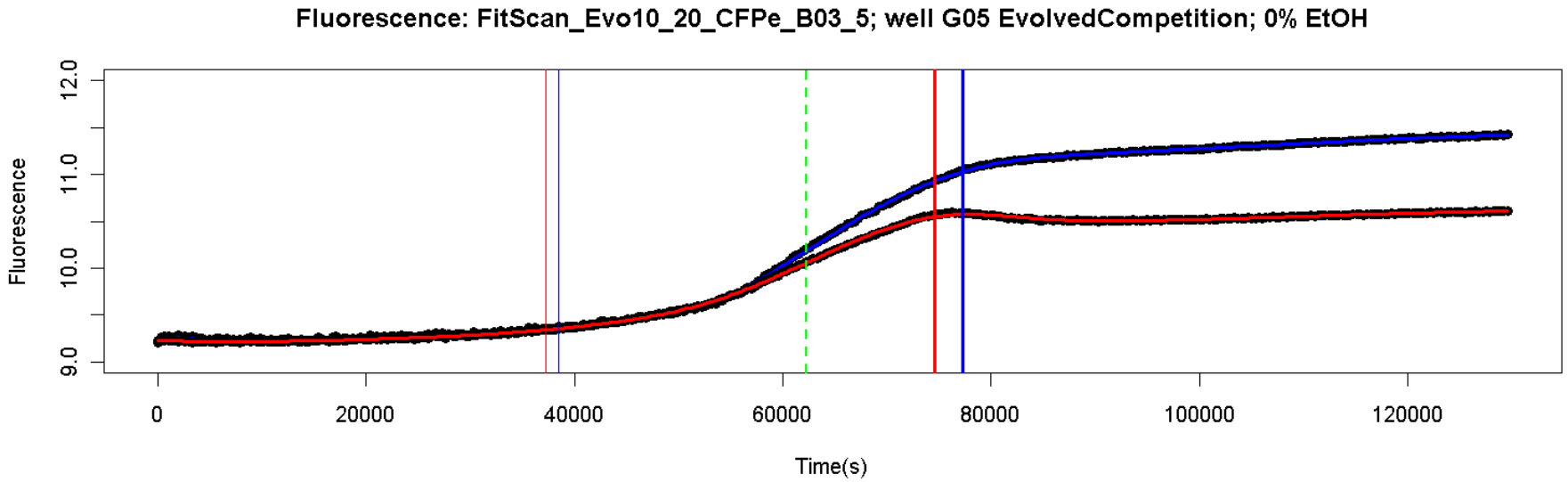


Figure 5.4 Time course of log fluorescence for evolved strain Evo10_20_CFPe_B03_5 competing with ancestral YFP in YPD (no added ethanol).
The measured fluorescence values are represented by points and the blue and red curves are the pspline smoothed fits to those points for the evolved CFP strain and ancestral YFP strain respectively. The fine vertical lines in blue (for the Evo10_20_B03_CFPe_5 strain) and red (for the ancestral YFP strain) show where the lag phase has reliably ended, where the growth rate crosses the $1x10^{-5}$ rate threshold in the first differential (see Figure 5.5). The green vertical dotted line shows the maximum rate for both strains coincides. The heavy vertical lines show the onset of diauxy for the CFP strain (blue) and YFP (red).

Whilst smoothing the data reduces the likelihood that noise will be mistaken for growth, there is also an upper limit to the legitimate smoothing before the data is distorted by artifacts. The tightest inflexion in the time course curves is the decline in the rate of growth at the onset of diauxy in YFP fluorescence and that sets the upper bound to the smoothing. If the data is over-smoothed the timing of YFP diauxy will advance. All strains were evaluated with the same Pspline smoothing parameter, $1x10^{19}$, a compromise between good smoothing during the lag and over-smoothing at diauxy; the curves show that there is some over-smoothing at YFP diauxy (see Figure 5.4) in return for less variation during the lag. This optimal smoothing value is substantially greater than the values obtained by minimisation of the generalized cross-validation criterion, which may be due to some autocorrelation in the data.

Determining the lag duration and the maximum rate

Detecting the point at which the growth of a small population of cells is detectable is problematic. If the threshold for the start of detectable growth is set to zero then any trivial variation in the fluorescence can appear to be growth.

Some have reported lag detection by extrapolating a linear fit of the rate backwards from the maximum, the intersection with zero growth gives the time of exit from lag (Lodge & Hinshelwood, 1943; Robinson et al. 1998). Whilst this is a valid method with data derived from counts of colony forming units, this data uses fluorescence as a proxy for growth and the calibration curve of cell count against fluorescence is non-linear. The extrapolation method of lag determination also depends upon a constant (semi-log transformed) rate of growth. Many of the curves in these experiments showed an acceleration of growth during the 'log' phase (see Figure 5.5). Extrapolating from such a rate curve would show a misleadingly late time for onset of growth.

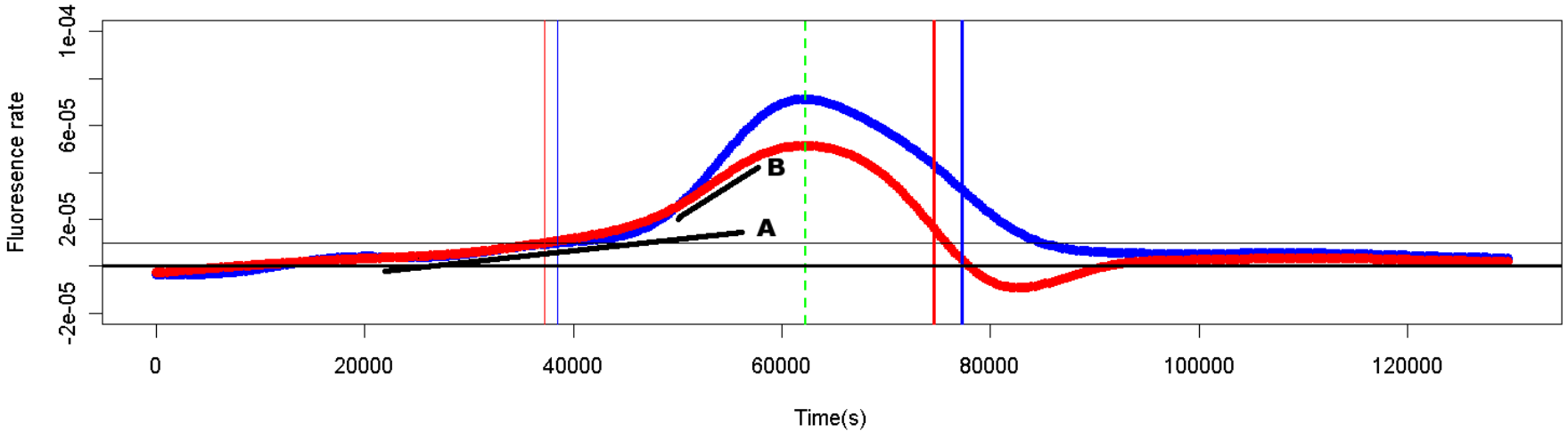


Figure 5.5 Rate of CFP (blue) and YFP fluorescence (red) change (first differential of Figure 5.4) for a competition between the evolved CFP strain Evo10_20_B03_CFPe_5 and the ancestral YFP strain with no added ethanol.

Rate values produced as the first differential of pspline smoothed time course. The horizontal line at a fluorescence growth rate of $1x10^{-5}$ fluorescence units/second is the threshold rate for the end of the lag phase. The fine vertical lines in blue (for the Evo10_20_B03_CFPe_5 strain) and red (for the ancestral YFP strain) show where the lag phase has reliably ended, where the growth rate crosses the $1x10^{-5}$ rate threshold. The green vertical dotted line shows the maximum rate for both strains coincides. The heavy vertical lines show the onset of diauxy for the CFP strain (blue) and YFP (red). The red YFP rate plot does not have a constant log rate; the black line marked A is approximately parallel to the early rate which subsequently accelerates black line B is approximately parallel to the higher rate.

Setting a growth rate threshold above zero will show an end of the lag phase that is later than the true value. However, since growth cannot be detected for approximately 4-5 divisions after the end of lag this seems a trivial sacrifice for a more reliable value. Accordingly, the threshold for the end of the lag phase was set at an increase in fluorescence of $1x10^{-5}$ log(fluorescence)/s; this is approximately 1 fluorescence unit per second.

The time and value of maximum rate of growth for each fluorophore was derived from the highest value in the first differential.

Diauxy

The onset of diauxy for each strain was derived from the minimum value in the second differential (see Figure 5.6).

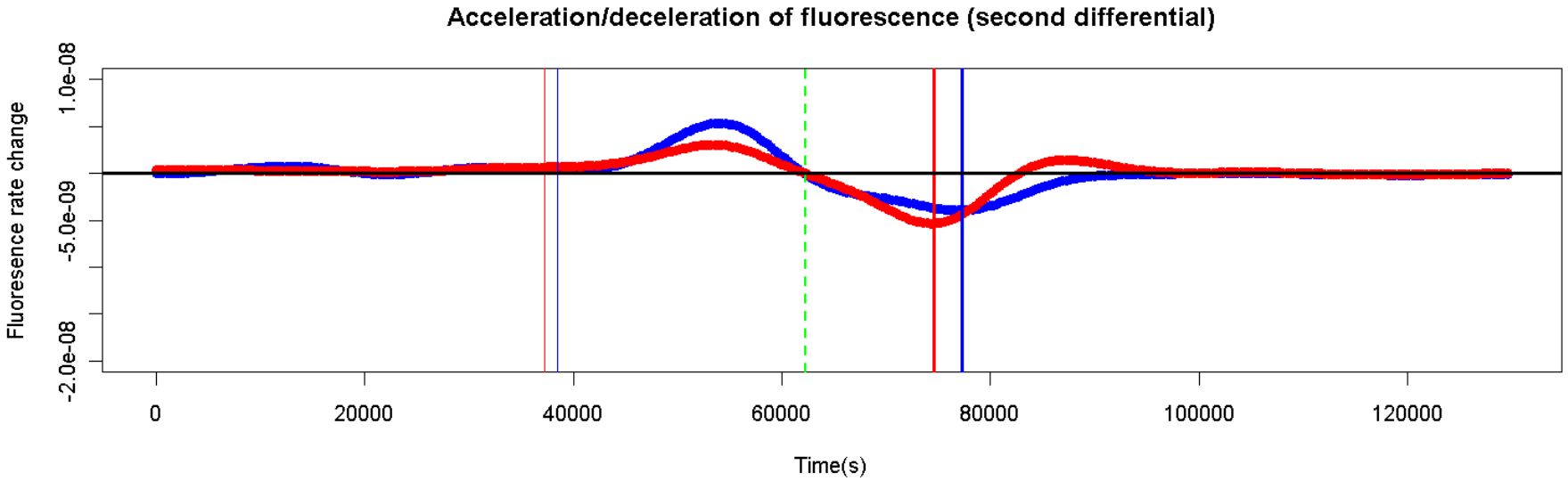


Figure 5.6 Second differential for a competition between the evolved CFP strain Evo10_20_B03_CFPe_5 and the ancestral YFP strain with no added ethanol
The fine vertical lines in blue (for the Evo10_20_B03_CFPe_5 strain) and red (for the ancestral YFP strain) show where the lag phase has reliably ended, where the growth rate crosses the $1x10^{-5}$ rate threshold in the first differential (see Figure 5.5). The green vertical dotted line shows the maximum rate for both strains coincides (second differential is zero). The heavy vertical lines show the onset of diauxy for the CFP strain (blue) and YFP (red) at the maximum rate of deceleration (second differential is at a minimum).

Growth

Values were derived for the increase in fluorescence as a proxy for growth (see Introduction, Section 2.5.4 for a discussion of this) at the onset of diauxy and at 36 hours (which was after the onset of diauxy for all strains). The bulk of fluorescence increase, putative growth, occurs between the exit from lag and the onset of diauxy. Accordingly, this duration and the overall rate together determine the amount of growth. Since fluorescence is a proxy for the amount of growth, the fluorescence by a set time after diauxy is a measure of the growth of a strain. A time of 36 hours was chosen as the set time for expressing growth; this is after the onset of diauxy for all strains at up to 10% ethanol.

5.2.2. Cell Size

Cell cultures were grown to the slow growth phase. Whilst YPD without ethanol could provide a dense slow growth culture in 18-24 hours, above an 8% added ethanol stress longer incubations were required. These cell cultures were counted using the usual methods (see Chapter 3, section 3.2). The mean cell diameters were plotted using the default settings in the geom_density function in the ggplot2 library (version 0.9.3.1) in R (version 2.15.2).

The initial and log phase mean cell diameter for two strains, the ancestral YFPa and Evo10_20_F10_CFPe_2, was also evaluated. Overnight cultures were diluted to an OD_{600} of 0.1 in YPD medium containing no ethanol and put separately into multiple wells on a 96 well plate. The cell sizes were measured after 30 minutes, 4 hours and 50 minutes, and 5 hours and 50 minutes.

5.2.3. Biolog GenIII plate phenotyping

The Biolog (21124 Cabot Blvd., Hayward, CA 94545, see www.biolog.com) GenIII plate was used for evolved strain evaluation; each plate has 71 different carbon sources and 23 stressors. Different growth in each well gives a rapid phenotypic profile which can be used to indicate changes in phenotype between ancestors and derived strains. Biolog will not divulge the exact compositions of the growth mediums and this reticence extends to the concentrations of the stressors or carbon sources. Yeast grows reliably in GenIII plates but their intended use is for bacterial identification and accordingly the growth media are unlikely to be optimal for yeast growth and yeast does not grow best in the GenIII "positive control" well.

15ml of YPD cultures of the two ancestors and eight evolved strains were incubated overnight at 30^{o}C with 200 rpm shaking. The cultures were washed three times by

centrifugation at 3000 rpm for 5 minutes followed by vortex resuspension of the pellet in 15ml of sterile distilled water. The OD_{600} of a 1/100 dilution of each washed suspension was used to calculate the dilution required to produce suspensions with an OD_{600} of 0.05. Each well of a room temperature Biolog GenIII plate was filled with 120µl of one suspension. A dye is available for measuring respiration in GenIII plates but this was not used in these experiments. Care was taken to ensure that a stable baseline OD_{600} for each well was established by monitoring each plate in a plate reader for 1 hour, taking a reading every 5 minutes with double orbital shaking at 200rpm for 1 min each reading. This ensured that the contents of each well were properly dissolved. The plates were then incubated for 24 hours at 30°C, shaking at 200rpm before the final OD_{600} was measured using the same plate reader. Each strain was trialled twice and all of the plates in each replicated came from a single batch. .

Growth in each well was measured by the change in OD_{600} over 24 hours.

5.3 Results

At the end of each evolution series evolved populations from a number of wells were plated and checked under a Leica MZ10 fluorescence stereo microscope. No colonies were found that lacked a fluorescent marker suggesting that mutations or deletion of the fluorophore gene does not confer a strong selective advantage.

5.3.1. Ancestral Competitions

The first two competitive fitness analyses, with strains Evo8_14_C06_CFPe_2 and Evo8_14_E07_CFPe_4, were run without an ancestral competition for comparison (figures for these strains show an ancestral competition from other plates). Two ancestral competitions were run on the PolarStar Omega, one from evolution 6 (Evo6_11_B06_CFPe_4) and one from evolution 8 (Evo8_14_F03_CFPe_2). Four ancestral competitions were run on the BMG FluoStar: one strain from evolution 6, Evo6_11_YFPe_D06_4, and three from evolution ten Evo10_20_CFPe_B03_5, Evo10_20_CFPe_F10_2, Evo10_20_YFPe_G10_2.

Ancestral strains: growth lags

See Figure 5.7 for a plot of the growth lag for each ancestral strain with increasing ethanol content of the media. This shows that the lag increases linearly with increasing ethanol concentration. The ancestral YFP strain starts detectable growth about 2.5 hours sooner than the CFP strain across the ethanol range from 0 to 10%; for the YFP strain the intercept of the linear fit is 43,705s+/-1529s.e., whilst that of

the CFP strain is 52942s +/-1958 s.e. However, this apparent consistent YFP advantage may be due to a difference in the sensitivity of detection for each fluorophore when the fluorescence due to yeast growth is very small. The gradients of change of lag with ethanol are not substantially different; for the ancestral CFP strain it is 3134s/ethanol(%) +/-452 s.e., against 3222s/ethanol(%)+/-353se.) for the YFP strain.

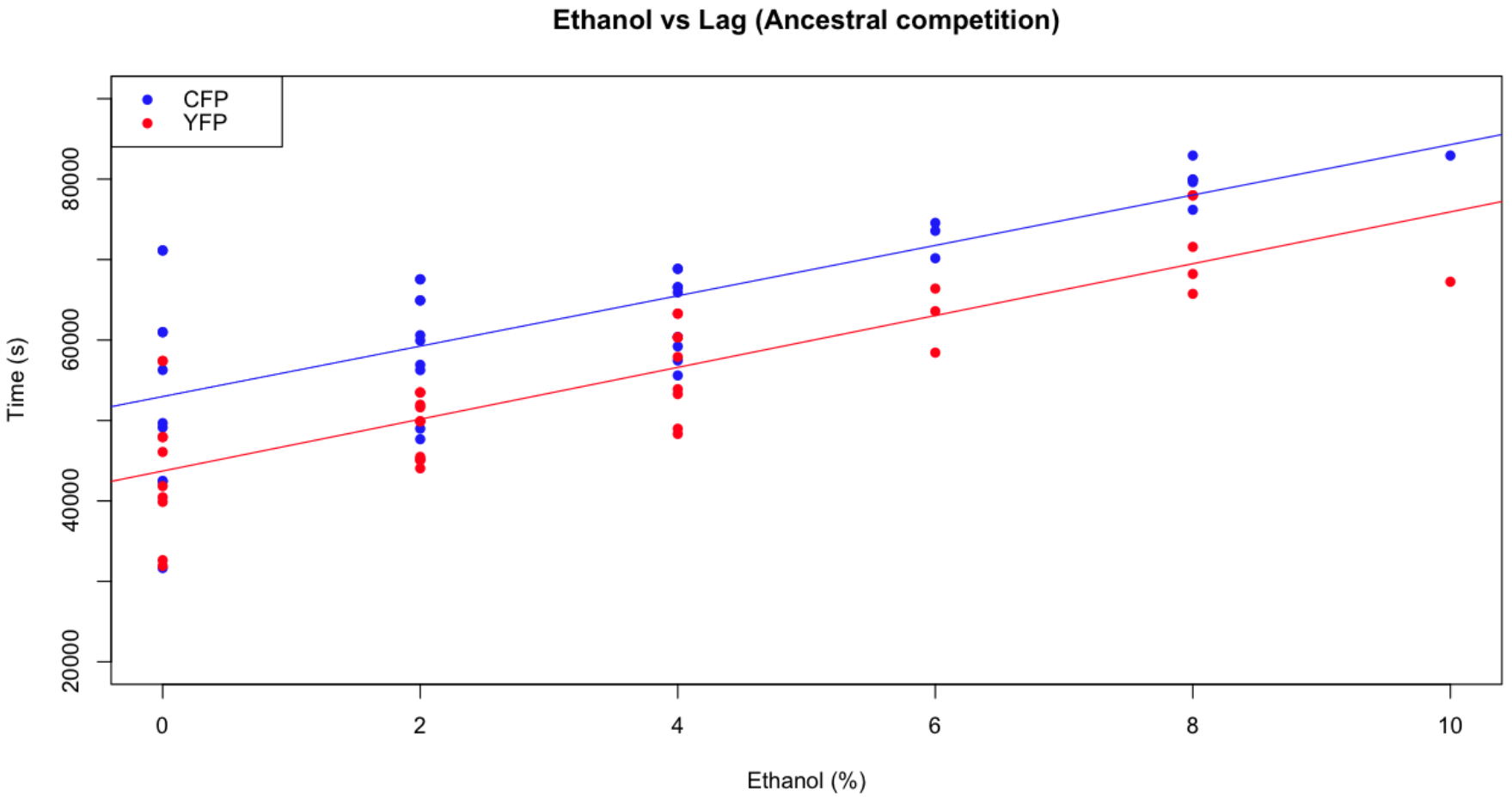


Figure 5.7 A plot of the lag phase before the onset of detectable growth during the competition of ancestral CFP and YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG FluoStar plate reader. The points in blue are for the CFP strain and the points in red are for the YFP strain.

When R (v 2.15.2) is used to perform an analysis of covariance of the model:

lag~ethanol*fluorophore

The following summary is produced (see Table 5.1)

.	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	52942.33	1756.87	30.134	< 2e-16 ***
Ethanol	3134.12	405.87	7.722	5.06e-11 ***
fluorophoreYFP	-9237.02	2484.59	-3.718	0.000395 ***
ethanol:fluorophoreYFP	87.82	573.99	0.153	0.878826

Table 5.1 Analysis of covariance for the model lag~ethanol*fluorophore for the onset of detectable growth during the competition of ancestral CFP and YFP strains in media with different ethanol concentrations

There is no significant difference between the slopes of the lag curve for each fluorophore (p=0.879) but there is a significant difference in the intercept, the lag for the YFP fluorophore strain ends an estimated lag 9237s (just over 2.5 hours) earlier than for the CFP strain. However, some caution should be exercised in interpreting this data. It contains data from three plates and the ancestral competitions were accordingly plated out from three overnight cultures. Accordingly the points are not independent and there is considerable pseudo-replication in this dataset which may weaken the statistical inference.

Ancestral strains: onset of diauxy

Similarly the onset of diauxy is slightly earlier for the ancestral YFP strain (intercept 82970s+/-1726s.e.) than the CFP strain (intercept 87364s +/-1767s.e.), see Figure 5.8, a difference of a 1.2 hours (approximately half the difference seen at the end of lag).

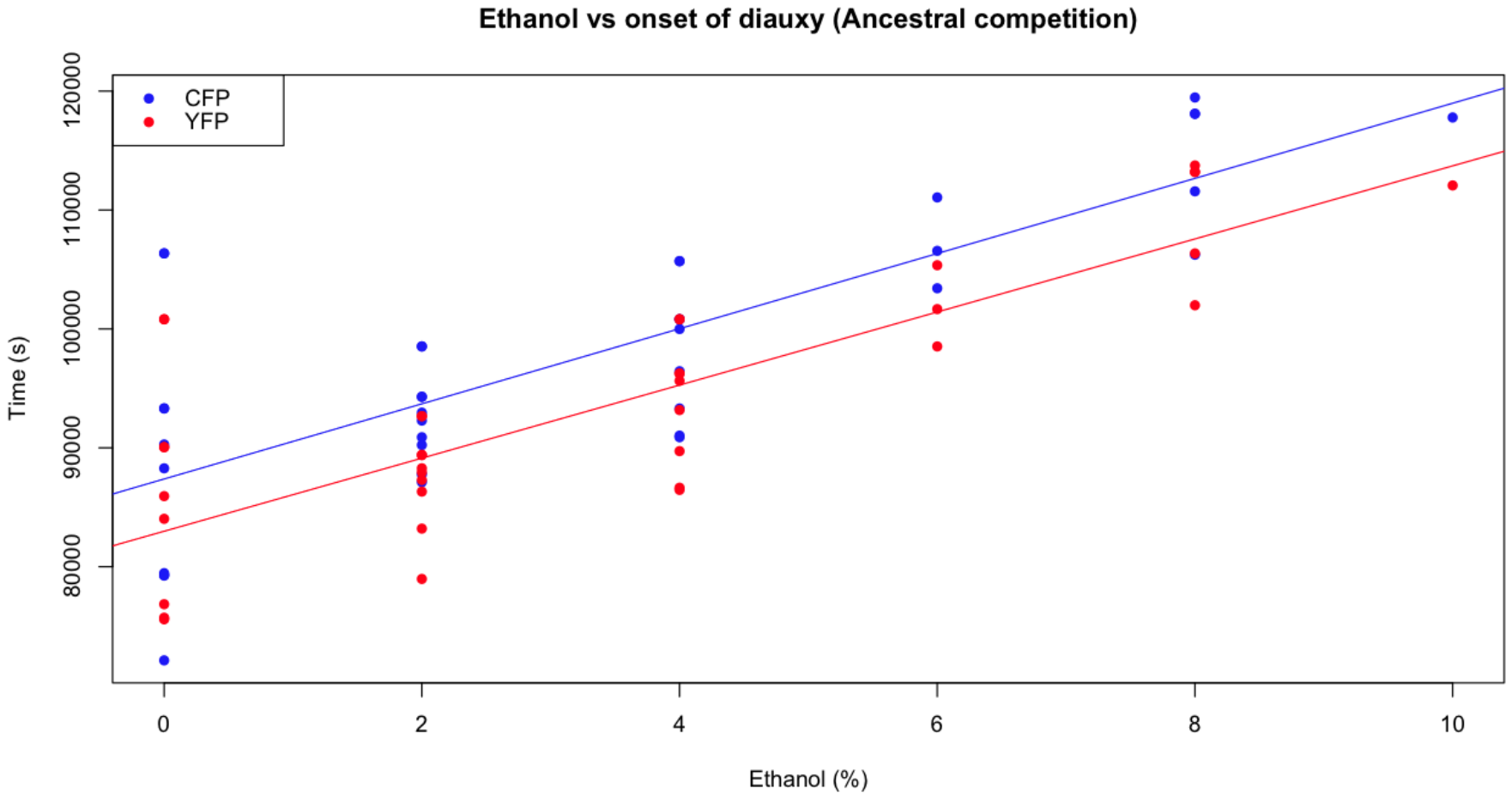


Figure 5.8 A plot of the onset of diauxy during the competition of ancestral CFP and YFP strains in media with different ethanol concentrations. Competitions were run on the BMG FluoStar plate reader. The points in blue are for the CFP strain and the points in red are for the YFP strain.

Ancestral strains: maximum growth rates

The maximum growth rate is more variable than the other growth metrics. As the concentration of ethanol increases the maximum rates of fluorescence growth increase for both strains but aside from three outlier points with low values at 6, 8, 10% from the Evo10_20_B03_CFPe_5 plate, the rates of growth for each strain appear very similar with the points from each competing strain showing considerable overlap.

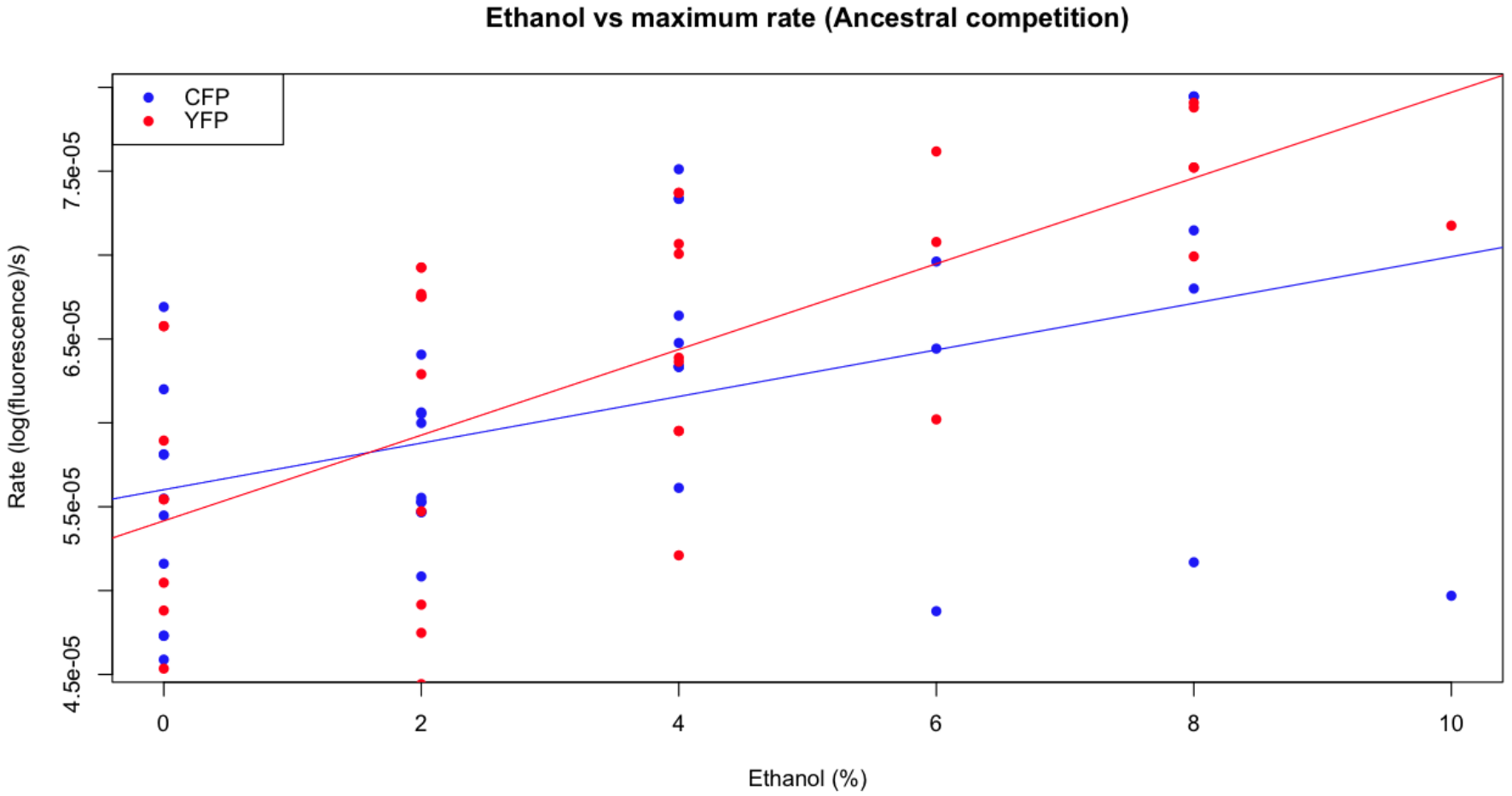


Figure 5.9 A plot of the maximum rate of growth during the competition of ancestral CFP and YFP strains in media with different ethanol concentrations. Competitions were run on the BMG FluoStar plate reader. The points in blue are for the CFP strain and the points in red are for the YFP strain.

Ancestral strains: fluorescence at 36 hours

There is a small upward gradient in the fluorescence at 36 hours with ethanol, perhaps reflecting the increased amount of carbon in the medium. There is no clear difference between the strains (see Figure 5.10).

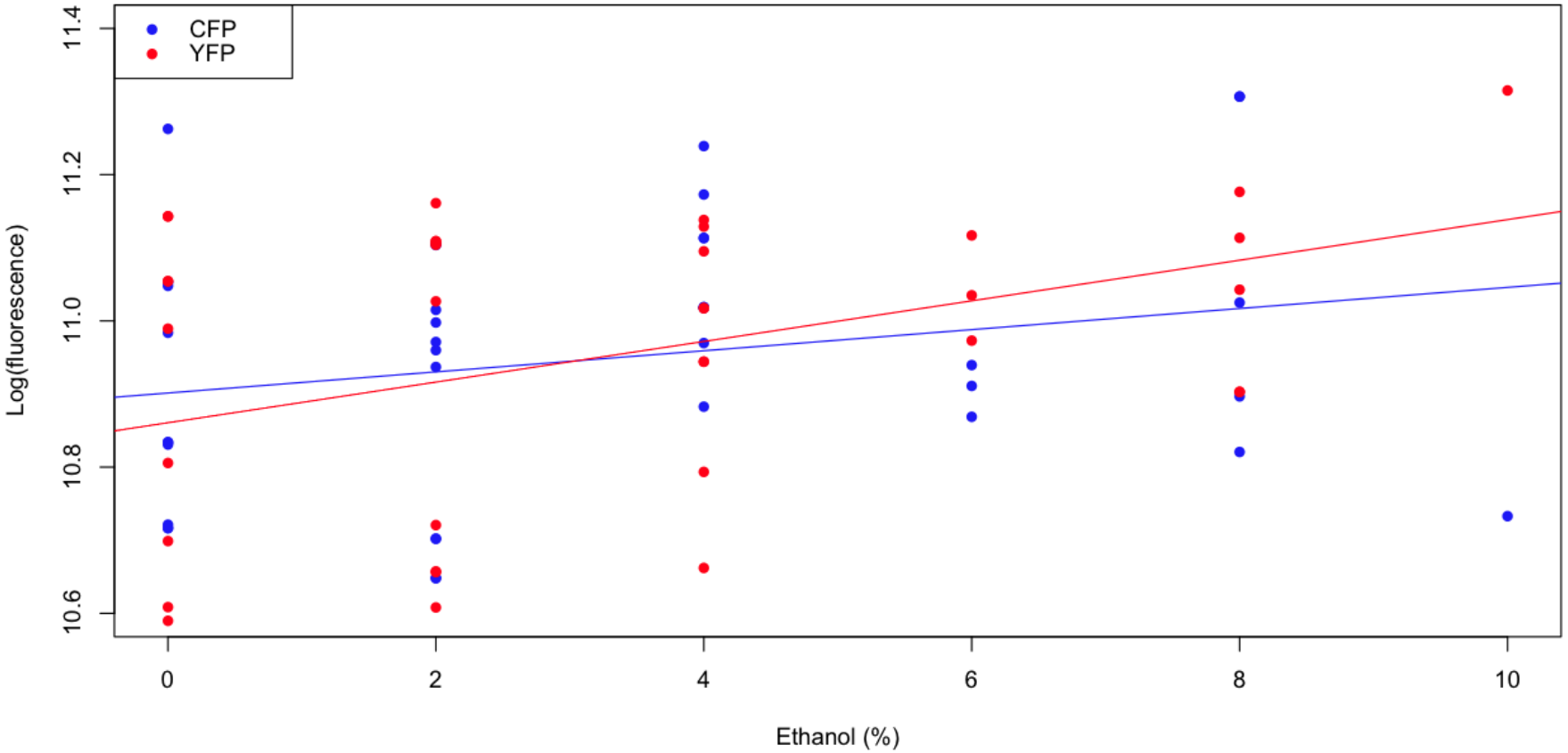


Figure 5.10 A plot of the fluorescence at 36 hours during the competition of ancestral CFP and YFP strains in media with different ethanol concentrations. Competitions were run on the BMG FluoStar plate reader. The points in blue are for the CFP strain and the points in red are for the YFP strain.

Ancestral growth summary

The ancestral YFP consistently emerges from lag 2.5 hours earlier than the ancestral CFP across the ethanol stress range and enters diauxy 1.2 hours earlier. However, there is considerable overlap between growth rates and the fluorescence at 36 hours across the range of added ethanol.

5.3.2. Evolved strain analysis: evolution 6

These strains were evolved in a constant ethanol stress of 4.5% for eleven plates with an estimated bottleneck size of 2000 cells (see Chapter 4, section 4.4.2 for a discussion of this evolution experiment). Two strains were isolated and sequenced, the CFP tagged strain Evo6_11_B06_CFPe_4 and the YFP strain Evo6_11_D06_YFPe_4. All analysed strains were isolated from plate 11.

Evo6_11_B06_CFPe_4: introduction

The plate analysing the competitive fitness of Evo6_11_B06_CFPe_4 in different ethanol concentrations was run on the BMG Omega plate reader.

Evo6_11_B06_CFPe_4: growth lag

The onset of detectable growth is earlier in the Evo6_11_B06_CFPe_4 (intercept=24096s+/-1564se.) than in its ancestral YFP competitor (Intercept=35222s+/-1082se.), a reduction of approximately 3 hours (see Figure 5.11). The slope of the evolved strain and its YFP competitor are very similar. Without an ethanol stress, the onset of detectable growth by the evolved CFP strain is very similar to the ancestral strain. However, with more ethanol stress its growth onset is better than the ancestor.

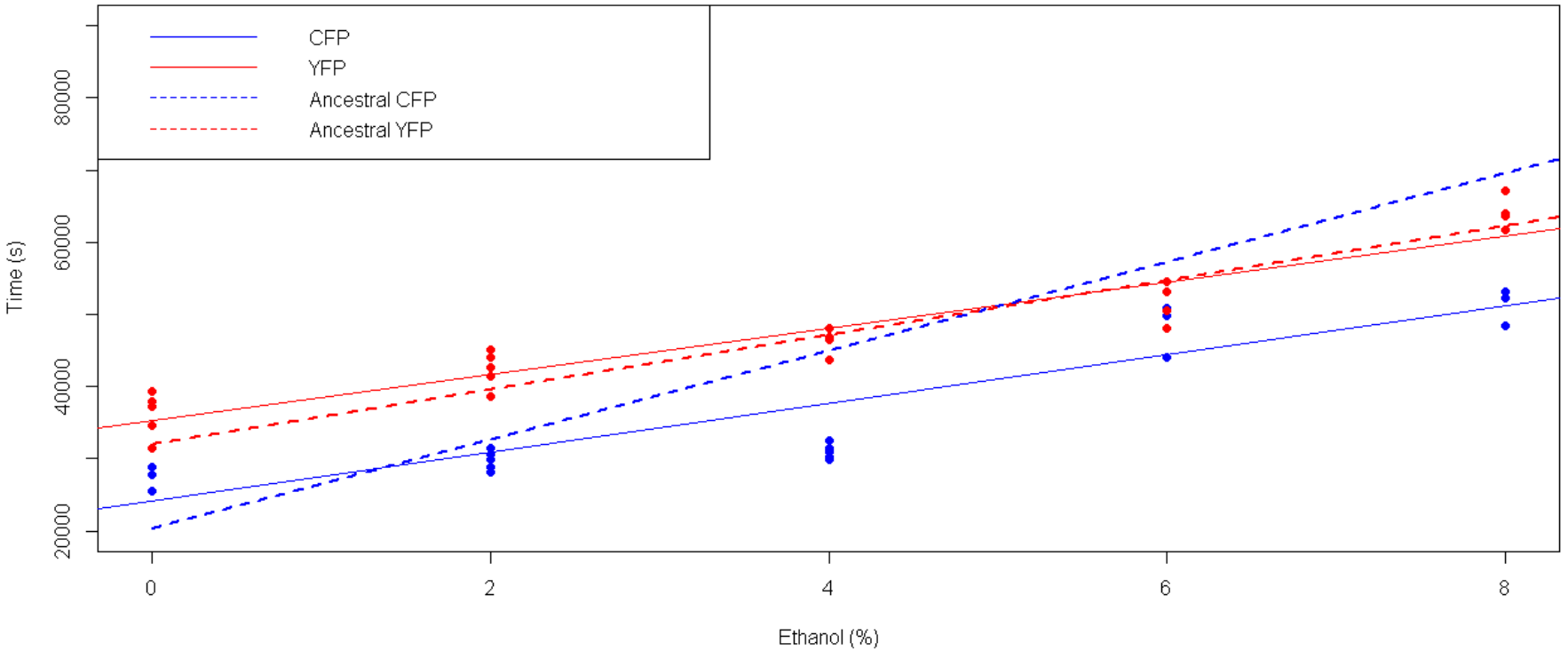


Figure 5.11 A plot of the lag phase before the onset of detectable growth during the competition of Evo6_11_B06_CFPe_4 and ancestral YFP strains in media with different ethanol concentrations.

The points in blue are for the Evo6_11_B06_CFPe_4 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the lag data for Evo6_11_B06_CFPe_4 and the ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG Omega plate reader.

Evo6_11_B06_CFPe_4: onset of diauxy

The onset of diauxy is very similar for both of the competing strains (see Figure 5.12), suggesting both are responding to the exhaustion of glucose rather than terminating fermentation prematurely due to ethanol stress. At ethanol concentrations above 2% the onset of diauxy is sooner than that in the ancestral competition.

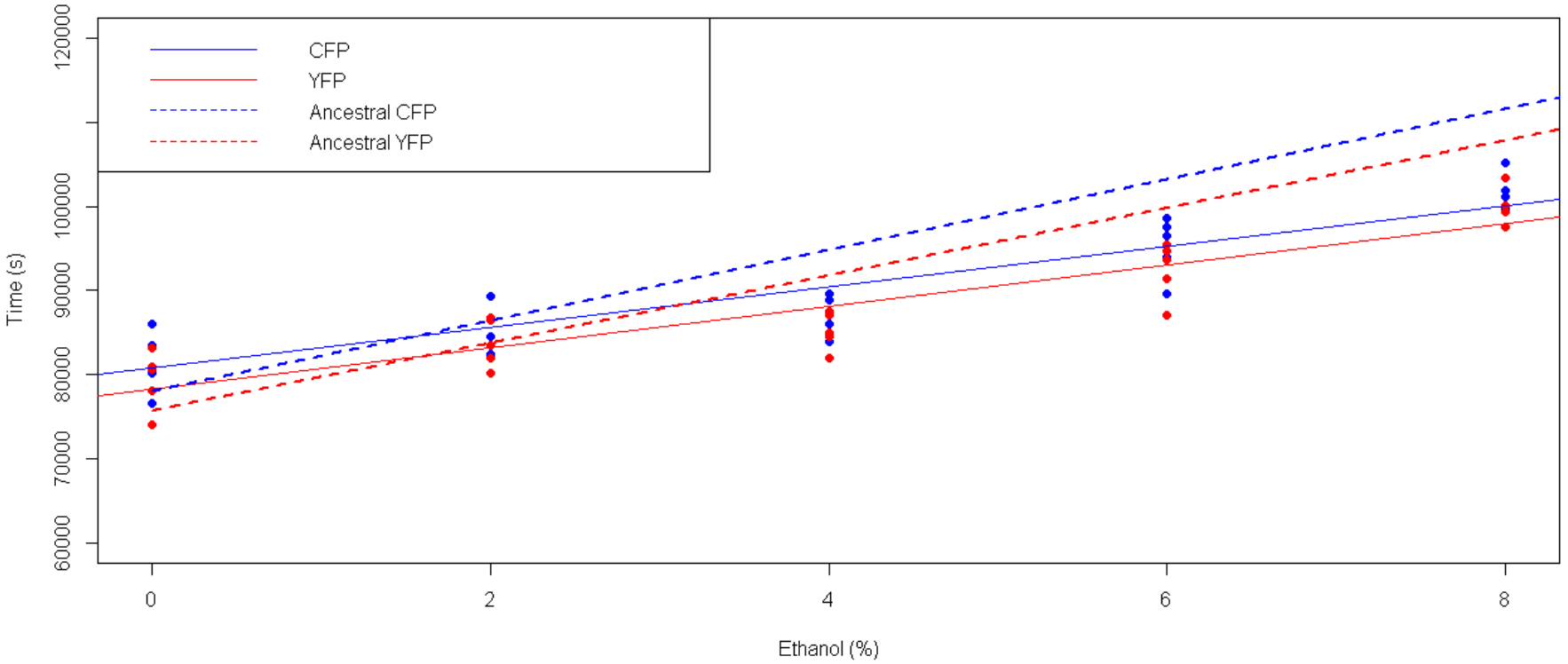


Figure 5.12 A plot of the time of onset of diauxy during the competition of Evo6_11_B06_CFPe_4 and ancestral YFP strains in media with different ethanol concentrations.
The points in blue are for the Evo6_11_B06_CFPe_4 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the diauxy data for Evo6_11_B06_CFPe_4 and the ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG Omega plate reader.

Evo6_11_B06_CFPe_4: growth duration

The combined effect of earlier onset and similar diauxy is that Evo6_11_B06_CFPe_4 grows rapidly for longer than the ancestral YFP strain in competition with it (see Figure 5.13). The duration of rapid growth is shorter for the ancestral YFP when it competes with Evo6_11_B06_CFPe_4 (see Figure 5.13) when ethanol is added to the growth medium.

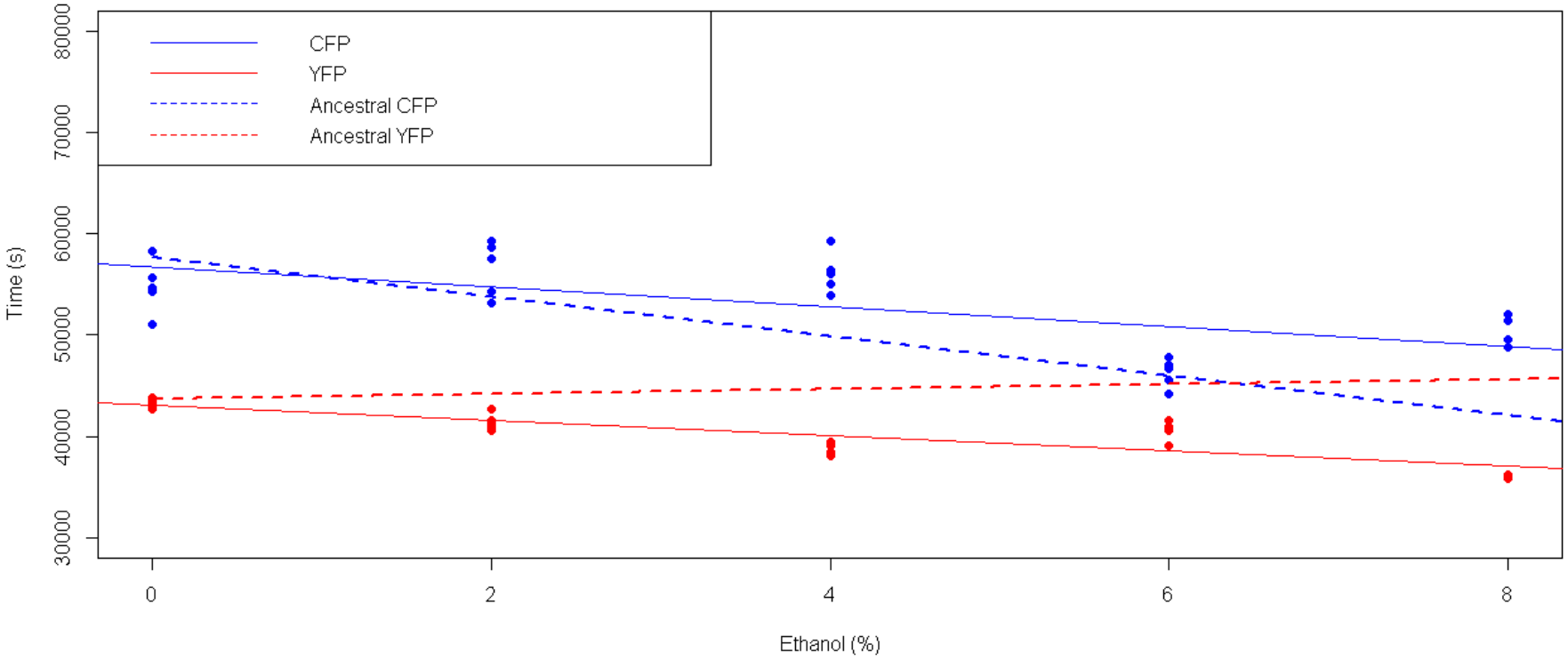


Figure 5.13 A plot of the duration of growth during the competition of Evo6_11_B06_CFPe_4 and ancestral YFP strains in media with different ethanol concentrations.
The points in blue are for the Evo6_11_B06_CFPe_4 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the growth duration data for Evo6_11_B06_CFPe_4 and the ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG Omega plate reader.

Evo6_11_B06_CFPe_4: maximum growth rate

The maximum rate of growth for both strains in the evolved competition is apparently lower than in the ancestral competition (see Figure 5.14). However, the rate of growth of Evo6_11_B06_CFPe_4 is substantially higher than its YFP competitor and the rate of Evo6_11_B06_CFPe_4 has a small positive slope with increasing ethanol stress.

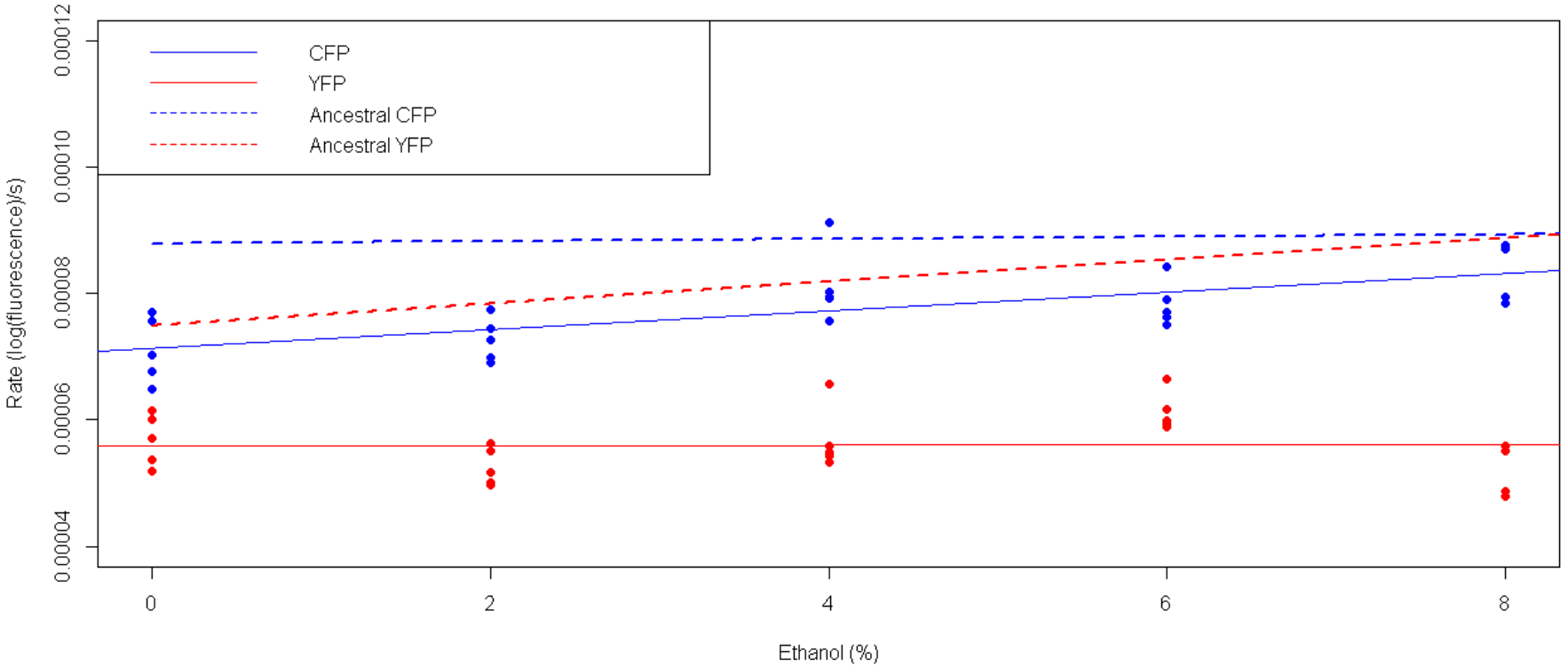


Figure 5.14 A plot of the maximum rate of growth during the competition of Evo6_11_B06_CFPe_4 and ancestral YFP strains in media with different ethanol concentrations.
The points in blue are for the Evo6_11_B06_CFPe_4 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the growth data for Evo6_11_B06_CFPe_4 and the ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG Omega plate reader.

Evo6_11_B06_CFPe_4: fluorescence at 36 hours

As measured by the fluorescence at 36 hours, Evo6_11_B06_CFPe_4 grows as well as the ancestral CFP in competition (perhaps slightly better with 6-8% ethanol). The growth of the ancestral YFP is poorer when competing with Evo6_11_B06_CFPe_4 than the ancestral CFP, a growth deficit that is maintained, and may slightly increase, as more ethanol is added to the growth medium.

194 of **424**

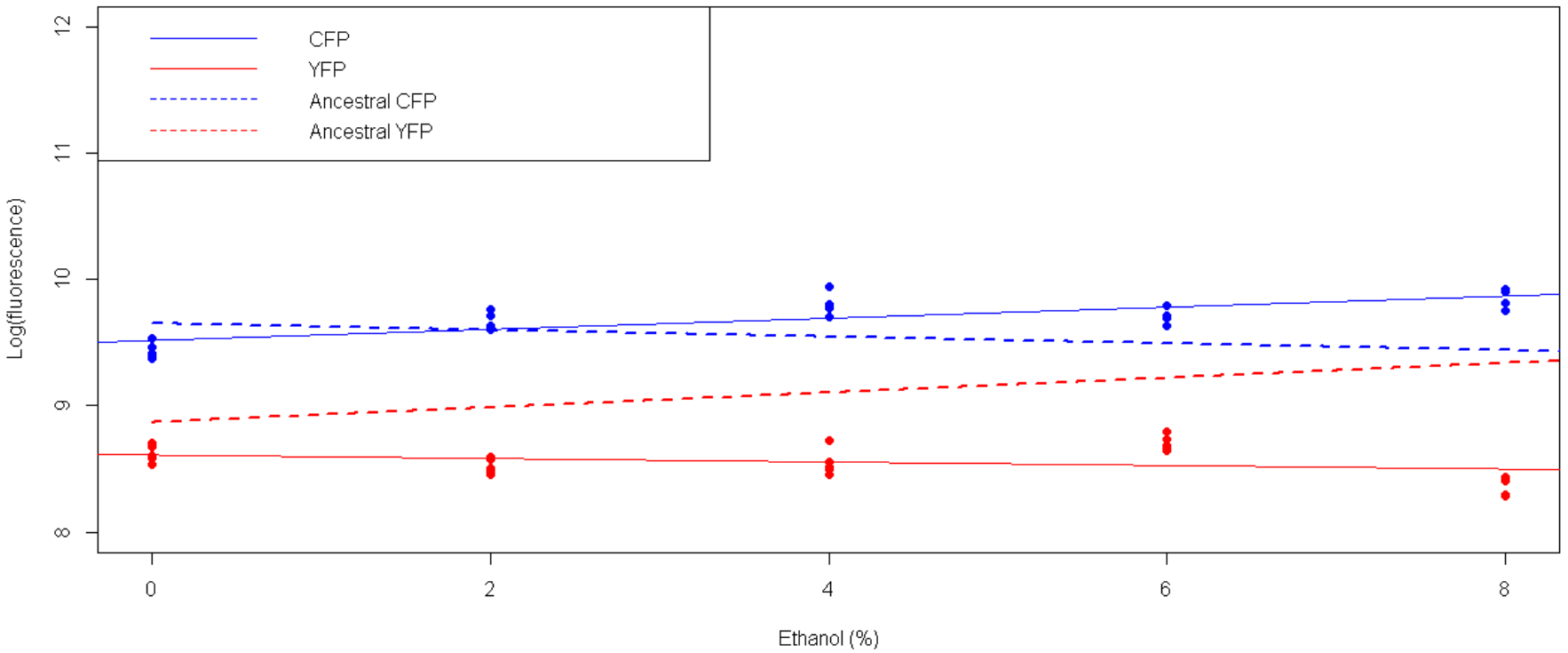


Figure 5.15 A plot of the fluorescence at 36 hours during the competition of Evo6_11_B06_CFPe_4 and ancestral YFP strains in media with different ethanol concentrations.
The points in blue are for the Evo6_11_B06_CFPe_4 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the 36-hour fluorescence for Evo6_11_B06_CFPe_4 and the ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG Omega plate reader.

Evo6_11_B06_CFPe_4: summary

Evo6_11_B06_CFPe_4 emerges from lag sooner than an ancestral YFP competitor enabling Evo6_11_B06_CFPe_4 to grow for longer. Evo6_11_B06_CFPe_4 also has a higher maximum rate of growth than its ancestral competitor, though its maximum growth rate is less than the rate shown by either of the ancestral strains when they are in competition with each other. The fluorescence at 36 hours of Evo6_11_B06_CFPe_4 shows a small increase with ethanol stress between 0 to 8% while there is a small decline in the corresponding fluorescence of its ancestral competitor.

Evo6_11_D06_YFPe_4: Introduction

The plate analysing the competitive fitness of Evo6_11_D06_YFPe_4 in different ethanol concentrations was run on the BMG FluoStar plate reader. All evolution 6 populations were subjected to an added ethanol stress of 4.5%

Evo6_11_D06_YFPe_4: growth lag

For Evo6_11_D06_YFPe_4 the slope of the time to detectable growth is very similar to that of the ancestors and the ancestral CFP competitor. However, there is a reduction in the intercept for Evo6_11_D06_YFPe_4 (26315s+/-884s.e.) compared with the ancestral CFP strain (43705s+/-1529s.e.) of approximately four hours and fifty minutes (see Figure 5.16).

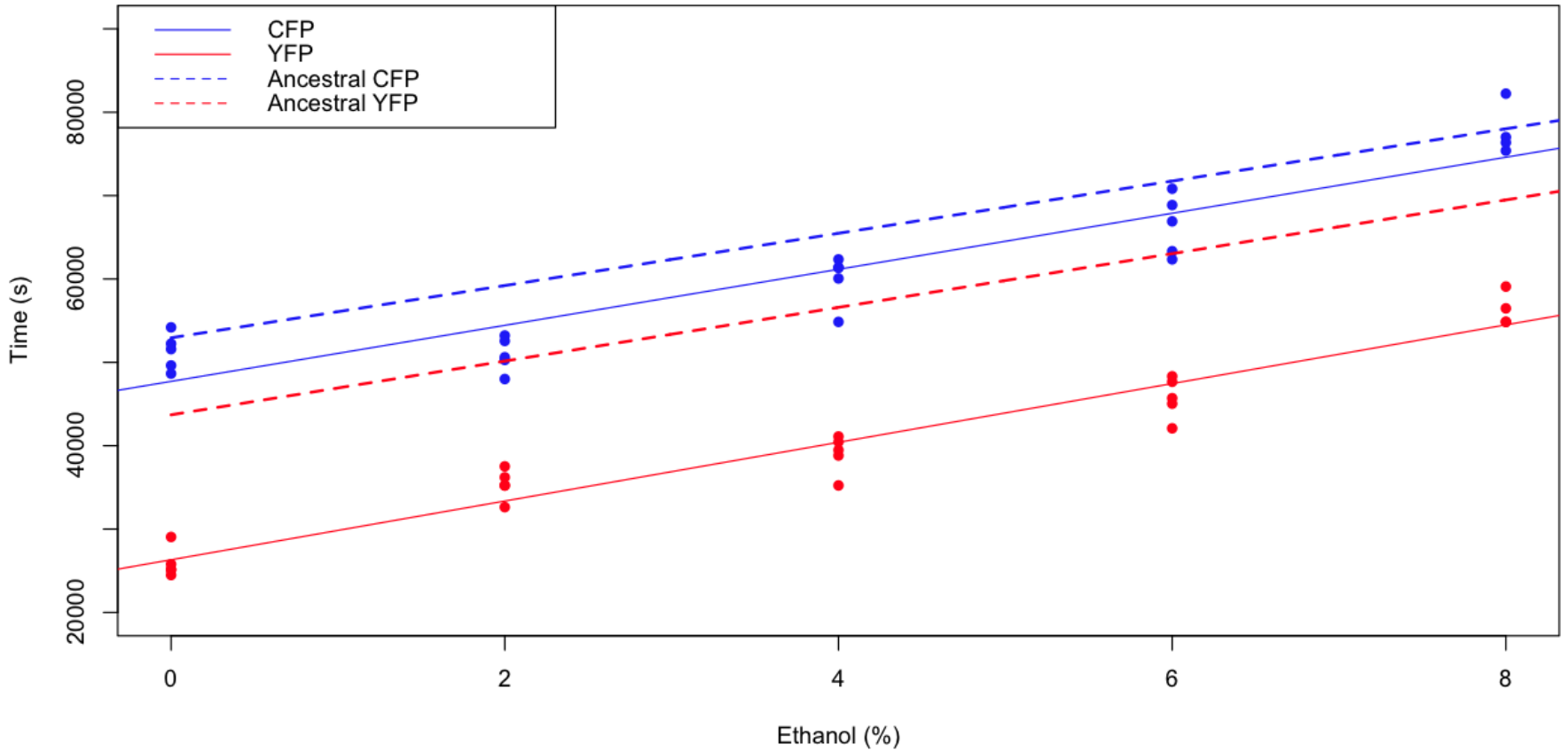


Figure 5.16 A plot of the lag phase before the onset of detectable growth during the competition of the ancestral CFP and Evo6_11_D06_YFPe_4 strains in media with different ethanol concentrations.
The points in blue are for the ancestral CFP and the points in red are for the Evo6_11_D06_YFPe_4 strain. The blue and red solid lines represent the linear best fit for the lag data for the ancestral CFP and Evo6_11_D06_YFPe_4 respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo6_11_D06_YFPe_4: onset of diauxy

The corresponding time to the onset of diauxy is reduced for both strains. The slopes of the evolved strain (3523s/% +/-187s.e.) and CFP ancestor (3362s/% +/-287s.e.) are not substantially different. There is a similar difference in the intercepts of the evolved strain and ancestor but this is of similar magnitude to the difference seen in the ancestors (see Figure 5.17).

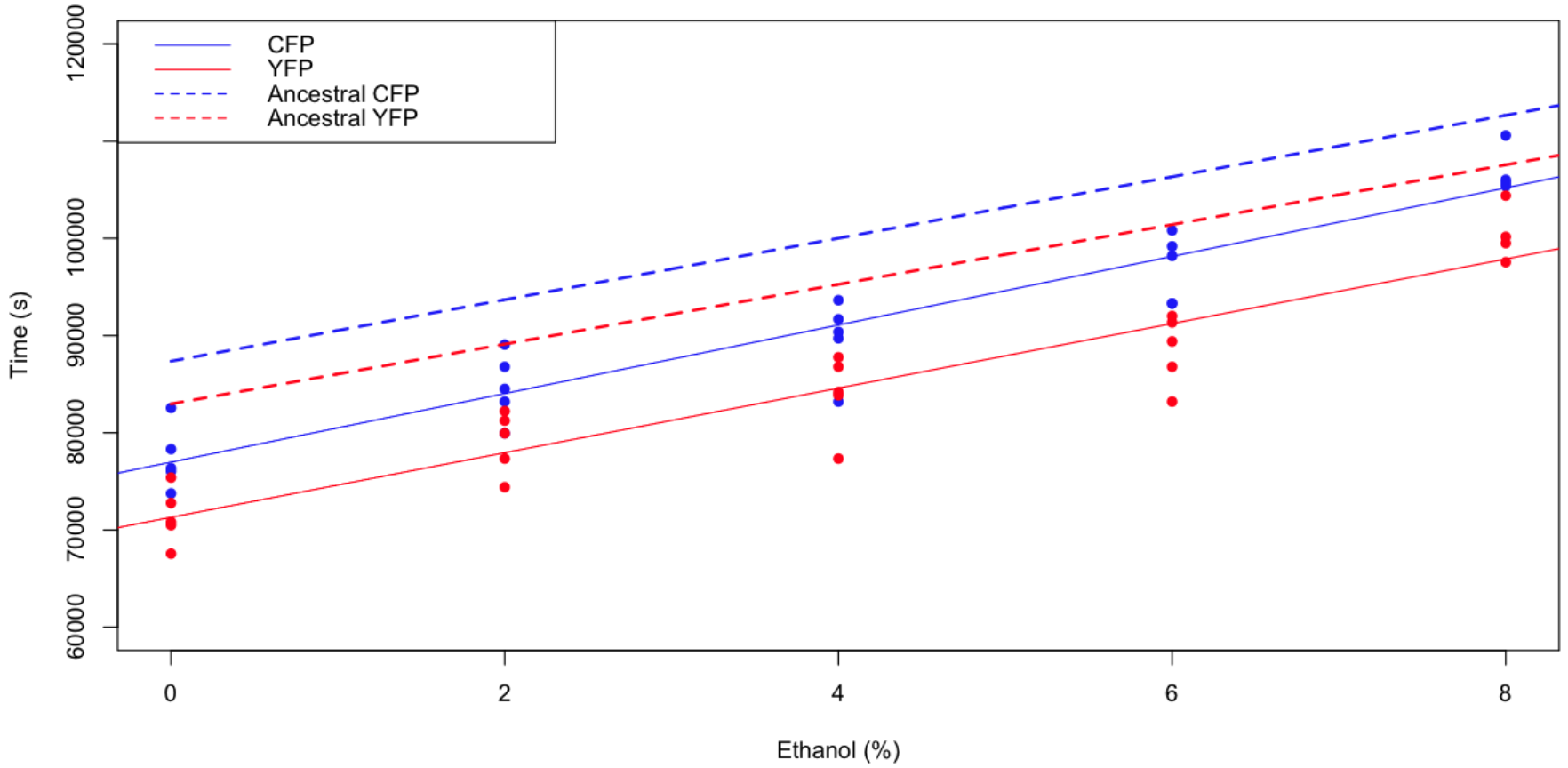


Figure 5.17 A plot of the time of onset of diauxy during the competition of the ancestral CFP and CompEvo6_11_D06_YFPe_4 strains in media with different ethanol concentrations.
The points in blue are for the ancestral CFP and the points in red are for the CompEvo6_11_D06_YFPe_4 strain. The blue and red solid lines represent the linear best fit for the onset of diauxy for the ancestral CFP and CompEvo6_11_D06_YFPe_4 respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo6_11_D06_YFPe_4: growth duration

As a result of the earlier onset of detectable growth, Evo6_11_D06_YFPe_4 grows for approximately four and a half hours longer than the competing ancestral CFP strain (see Figure 5.18) and just under an hour and a half longer than the ancestral YFP strain in the ancestral competition. The time the ancestral CFP spent growing rapidly is shorter when competing with Evo6_11_D06_YFPe_4 than with the ancestral YFP.

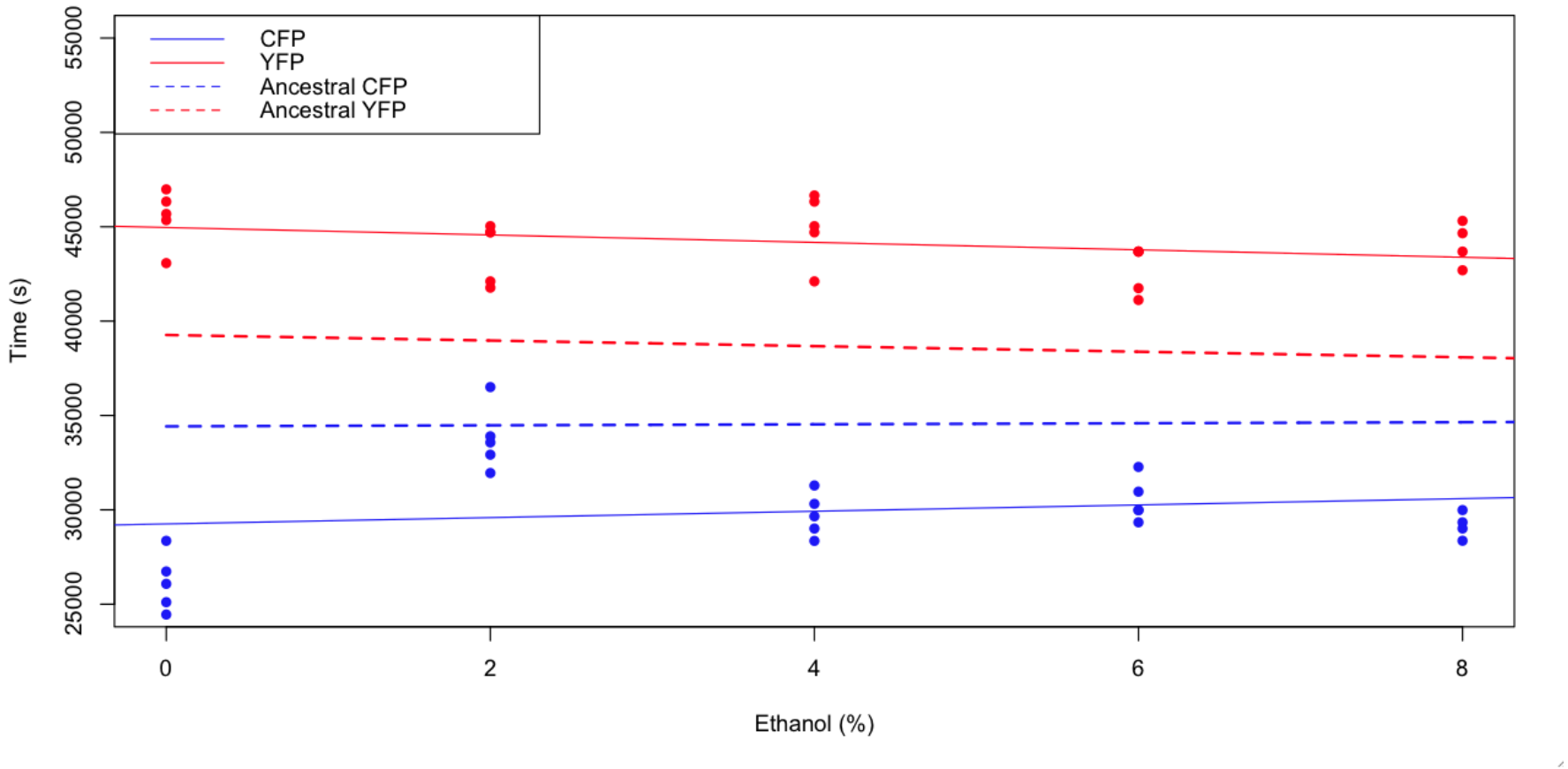


Figure 5.18 A plot of the duration of growth during the competition of the ancestral CFP and CompEvo6_11_D06_YFPe_4 strains in media with different ethanol concentrations.
The points in blue are for the ancestral CFP and the points in red are for the CompEvo6_11_D06_YFPe_4 strain. The blue and red solid lines represent the linear best fit for the growth duration data for the ancestral CFP and CompEvo6_11_D06_YFPe_4 respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo6_11_D06_YFPe_4: maximum growth rate

The maximum rate of growth of Evo6_11_D06_YFPe_4 is lower than the best fit of all ancestral rates but that of Evo6_11_D06_YFPe_4 is higher at all ethanol concentrations than its CFP ancestral competitor (see Figure 5.19 below).

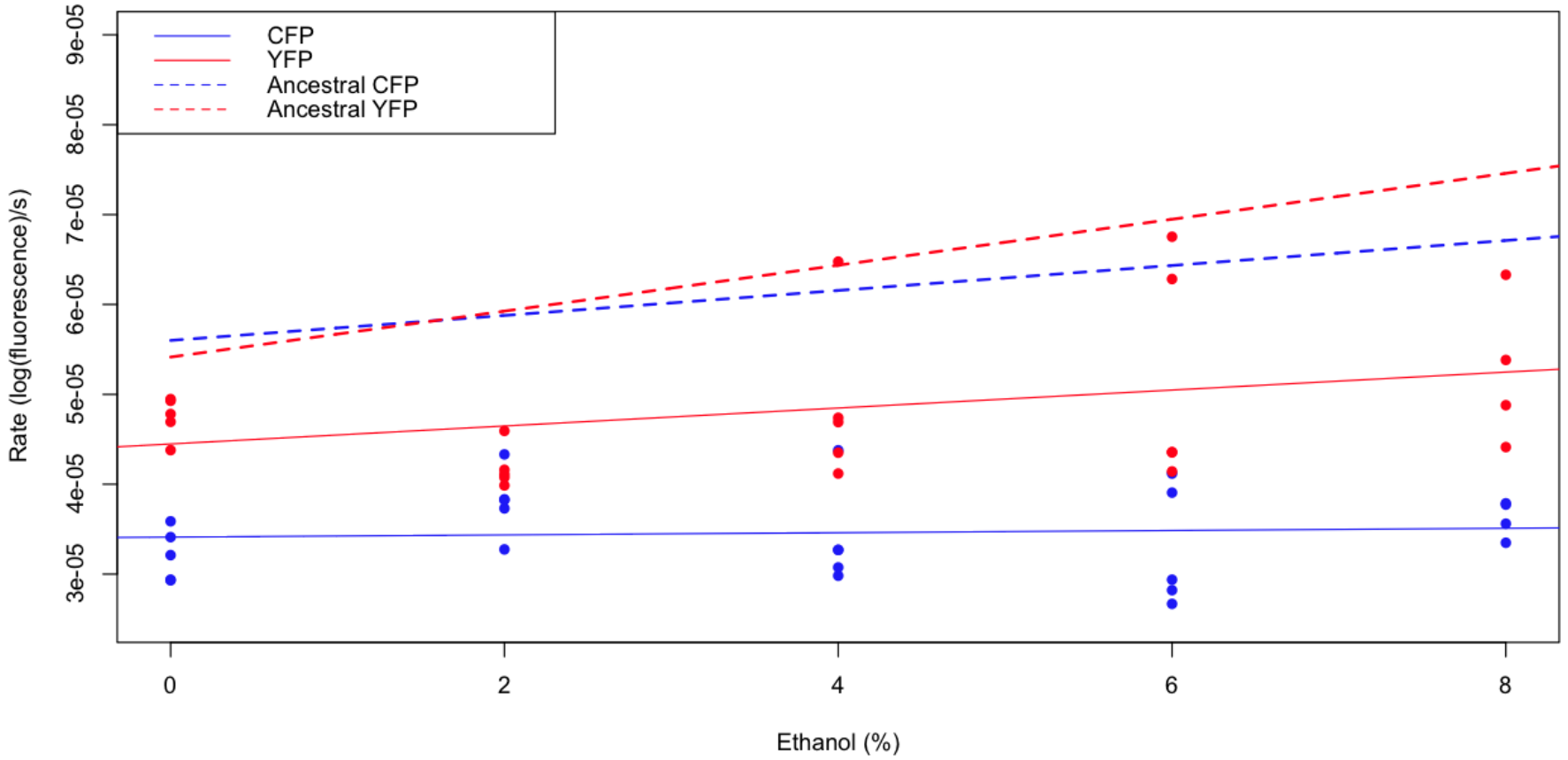


Figure 5.19 A plot of the maximum growth rate during the competition of the ancestral CFP and CompEvo6_11_D06_YFPe_4 strains in media with different ethanol concentrations.
The points in blue are for the ancestral CFP and the points in red are for the CompEvo6_11_D06_YFPe_4 strain. The blue and red solid lines represent the linear best fit for the growth rate data for the ancestral CFP and CompEvo6_11_D06_YFPe_4 respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo6_11_D06_YFPe_4: fluorescence at 36 hours

The fluorescence of Evo6_11_D06_YFPe_4 at 36 hours is very close to the best fit of the ancestors but the CFP fluorescence is only a little more than half that of Evo6_11_D06_YFPe_4 (see Figure 5.20); the log intercept of the CFP regression is 10.25+/-0.059s.e while that Evo6_11_D06_YFPe_4 is 10.78+/-0.046s.e.; fluorescence values of $2.8x10^3$ vs $4.8x10^3$. In competition with the evolved strain, the CFP fluorescence slope is not significantly different from zero (p=0.351) but the evolved strain YFP fluorescence slope is (0.049976+/-0.009701s.e., p=3.65e-05).

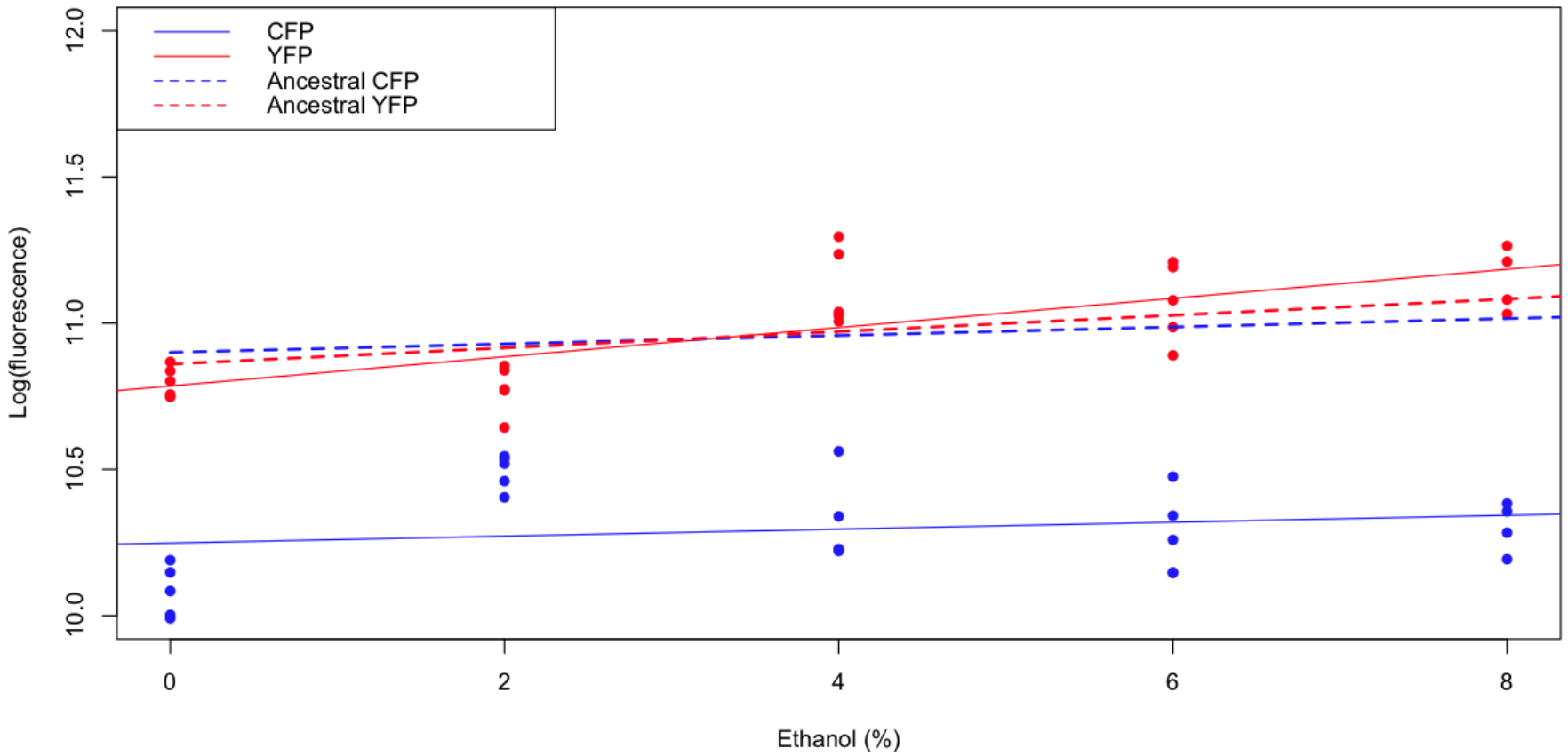


Figure 5.20 A plot of the fluorescence at 36 hours during the competition of the ancestral CFP and CompEvo6_11_D06_YFPe_4 strains in media with different ethanol concentrations.
The points in blue are for the ancestral CFP and the points in red are for the CompEvo6_11_D06_YFPe_4 strain. The blue and red solid lines represent the linear best fit for the 36-hour fluorescence for the ancestral CFP and CompEvo6_11_D06_YFPe_4 respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo6_11_D06_YFPe_4: summary

Evo6_11_D06_YFPe_4 grows for longer and at a higher maximum rate than its ancestral competitor (though at a lower maximum rate than either ancestral strain competing together). Evo6_11_D06_YFPe_4 is fitter than the ancestral CFP competitor at all ethanol concentrations up to 8%.

Evolution 6: summary

Both Evo6_11_B06_CFPe_4 and Evo6_11_D06_YFPe_4 originated from starting strains with different fluorophores and were analysed in different plate readers. Despite this both strains have very similar growth characteristics when competed against the ancestor with the contrasting fluorophore. Both grow for longer because they show detectable growth ~4.5 hours before their ancestral competitor: the maximum growth rate of each evolved strain is poorer than their ancestor in competition but the maximum rate of their ancestral competitor is depressed much further. The result is that the evolved strains produce a similar fluorescence at 36 hours to their ancestor but their competitors perform substantially worse.

5.3.3. Evolved strain analysis: evolution 8

These strains were subjected to constant ethanol stresses of 4.5% and 6.5% and had control wells with no added ethanol. From the 0% ethanol control wells, one strain, Evo8_14_CFPe_E07_4 was isolated from. All analysed strains were isolated from plate 14.

Evo8_14_CFPe_E07 _4 (0% ethanol): Introduction

Evo8_14_CFPe_E07 _4 was isolated from a control population that had not been subjected to ethanol stress. The plate analysing competitive fitness in different ethanol concentrations was run on the BMG Omega plate reader; the plate lacked competitions between the two ancestral strains.

Evo8_14_CFPe_E07 _4 (0% ethanol): growth lag

The onset of detectable growth for the evolved strain Evo8_14_CFPe_E07_4 and its ancestral YFP competitor are almost identical. Growth onset is approximately $2x10^4$s (~5.5 hours) later for the evolved strain than the CFP strain in ancestral competition (see Figure 5.21).

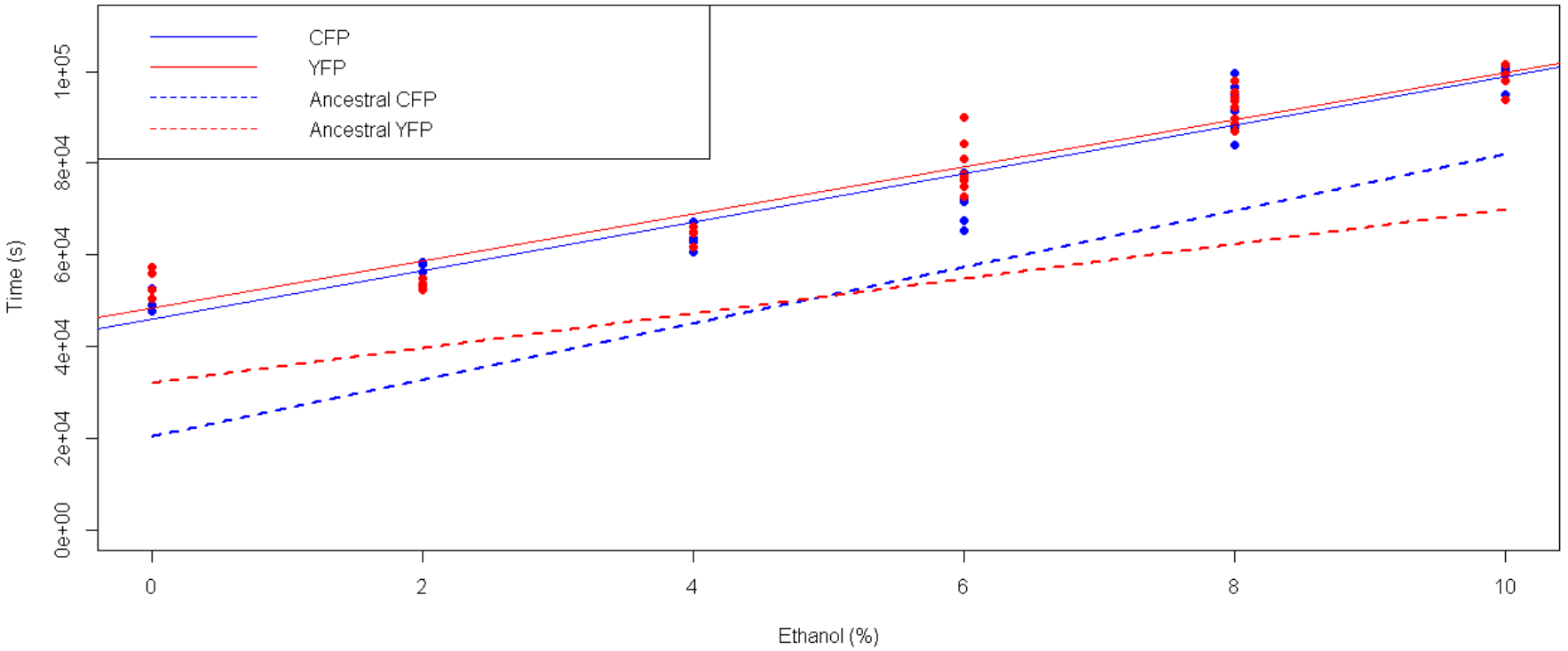


Figure 5.21 A plot of the lag phase before the onset of detectable growth during the competition of Evo8_14_E07_CFPe_4 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_E07_CFPe_4 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the lag data for Evo8_14_E07_CFPe_4 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

Evo8_14_CFPe_E07 _4 (0% ethanol): onset of diauxy

There is no perceptible difference in the onset of diauxy for either strain in the evolved competition. Diauxy also occurs later with a slightly greater slope than the ancestral competition; it is some ~4 hours later than the ancestral competition without added ethanol rising to ~5.5 hours later with 10% ethanol stress (see Figure 5.22).

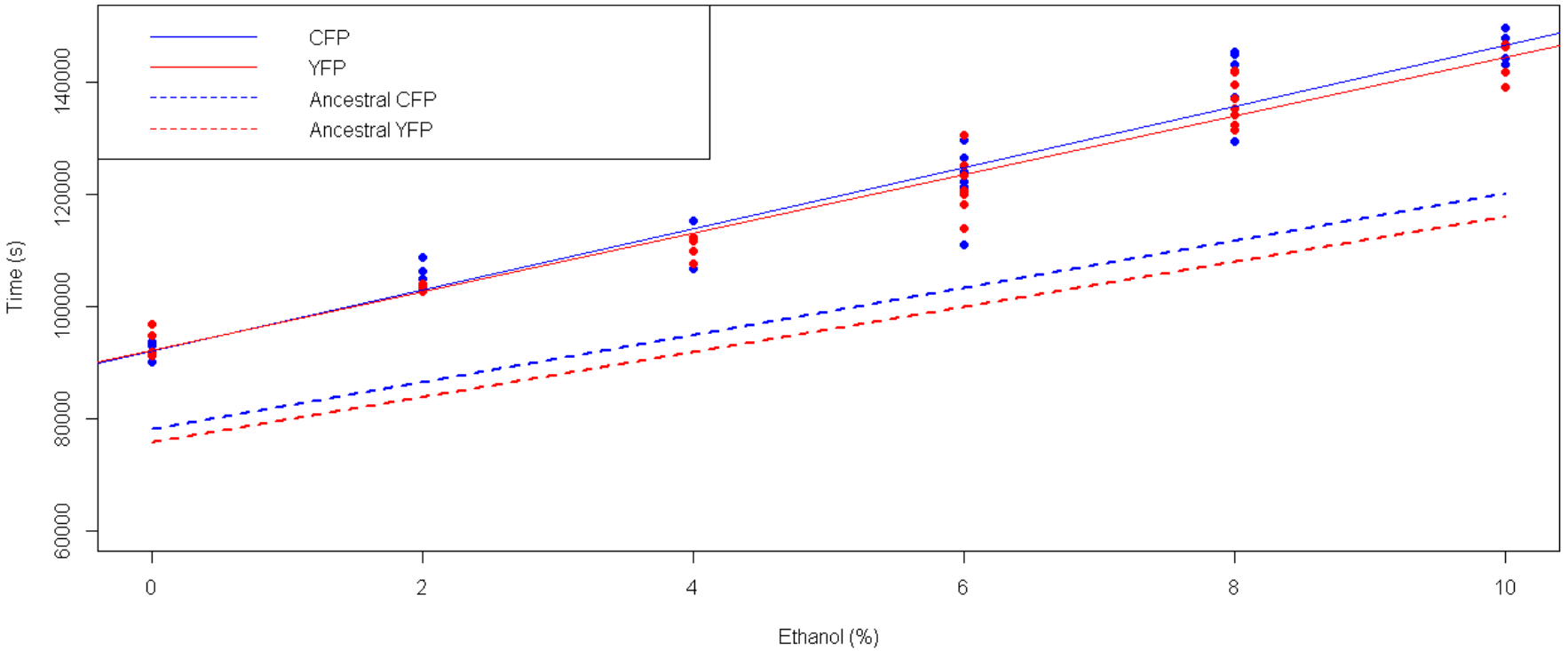


Figure 5.22 A plot of the onset of diauxy during the competition of Evo8_14_E07_CFPe_4 and ancestral YFP strains in media with different ethanol concentrations. Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_E07_CFPe_4 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the diauxy data for Evo8_14_E07_CFPe_4 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

Evo8_14_CFPe_E07 _4 (0% ethanol): growth duration

Both the onset of growth and diauxy are shifted later by similar amounts. The time of growth that results for each strain in the evolved competition is very similar to that of the YFP strain in the ancestral competition (see Figure 5.23).

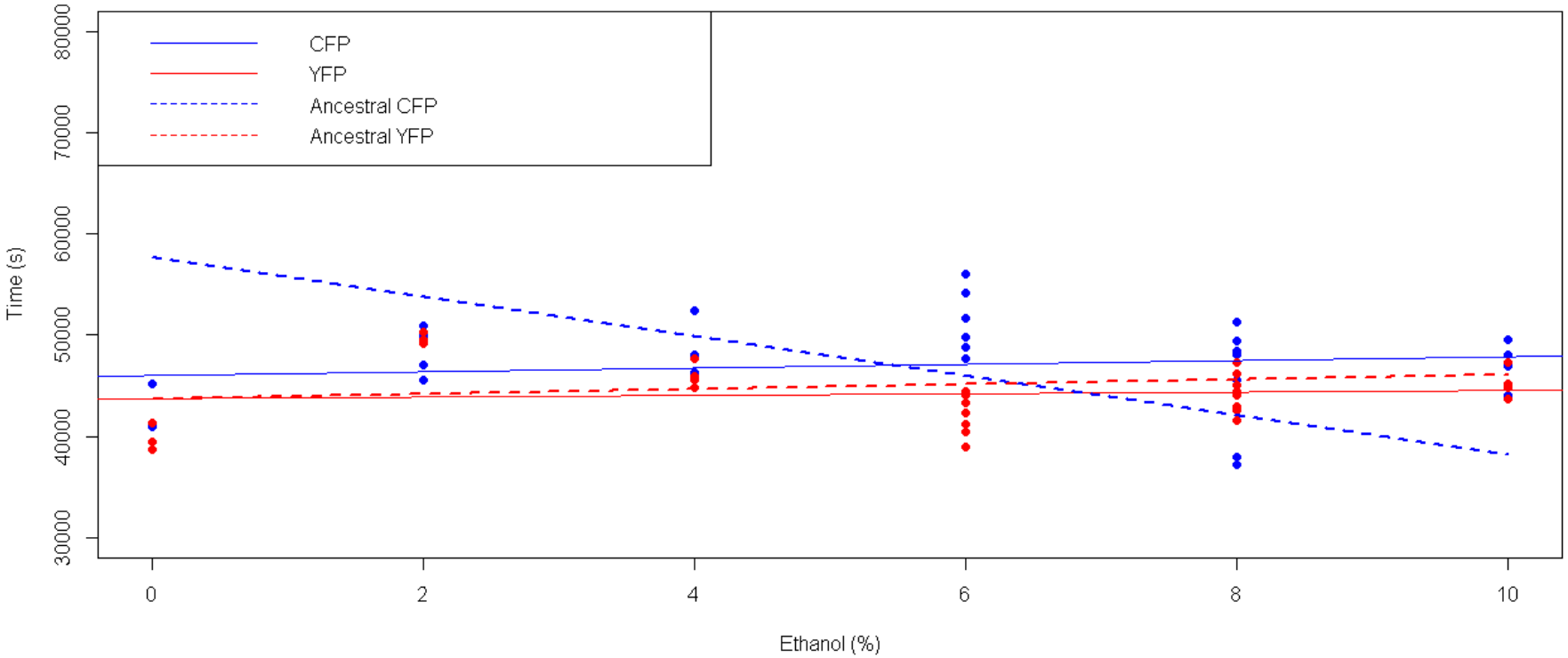


Figure 5.23 A plot of the duration of growth during the competition of Evo8_14_E07_CFPe_4 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_E07_CFPe_4 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the growth duration data for Evo8_14_E07_CFPe_4 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

Evo8_14_CFPe_E07 _4 (0% ethanol): maximum growth rate

The maximum growth rate of the evolved strain is slightly less than the CFP strain in the ancestral strain but the rate of the YFP competitor is a lot lower (see Figure 5.24). Neither strain in the evolved competition has a significant slope, though the ancestral YFP is close (CFP slope=4.739e-07+/-3.679e-07, p=0.208;YFP slope= 5.131e-07+/-2.538e-07se.,p=0.0522).

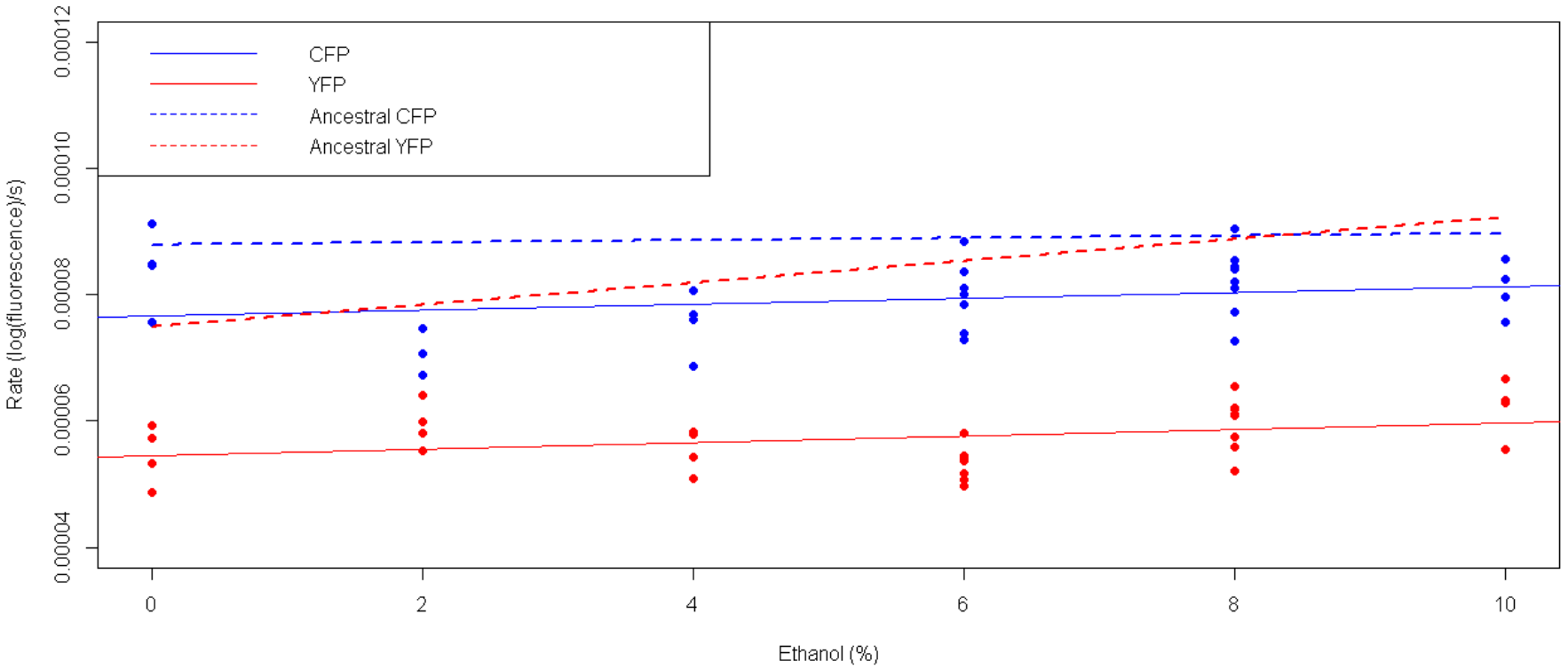


Figure 5.24 A plot of the maximum rate of growth during the competition of Evo8_14_E07_CFPe_4 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_E07_CFPe_4 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the growth data for Evo8_14_E07_CFPe_4 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

Evo8_14_CFPe_E07 _4 (0% ethanol): fluorescence at 36 hours

The result is an improvement in the growth of the evolved strain at higher concentrations of ethanol and a corresponding impairment of the growth of the YFP competitor under ethanol stress (see Figure 5.25).

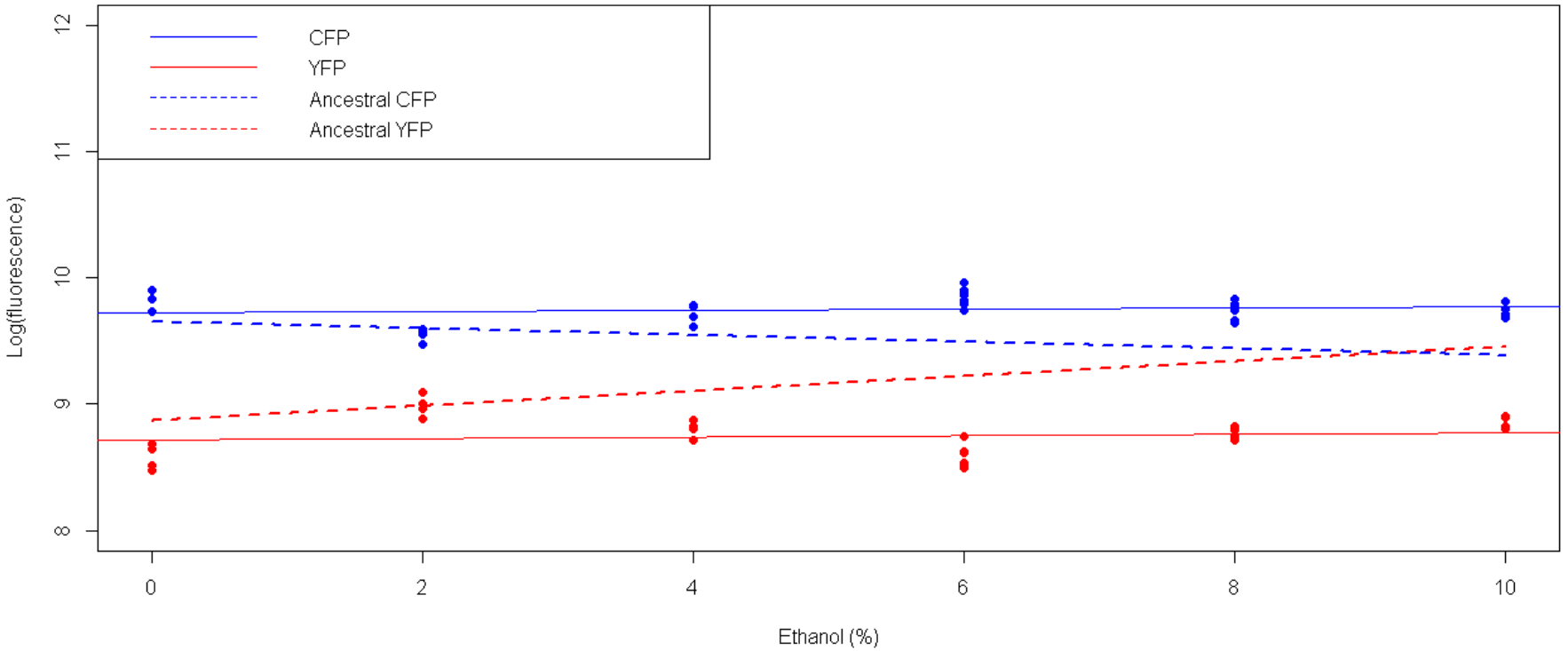


Figure 5.25 A plot of the fluorescence at 36 hours during the competition of Evo8_14_E07_CFPe_4 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_E07_CFPe_4 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the 36-hour fluorescence for Evo8_14_E07_CFPe_4 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

Evo8_14_CFPe_E07 _4 (0% ethanol): summary

The elapsed time of active growth for the evolved strain Evo8_14_E07_CFPe_4 is the same as its YFP competitor, its maximum rate of growth is slightly poorer but the rate of growth of its YFP competitor is substantially depressed; Evo8_14_E07_CFPe_4 outcompetes the YFP ancestor due to its faster growth rate.

Evo8_14_F03_CFPe_2 (4.5% ethanol): Introduction

This strain was isolated from a population that evolved under a 4.5% ethanol stress. The fitness analysis was run in the BMG Omega plate reader; the plate did not include a competition between the two ancestors.

Evo8_14_F03_CFPe_2 (4.5% ethanol): growth lag

The fitness analysis of Evo8_14_F03_CFPe_2 in varying ethanol concentrations shows that without an ethanol stress there is a very small reduction in the time to detect growth in the strain but this diminishes with ethanol stress (see Figure 5.26).

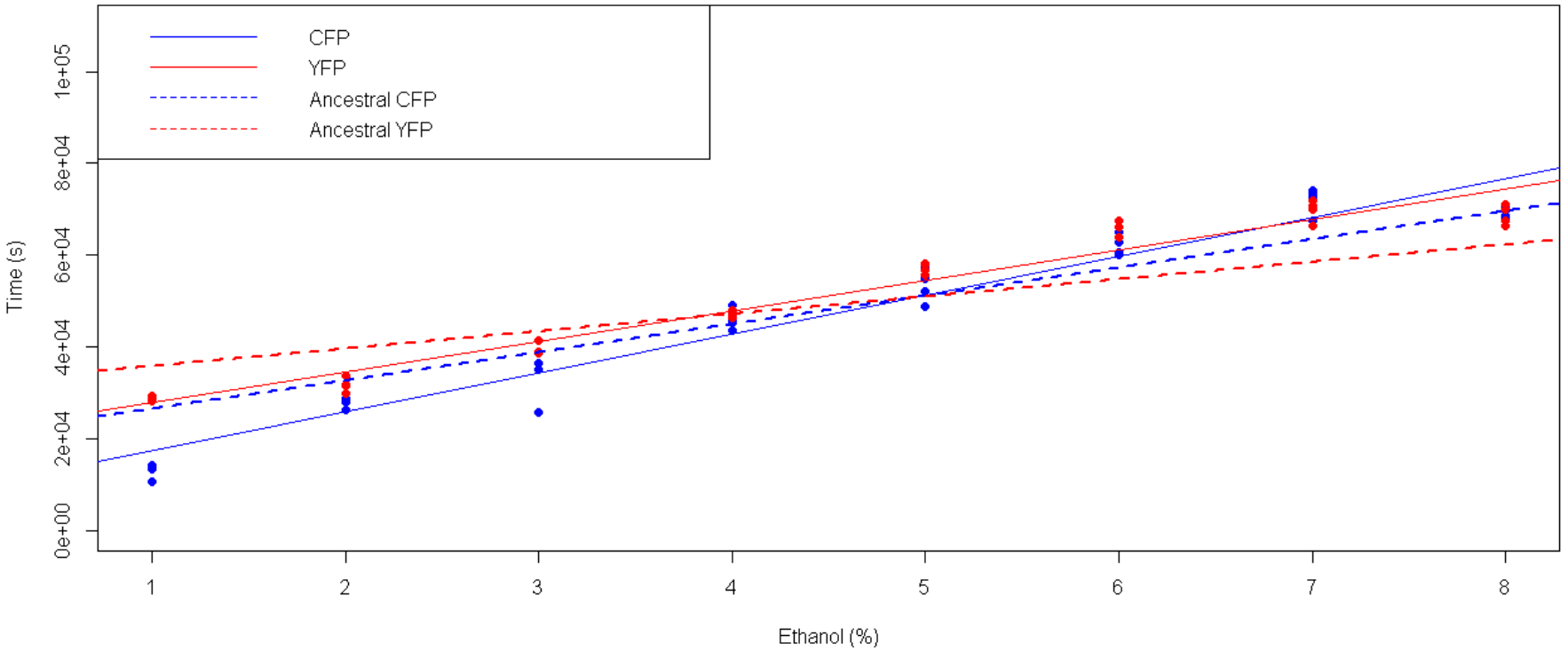


Figure 5.26 A plot of the lag phase before the onset of detectable growth during the competition of Evo8_14_F03_CFPe_2 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_F03_CFPe_2 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the lag data for Evo8_14_F03_CFPe_2 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

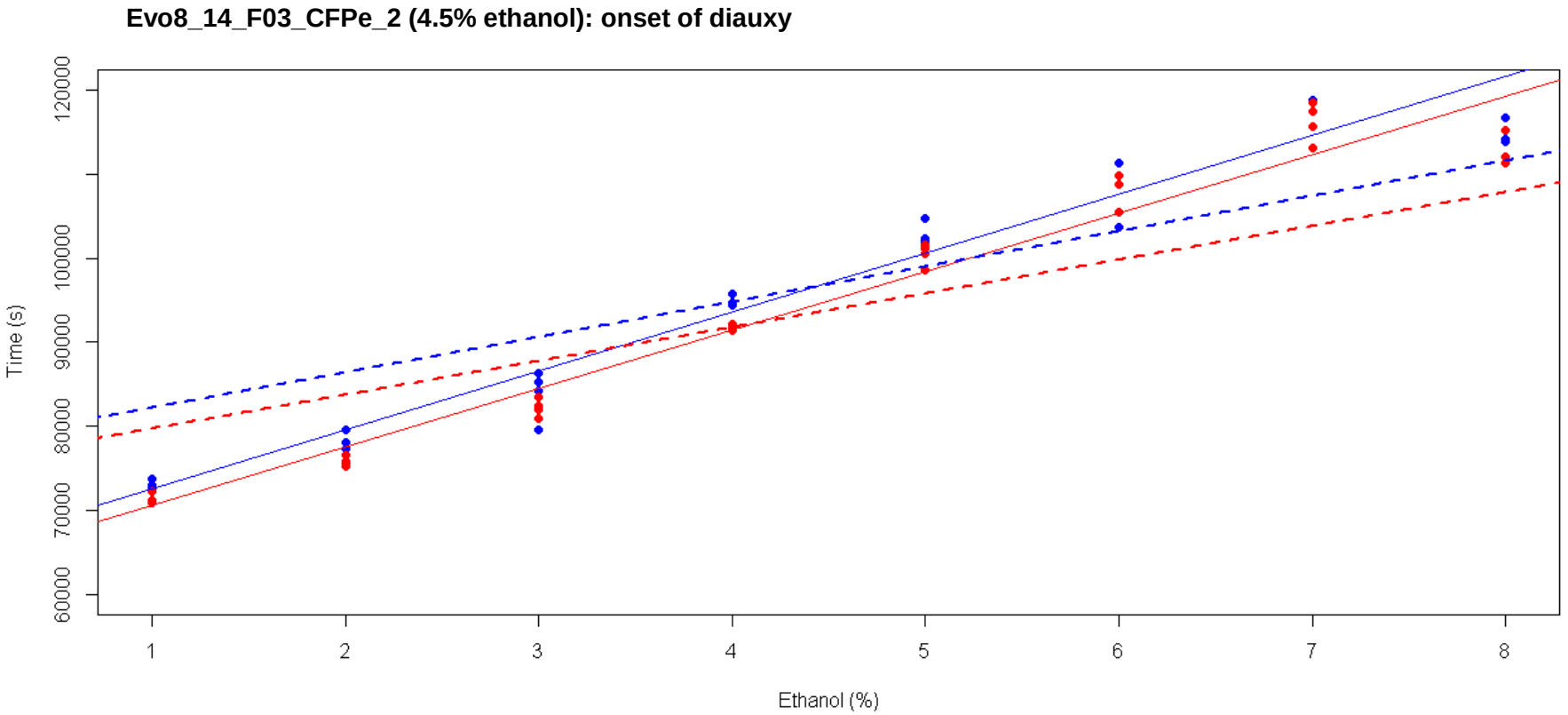


Figure 5.27 A plot of the time to onset of diauxy during the competition of Evo8_14_F03_CFPe_2 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_F03_CFPe_2 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the diauxy data for Evo8_14_F03_CFPe_2 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

Evo8_14_F03_CFPe_2 (4.5% ethanol): growth duration

Despite minor differences in the patterns of growth onset and diauxy, the elapsed time of growth is almost identical to the ancestral values derived from other plates (see Figure 5.28).

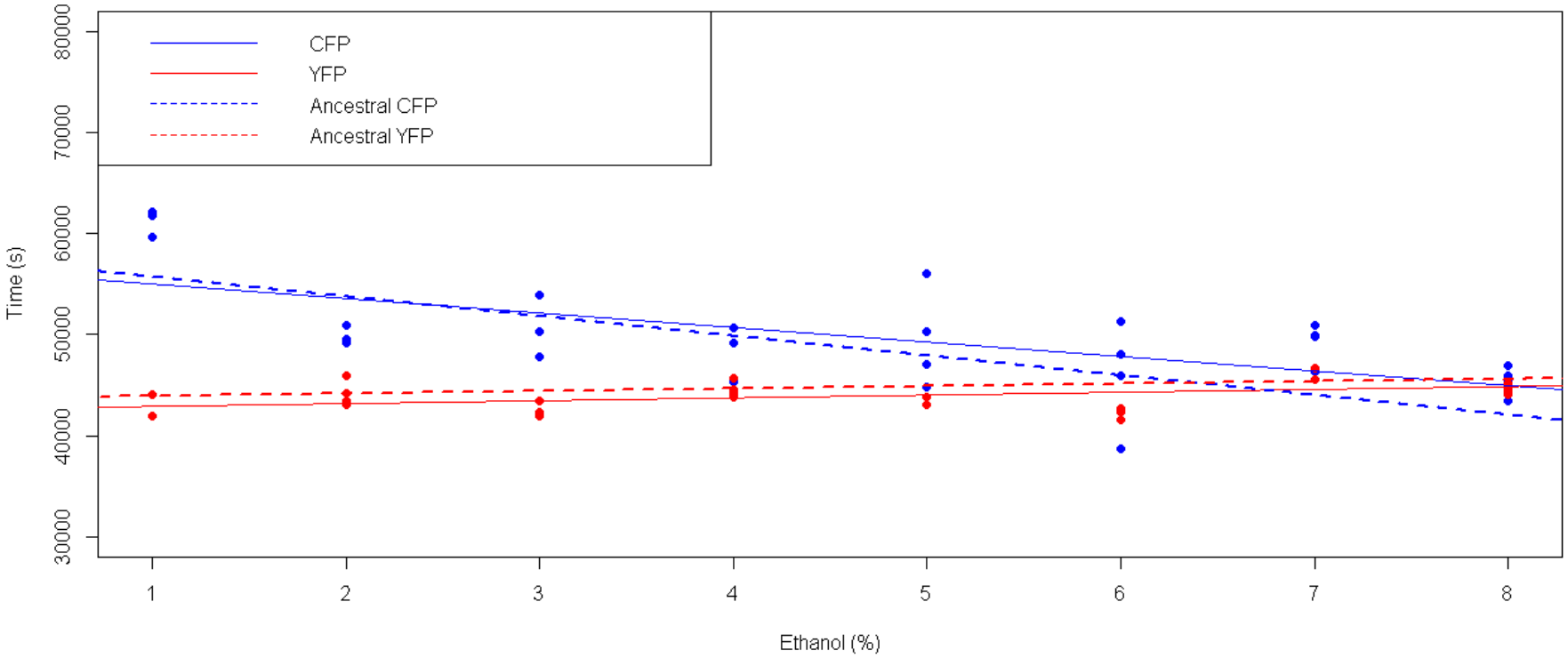


Figure 5.28 A plot of the time elapsed between detectable growth and onset of diauxy during the competition of Evo8_14_F03_CFPe_2 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_F03_CFPe_2 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the growth duration data for Evo8_14_F03_CFPe_2 and the ancestral YFP respectively. This plate did not have an ancestral competition so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

Evo8_14_F03_CFPe_2 (4.5% ethanol): maximum growth rate

Without ethanol, the maximum rate of growth of the evolved CFP strain is close to that of the CFP strain in the ancestral competition but declines with ethanol stress. However the ancestral YFP in evolved competition has a much lower rate than in the ancestral competition. For both strains in evolved competition the linear fit appears relatively poor, there may be a minimum between 4 and 5% (see Figure 5.29).

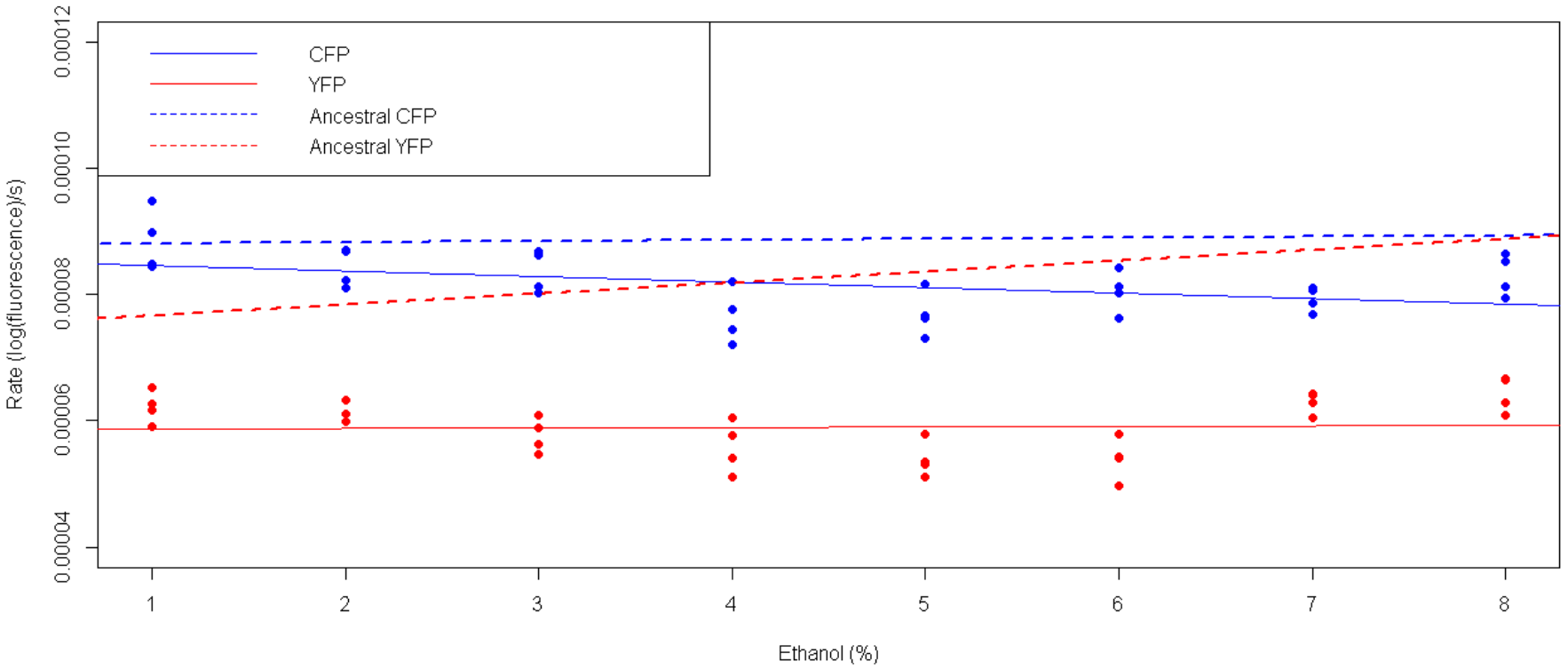


Figure 5.29 A plot of the maximum rate of growth during the competition of Evo8_14_F03_CFPe_2 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_F03_CFPe_2 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the growth data for Evo8_14_F03_CFPe_2 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

Evo8_14_F03_CFPe_2 (4.5% ethanol): fluorescence at 36 hours

The evolved strain shows a very similar fluorescence at 36 hours to its CFP ancestor but its YFP competitor shows depressed growth. The YFP strain in the ancestral competition has a significant positive slope (0.0586+/-0.0166se., p=0.0124*) which increases its fluorescence to parity with the ancestral CFP at higher ethanol concentrations. However, in the evolved competition the YFP strain has a very small increase (slope=0.01388+/-0.00652se.,p=0.0372*) and YFP fluorescence remains lower (see Figure 5.30).

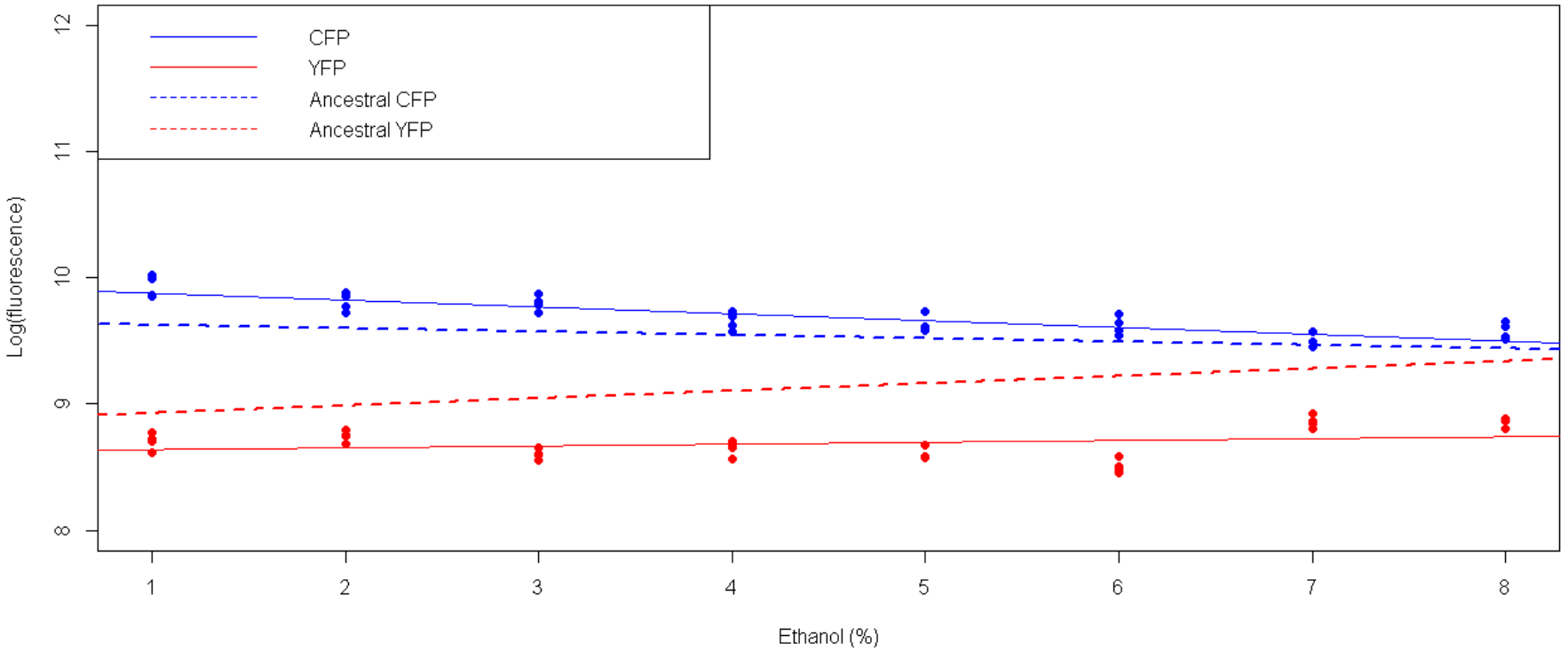


Figure 5.30 A plot of the fluorescence at 36 hours during the competition of Evo8_14_F03_CFPe_2 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_F03_CFPe_2 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the 36-hour fluorescence for Evo8_14_F03_CFPe_2 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

Evo8_14_F03_CFPe_2 (4.5% ethanol): summary

Like Evo8_14_E07_CFPe_4, the elapsed time of active growth for the evolved strain Evo8_14_F03_CFPe_2 is the same as its ancestor, and its maximum rate of growth is similar but the rate of growth of its YFP competitor is reduced.

Evo8_14_C06_CFPe_2 (4.5% ethanol): Introduction

Evo8_14_C06_CFPe_2 was isolated from a population subjected to 4.5% ethanol stress. The fitness analysis was run in the BMG Omega plate reader. The plate did not include a competition between the two ancestors.

Evo8_14_C06_CFPe_2 (4.5% ethanol): growth lag

The onset of detectable growth in the evolved competition is very similar for both strains and closely follows the onset of the CFP strain in the ancestral competition (see Figure 5.31)

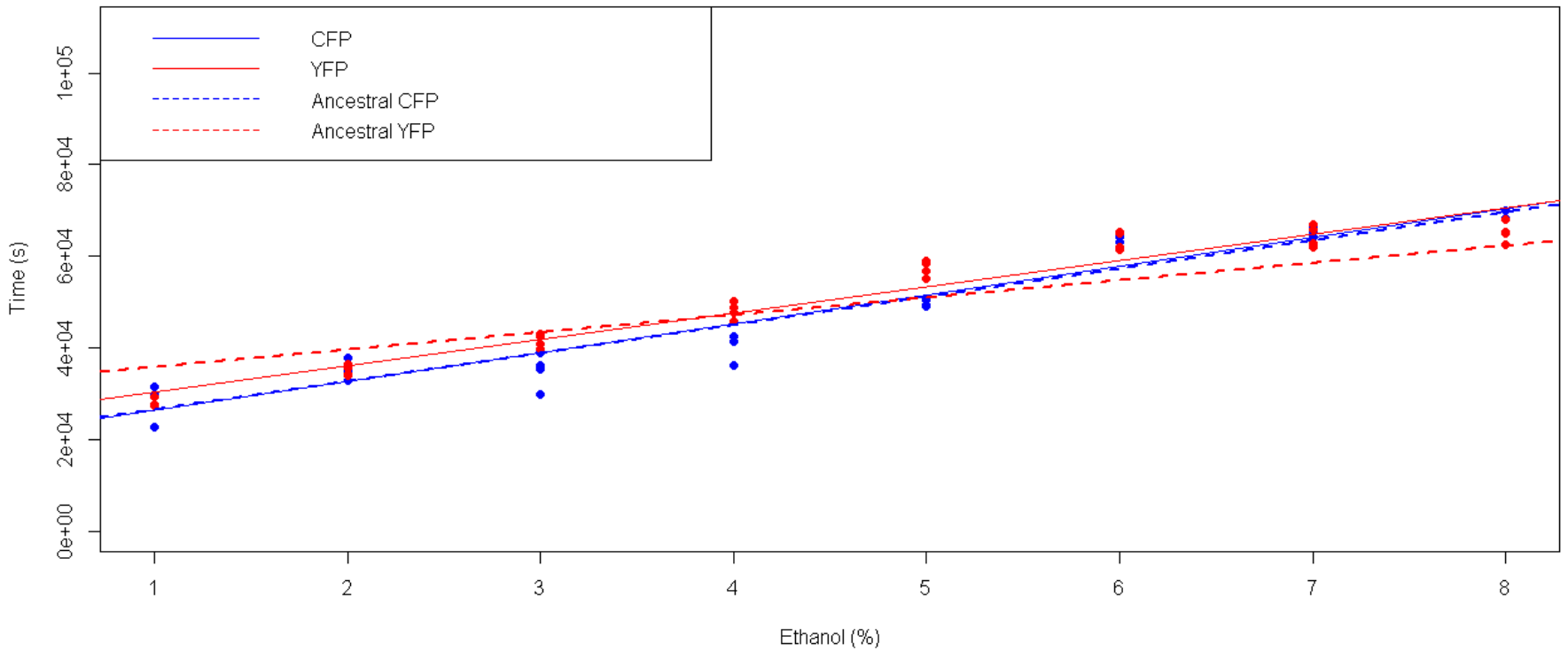


Figure 5.31 A plot of the lag phase before the onset of detectable growth during the competition of Evo8_14_C06_CFPe_2 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_C06_CFPe_2 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the lag data for Evo8_14_C06_CFPe_2 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

Evo8_14_C06_CFPe_2 (4.5% ethanol): onset of diauxy

The two strains in the evolved competition have very similar times to diauxy across the range of ethanol stress. Without an ethanol stress, they both enter diauxy ~2.5 hours earlier than the strains in the ancestral competition but the onset of diauxy converges with the ancestral pattern under ethanol stress up to 8% (see Figure 5.32).

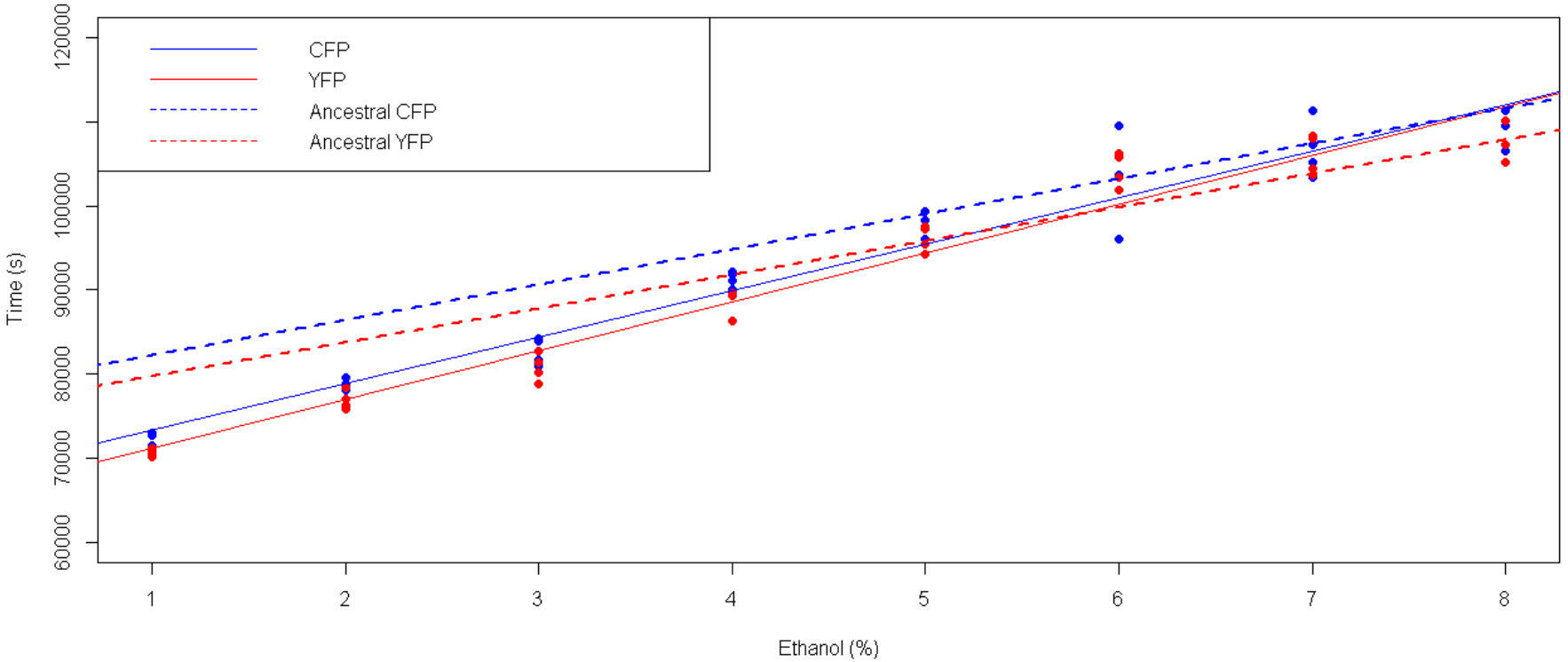


Figure 5.32 A plot of the time to onset of diauxy during the competition of Evo8_14_C06_CFPe_2 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_C06_CFPe_2 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the diauxy data for Evo8_14_C06_CFPe_2 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

Evo8_14_C06_CFPe_2 (4.5% ethanol): growth duration

The effect of growth onset and diauxy is that the time between detectable onset and diauxy is very similar for both strains in the evolved competition (see Figure 5.33).

Evo8_14_C06_CFPe_2 (4.5% ethanol): maximum growth rate

The maximum rate of growth of the evolved strain is very similar to that of the CFP strain in the ancestral competition but the YFP strain is substantially depressed (see Figure 5.34). The linear fit for the YFP strain in the evolved competition is poor, it appears that the maximum rate falls to a minimum between 3%-5% ethanol and improves at higher ethanol stress

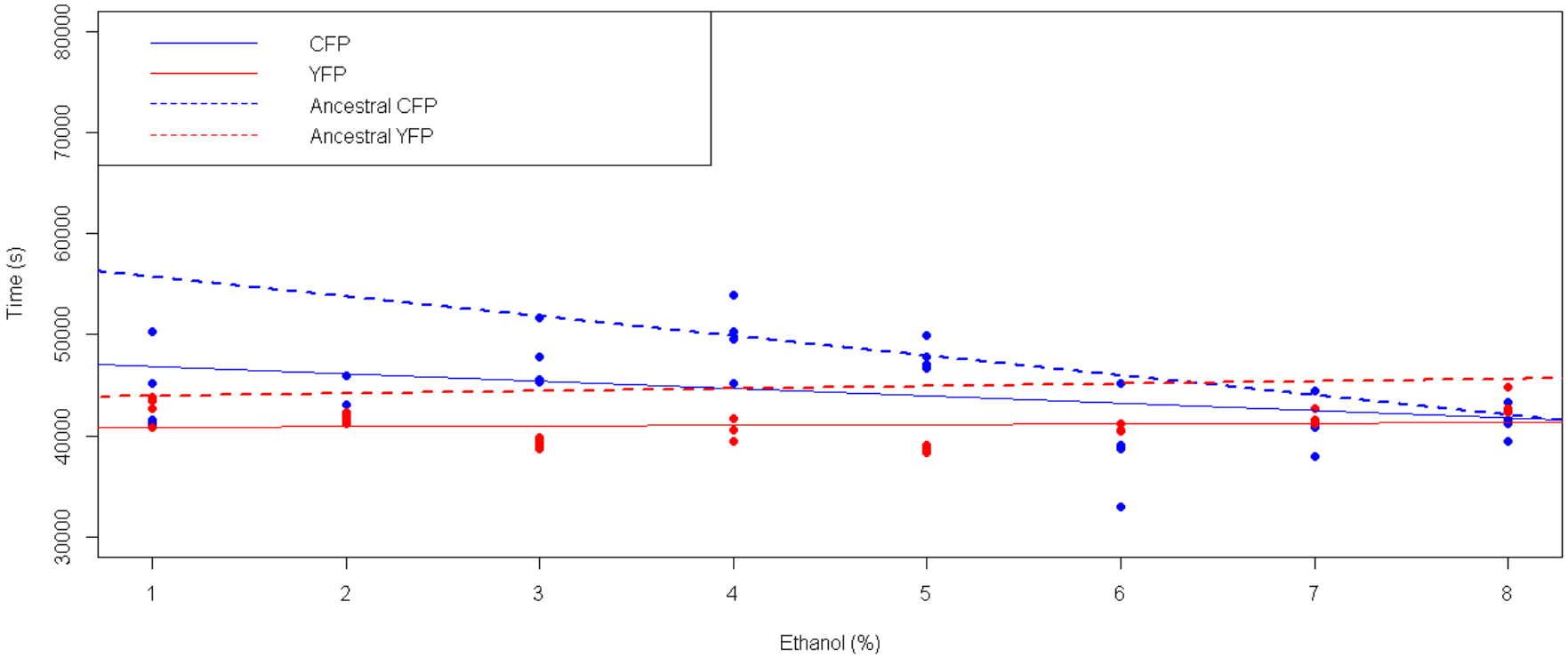


Figure 5.33 A plot of time elapsed between the start of detectable growth the onset of diauxy during the competition of Evo8_14_C06_CFPe_2 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_C06_CFPe_2 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the growth duration data for Evo8_14_C06_CFPe_2 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

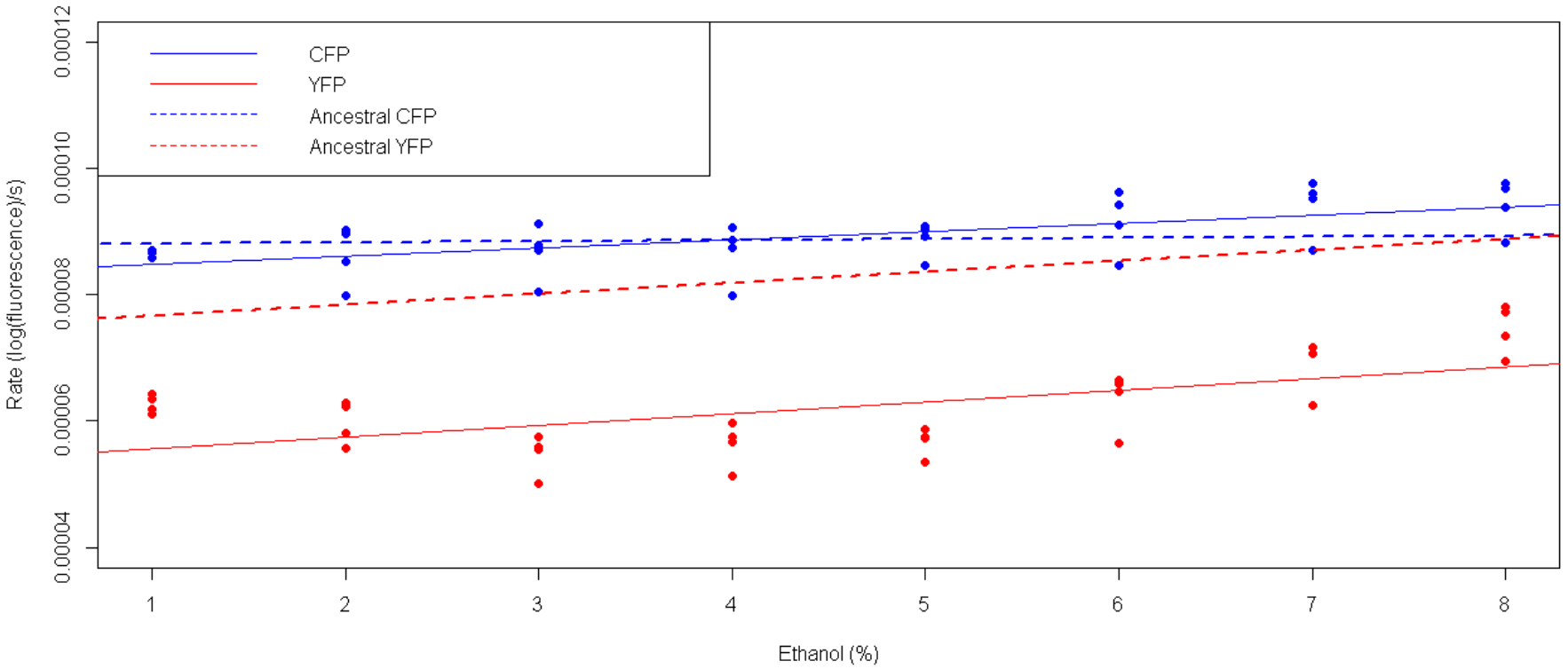


Figure 5.34 A plot of the maximum rate of growth during the competition of Evo8_14_C06_CFPe_2 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_C06_CFPe_2 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the growth data for Evo8_14_C06_CFPe_2 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

Evo8_14_C06_CFPe_2 (4.5% ethanol): fluorescence at 36 hours

The fluorescence at 36 hours is similar to the pattern of maximum rate, the evolved strain CFP fluorescence is very similar to the CFP strain in the ancestral competition but slightly better overall and with a smaller negative slope (-0.011806+/-0.003798se., p=0.00284**; ancestral slope=-0.02657+/-0.01068se., p=0.0473*). The difference in slope gives the evolved strain better fitness at higher ethanol stresses up to 8%. The YFP fluorescence in the evolved strain is poorer than the ancestral competition at all ethanol concentrations. However, the YFP linear regression fit is poor; the data looks as though there is a minimum between 3-5% ethanol with growth at 0% and 8% ethanol stress approaching the YFP strain in the ancestral competition (see Figure 5.35).

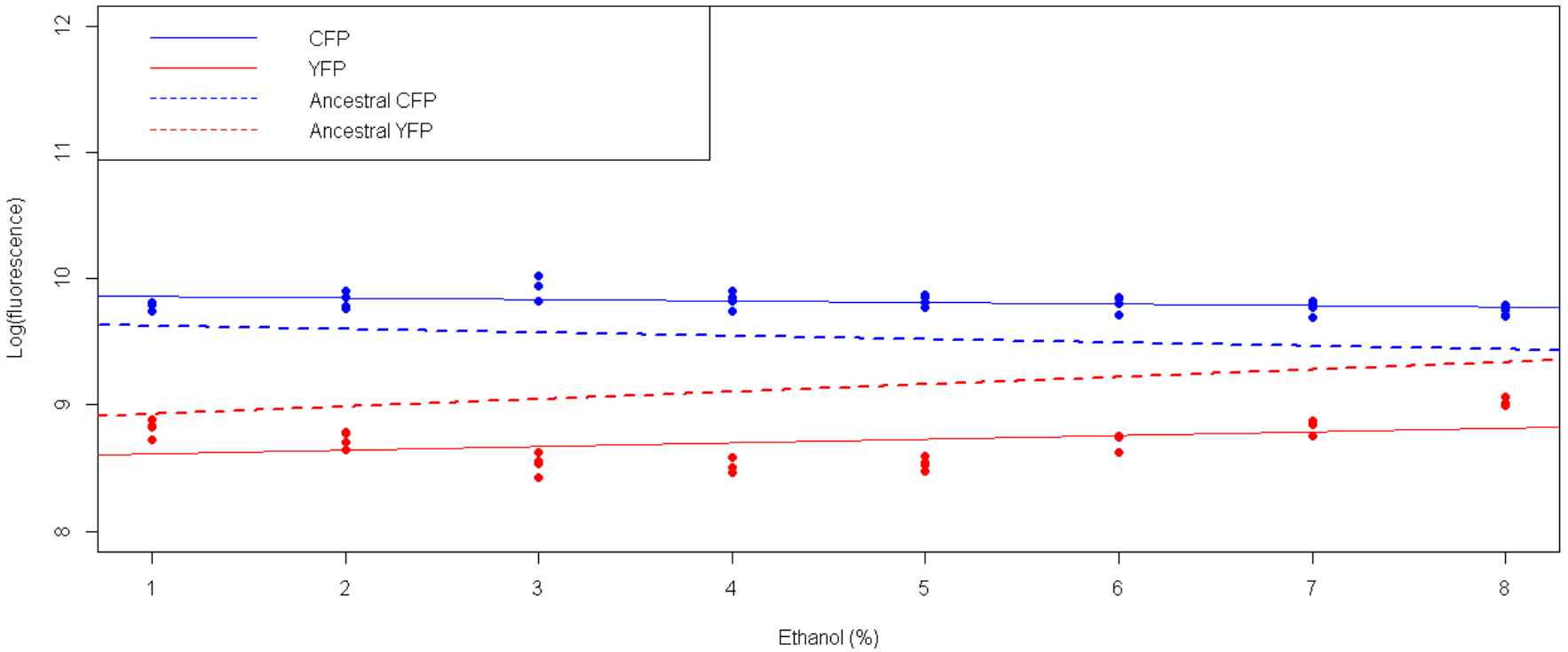


Figure 5.35 A plot of the fluorescence at 36 hours during the competition of Evo8_14_C06_CFPe_2 and ancestral YFP strains in media with different ethanol concentrations.
Competitions were run on the BMG Omega plate reader. The points in blue are for the Evo8_14_C06_CFPe_2 strain and the points in red are for the YFP strain. The blue and red solid lines represent the linear best fit for the 36-hour fluorescence for Evo8_14_C06_CFPe_2 and the ancestral YFP respectively. This plate did not have an ancestral competition on it so the dotted lines represent the linear best fit for the ancestral competitions from other plates read in the same machine.

Evolution 8 Summary

In summary, the growth patterns in ancestral competition of Evo8_14_C06_CFPe_2 are very similar to both Evo8_14_F03_CFPe_2 and Evo8_14_E07_CFPe_4; the elapsed time of active growth for the evolved strain is the same as its competitor, and its maximum rate of growth is very similar to its ancestor but the rate of growth of its YFP competitor is substantially depressed

5.3.4. Evolved strain analysis: evolution 10

All of the evolution 10 derived strains were subjected to an ethanol stress that increased from 0% to 9% with a 0.5% increase each plate transfer. All analysed strains were isolated from plate 20. The plates analysing the competitive fitness of evolution 10 strains in different ethanol concentrations were run on the BMG FluoStar plate reader.

Evo10_20_G10_YFPe_2: growth lag

There is a small reduction in the time to detectable growth in the evolved YFP strain, a reduction which increases with ethanol stress (see Figure 5.36).

Evo10_20_G10_YFPe_2: onset of diauxy

The onset of diauxy mirrors the early onset of growth (see Figure 5.37).

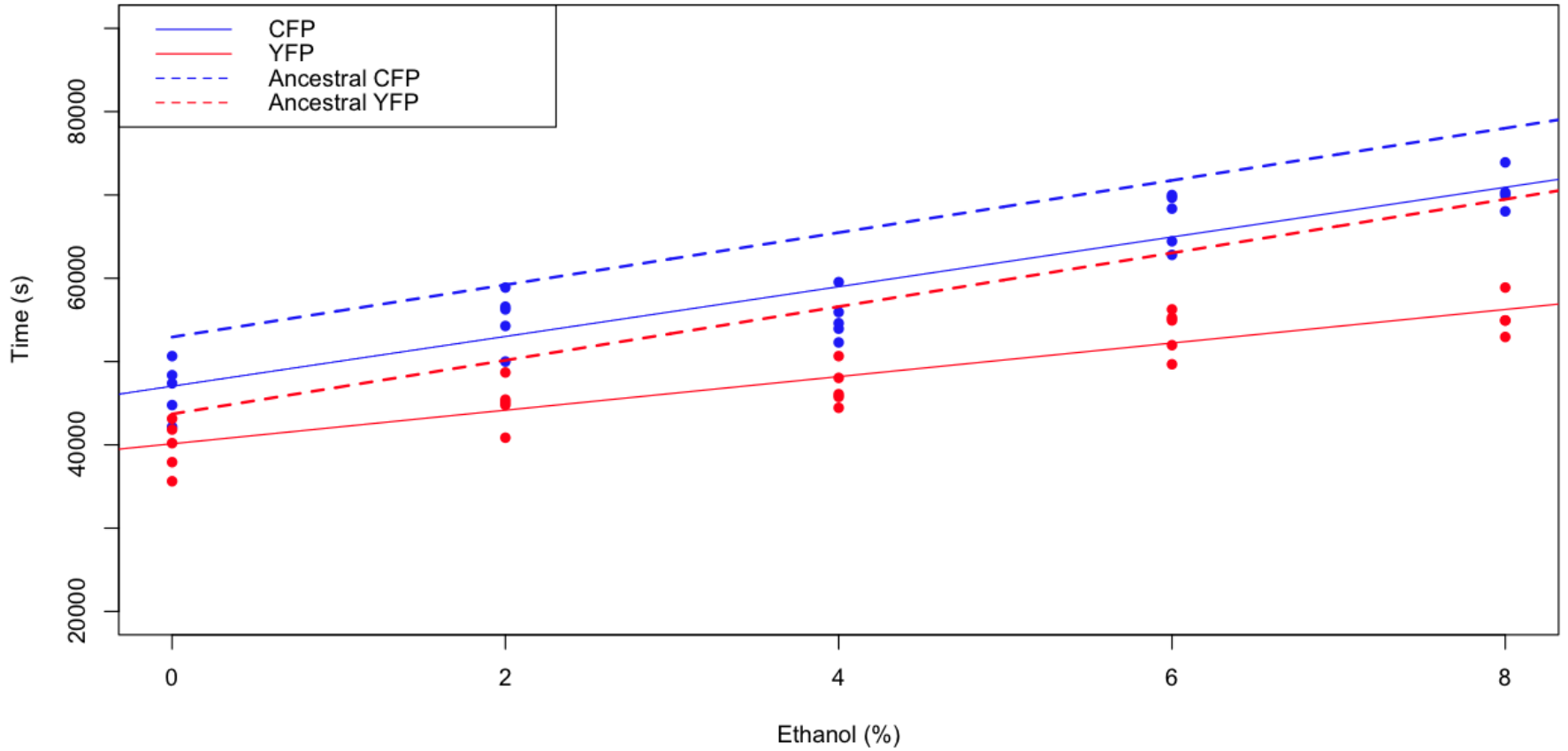


Figure 5.36 A plot of the lag phase before the onset of detectable growth during the competition of the ancestral CFP and Evo10_20_G10_YFPe_2 strains in media with different ethanol concentrations.
The points in blue are for the ancestral CFP and the points in red are for the Evo10_20_G10_YFPe_2 strain. The blue and red solid lines represent the linear best fit for the lag data for the ancestral CFP and Evo10_20_G10_YFPe_2 respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

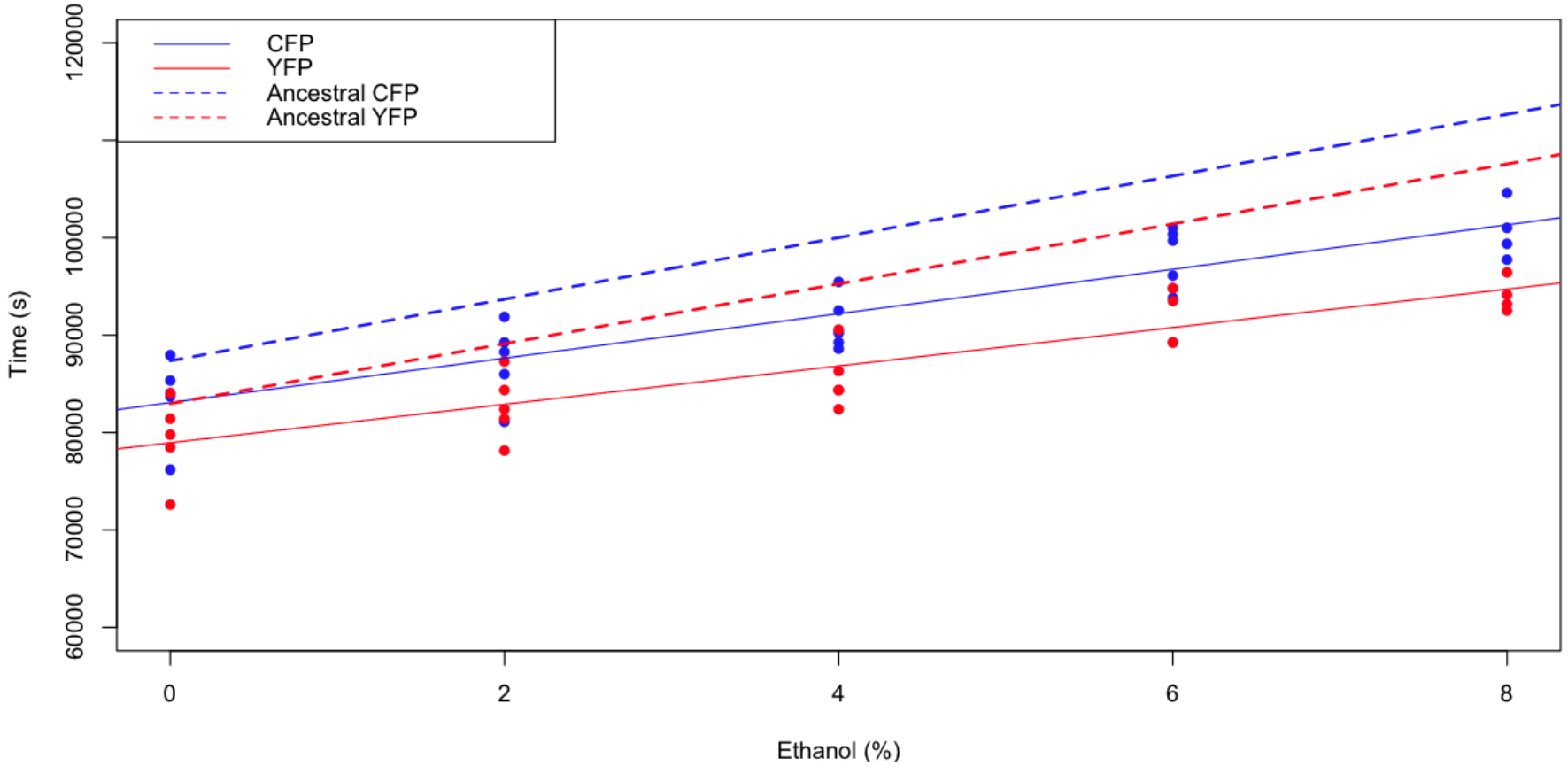


Figure 5.37 A plot of the time of onset of diauxy during the competition of the ancestral CFP and Evo10_20_G10_YFPe_2 strains in media with different ethanol concentrations.
The points in blue are for the ancestral CFP and the points in red are for the Evo10_20_G10_YFPe_2 strain. The blue and red solid lines represent the linear best fit for the onset of diauxy for the ancestral CFP and Evo10_20_G10_YFPe_2 respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo10_20_G10_YFPe_2: growth duration

The result of the shifted lag and diauxy is that the time spent by the evolved strain in active growth (See Figure 5.38) is the same as the YFP strain in the ancestral competition and is independent of ethanol stress within the range 0-8% (slope=-42.73+/-89.78se., p=0.639) but the CFP strain competing with Evo10_20_G10_YFPe_2 has a declining time spent in active growth as ethanol stress increases with a significant declining slope (-706.7+/-152se.,p=0.000121***).

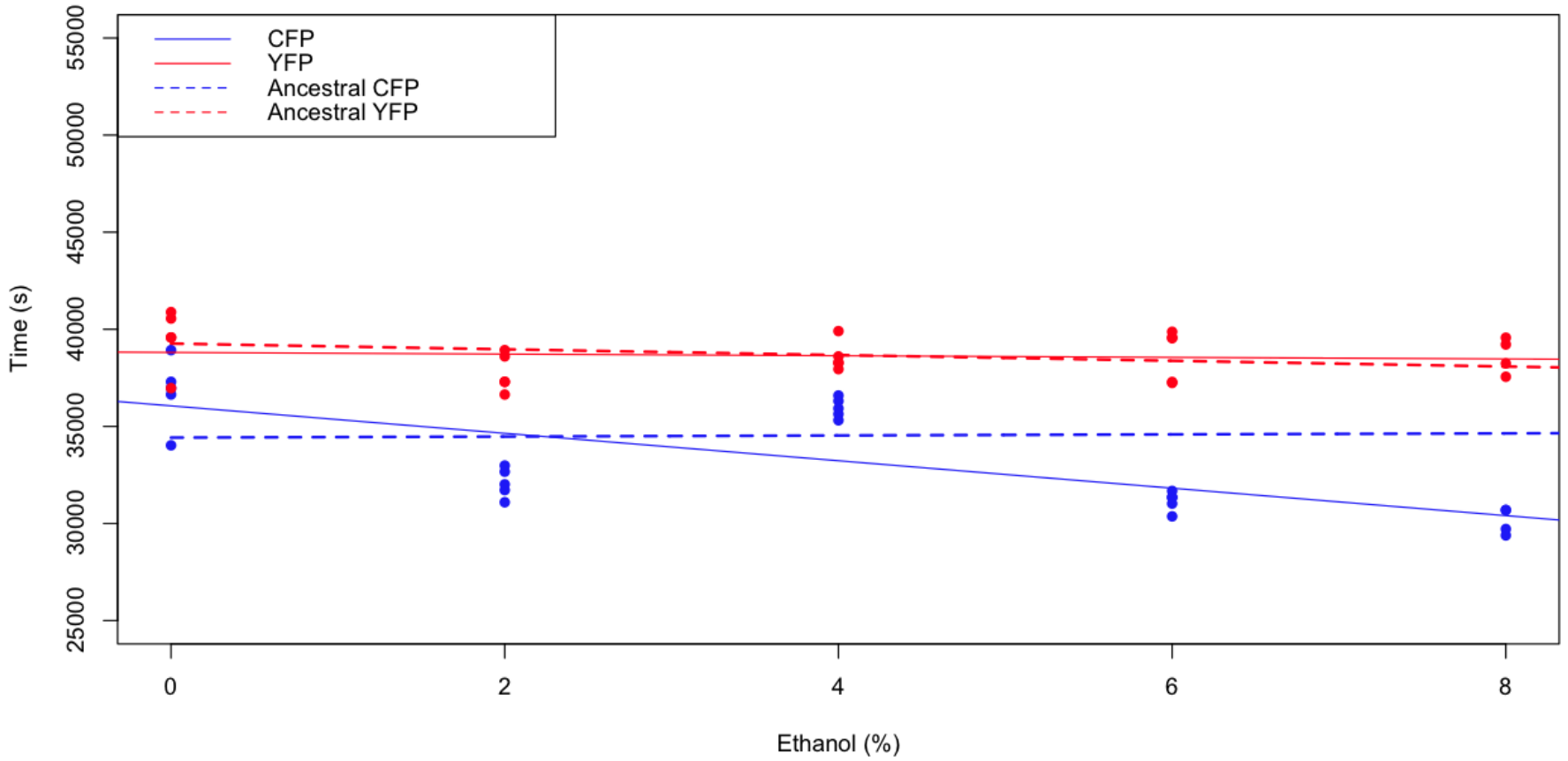


Figure 5.38 A plot of the elapsed time between growth being detected and the onset of diauxy during the competition of the ancestral CFP and Evo10_20_G10_YFPe_2 strains in media with different ethanol concentrations.
The points in blue are for the ancestral CFP and the points in red are for the Evo10_20_G10_YFPe_2 strain. The blue and red solid lines represent the linear best fit for the growth duration data for the ancestral CFP and Evo10_20_G10_YFPe_2 respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo10_20_G10_YFPe_2: maximum growth rate

Without an ethanol stress, the maximum rate of growth of Evo10_20_G10_YFPe_2 (intercept=$4.157x10^{-05}$+/-$3.539x10^{-06}$) is somewhat less than that of its ancestral CFP competitor (intercept= $5.263x10^{-05}$+/-$1.398x10^{-06}$se.). However, with increasing ethanol stress the maximum rate of the evolved YFP strain increases (slope= $3.740x10^{-06}$+/-$7.489x10^{-07}$se., p=$5.35x10^{-05}$***) while its ancestral CFP competitor declines (slope= -$7.255x10^{-07}$+/-$2.958x10^{-07}$se., p=0.0226*); see Figure 5.39. Above 4% ethanol the rate of growth of the evolved strain is clearly greater than its ancestral competitor.

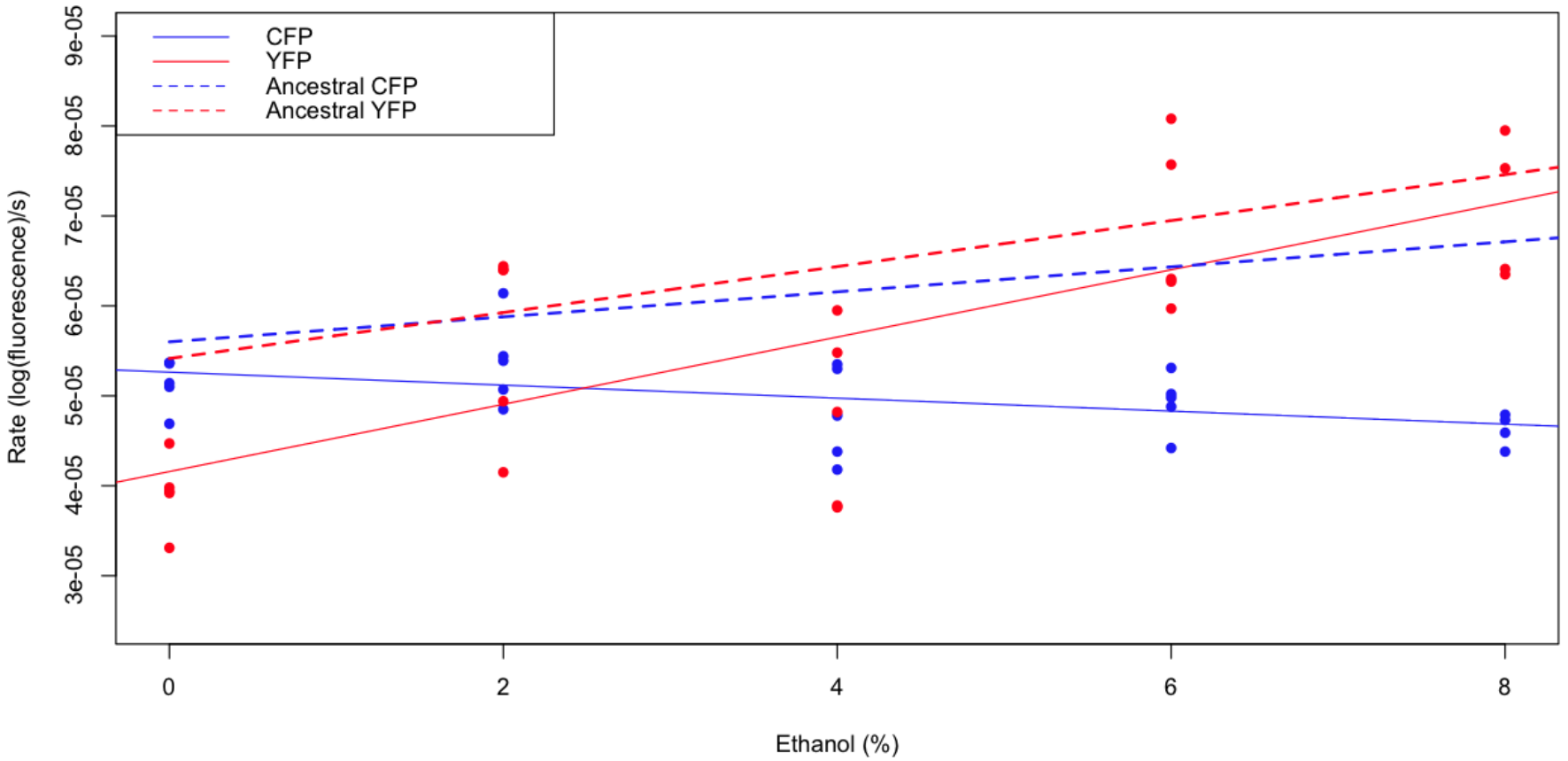


Figure 5.39 A plot of the maximum growth rate during the competition of the ancestral CFP and Evo10_20_G10_YFPe_2 strains in media with different ethanol concentrations.
The points in blue are for the ancestral CFP and the points in red are for the Evo10_20_G10_YFPe_2 strain. The blue and red solid lines represent the linear best fit for the growth rate data for the ancestral CFP and Evo10_20_G10_YFPe_2 respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo10_20_G10_YFPe_2: fluorescence at 36 hours

The fluorescence at 36 hours follows the pattern of growth rate. Without an added ethanol stress the CFP ancestor outcompetes Evo10_20_G10_YFPe_2 but above 4% the evolved strain is fitter (see Figure 5.40). However, it is only above ~6% ethanol that the total YFP fluorescence produced is greater than that in the ancestral competition.

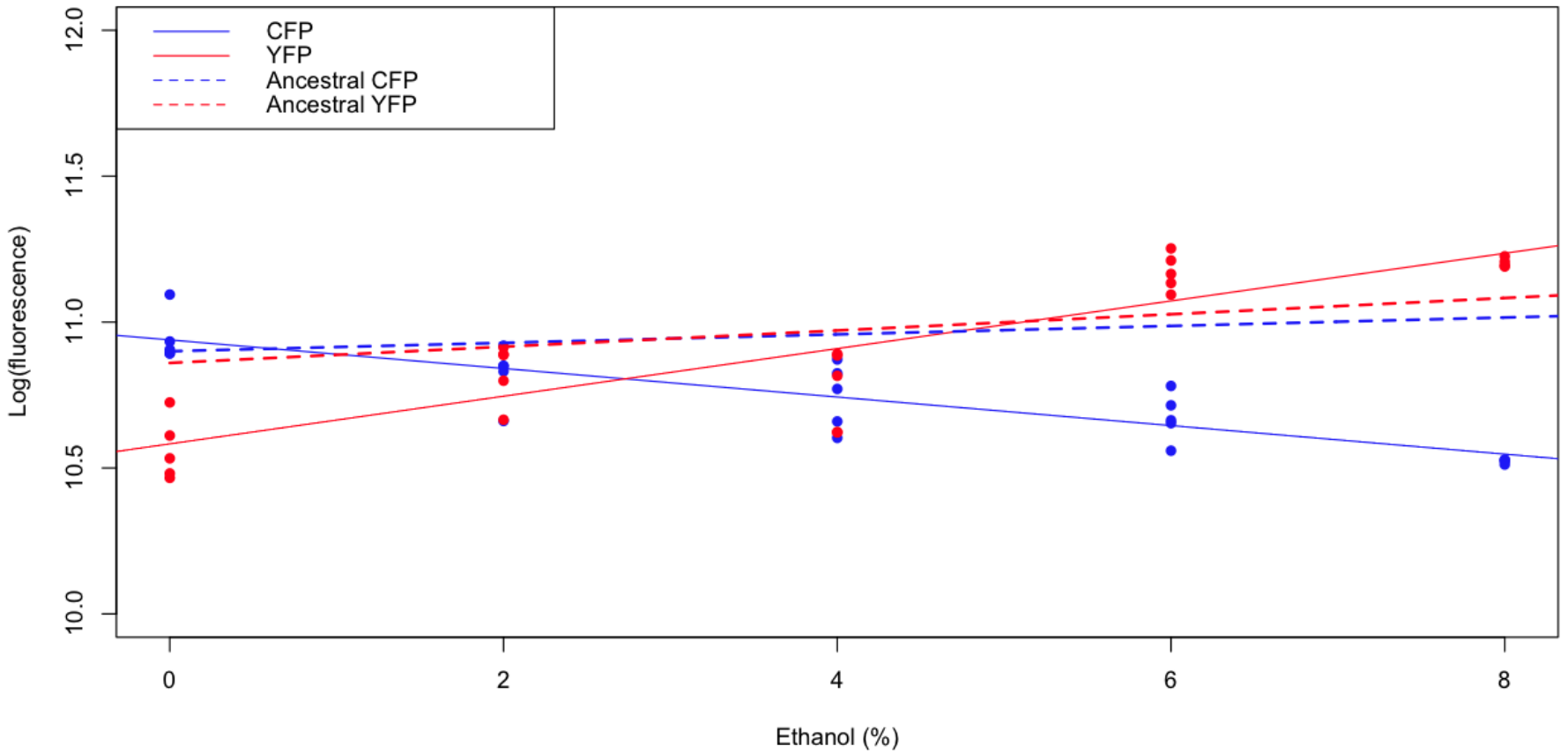


Figure 5.40 A plot of the fluorescence at 36 hours during the competition of the ancestral CFP and Evo10_20_G10_YFPe_2 strains in media with different ethanol concentrations.
The points in blue are for the ancestral CFP and the points in red are for the Evo10_20_G10_YFPe_2 strain. The blue and red solid lines represent the linear best fit for the 36-hour fluorescence for the ancestral CFP and Evo10_20_G10_YFPe_2 respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo10_20_G10_YFPe_2: summary

The maximum growth rate and fluorescence Evo10_20_G10_YFPe_2 at 36 hours are both poorer than its ancestral competitor in YPD but consistently improve with added ethanol.

Evo10_20_F10_CFPe_2: Introduction

The 6% ethanol medium used to make up part of the plate to analyse competitive fitness of the strain Evo10_20_F10_CFPe_2 was contaminated. The data from the other competitions was unaffected with no contaminated blanks in other growth media.

Evo10_20_F10_CFPe_2: growth lag

Without an added ethanol stress the onset of detectable growth by Evo10_20_F10_CFPe_2 takes slightly longer than the ancestral CFP strain in competition, but growth onset converges on the ancestral pattern as the ethanol concentration increases (see figure Figure 5.41).

Evo10_20_F10_CFPe_2: onset of diauxy

This is mirrored by the pattern of onset of diauxy (see Figure 5.42).

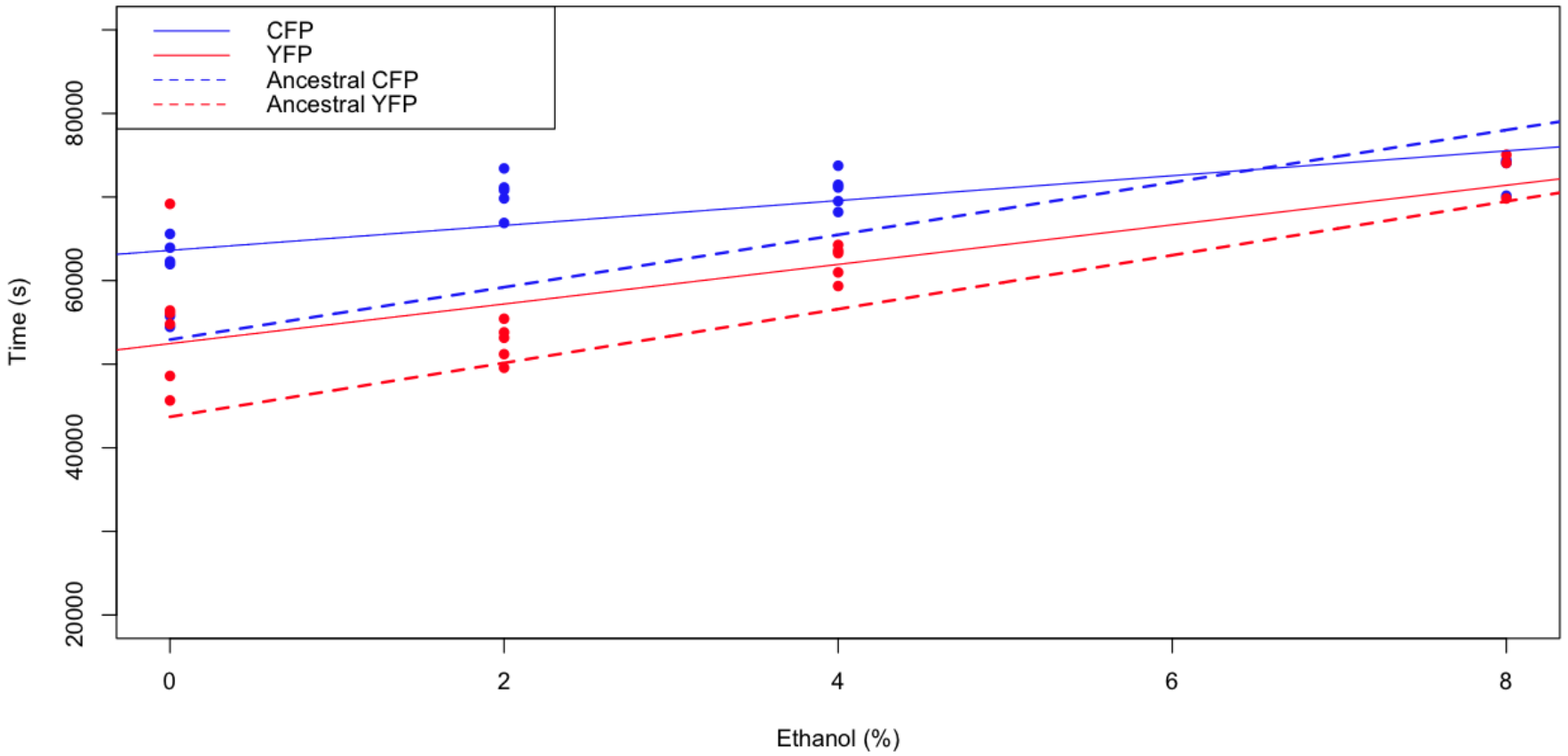


Figure 5.41 A plot of the lag phase before the onset of detectable growth during the competition of the evolved strain Evo10_20_F10_CFPe2 and ancestral YFP strains in media with different ethanol concentrations.
The points in blue are for Evo10_20_F10_CFPe2 and the points in red are for the ancestral YFP strain. The blue and red solid lines represent the linear best fit for the lag data for the Evo10_20_F10_CFPe2 and ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

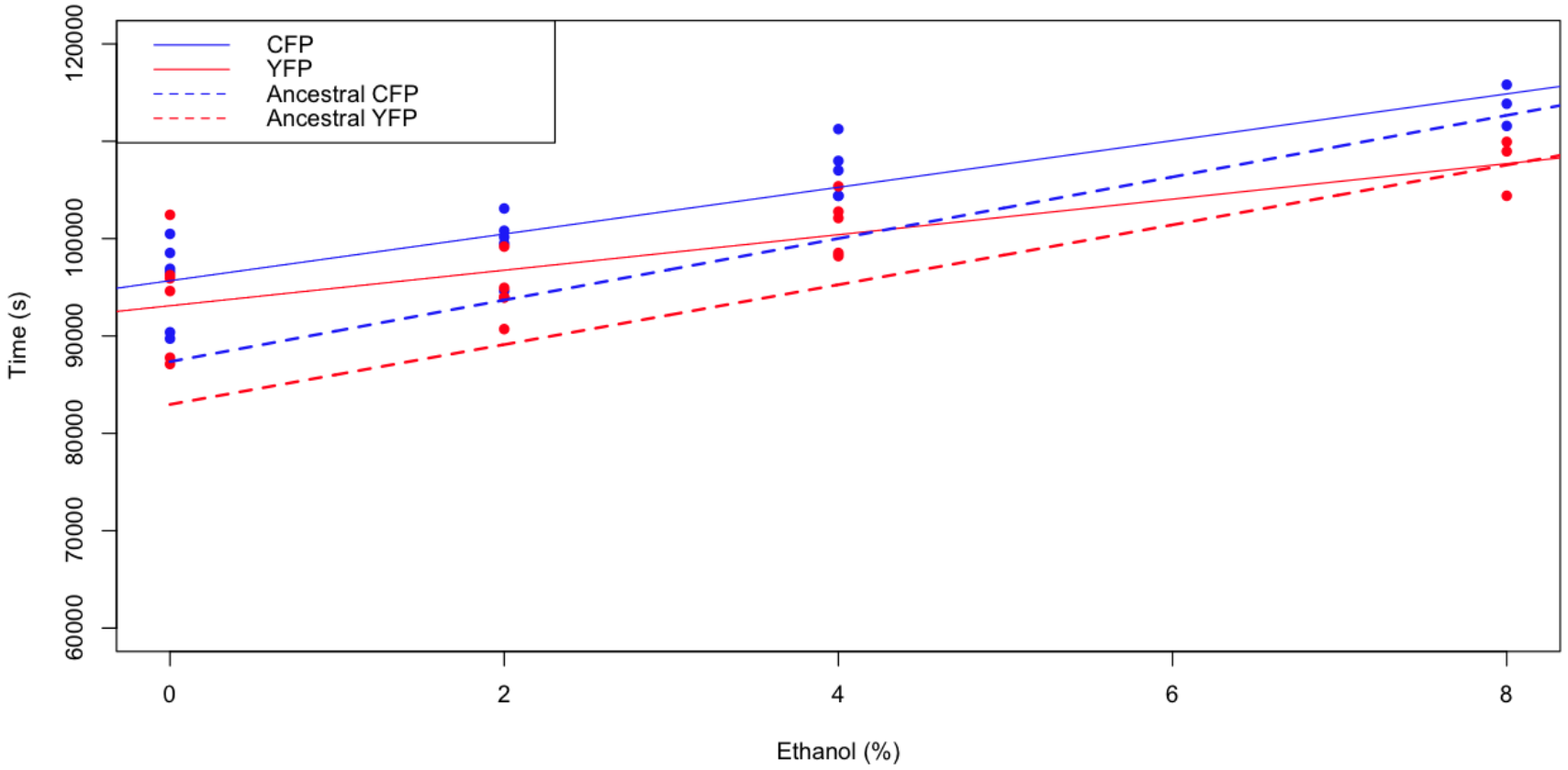


Figure 5.42 A plot of the time of onset of diauxy during the competition of the evolved strain Evo10_20_F10_CFPe2 and ancestral YFP strain in media with different ethanol concentrations.
The points in blue are for the Evo10_20_F10_CFPe2 and the points in red are for the ancestral YFP strain. The blue and red solid lines represent the linear best fit for the onset of diauxy for the Evo10_20_F10_CFPe2 and ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo10_20_F10_CFPe_2: growth duration

The corresponding changes in detectable growth onset and diauxy mean that the overall time spent growing, from detectable growth to onset of diauxy, is very similar to the ancestral competition (see figure Figure 5.43). However, there is a small but significant positive slope with ethanol stress in the evolved strain (910.6+/-284.5se.,p=0.00524**) and a smaller but just significant negative slope in its ancestral YFP competitor (-543.4+/-219.0se.,p=0.0239*)

Evo10_20_F10_CFPe_2: maximum growth rate

There is no significant increase in maximum rate for the evolved strain in competition (slope=1.305e-06+/-8.587e-07se., p=0.147) see Figure 5.44.

.

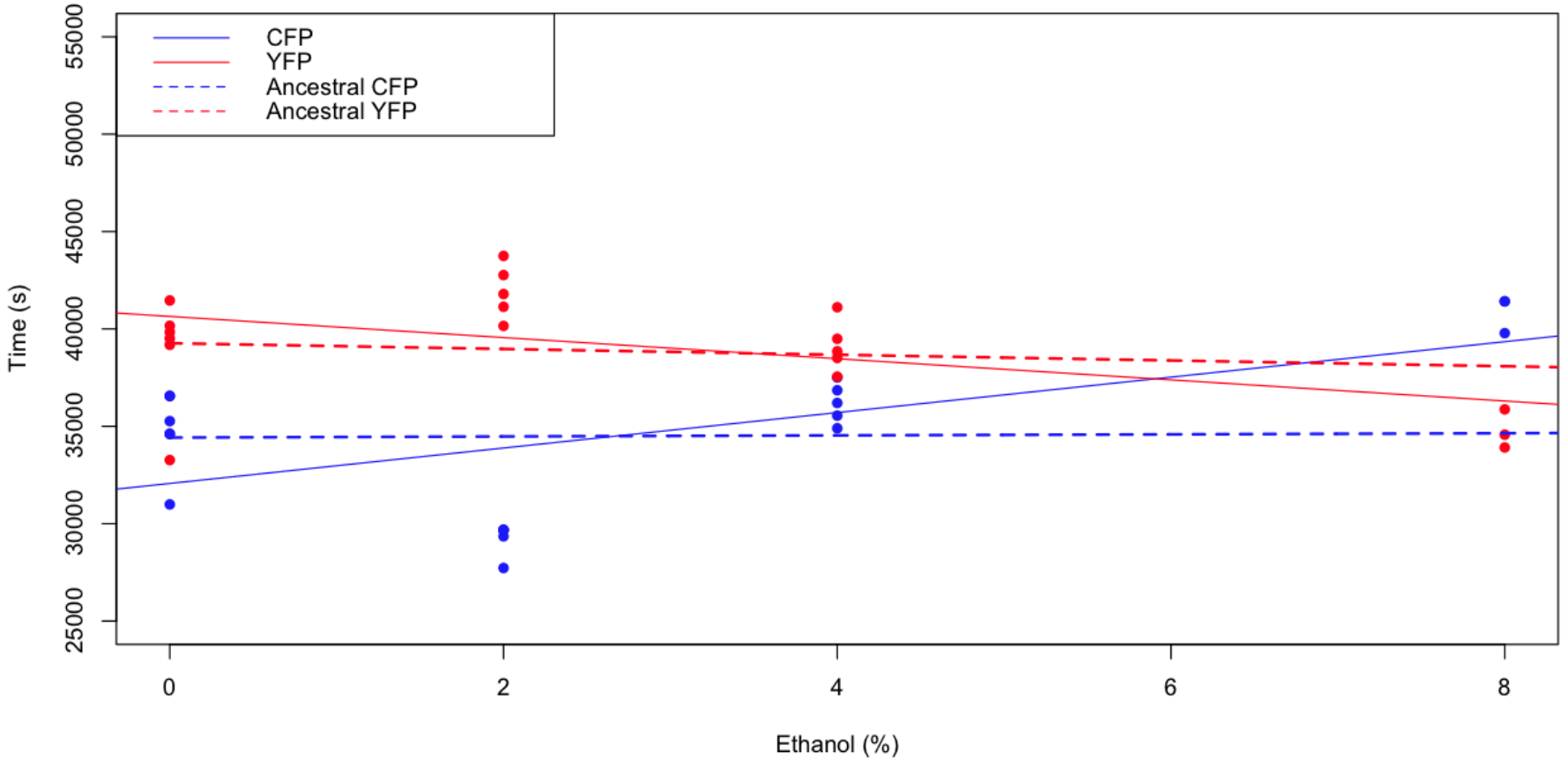


Figure 5.43 A plot of the elapsed time from detectable growth to onset of diauxy during the competition of the evolved strain Evo10_20_F10_CFPe2 and ancestral YFP strain in media with different ethanol concentrations.

The points in blue are for the Evo10_20_F10_CFPe2 and the points in red are for the ancestral YFP strain. The blue and red solid lines represent the linear best fit for the growth duration data for the Evo10_20_F10_CFPe2 and ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

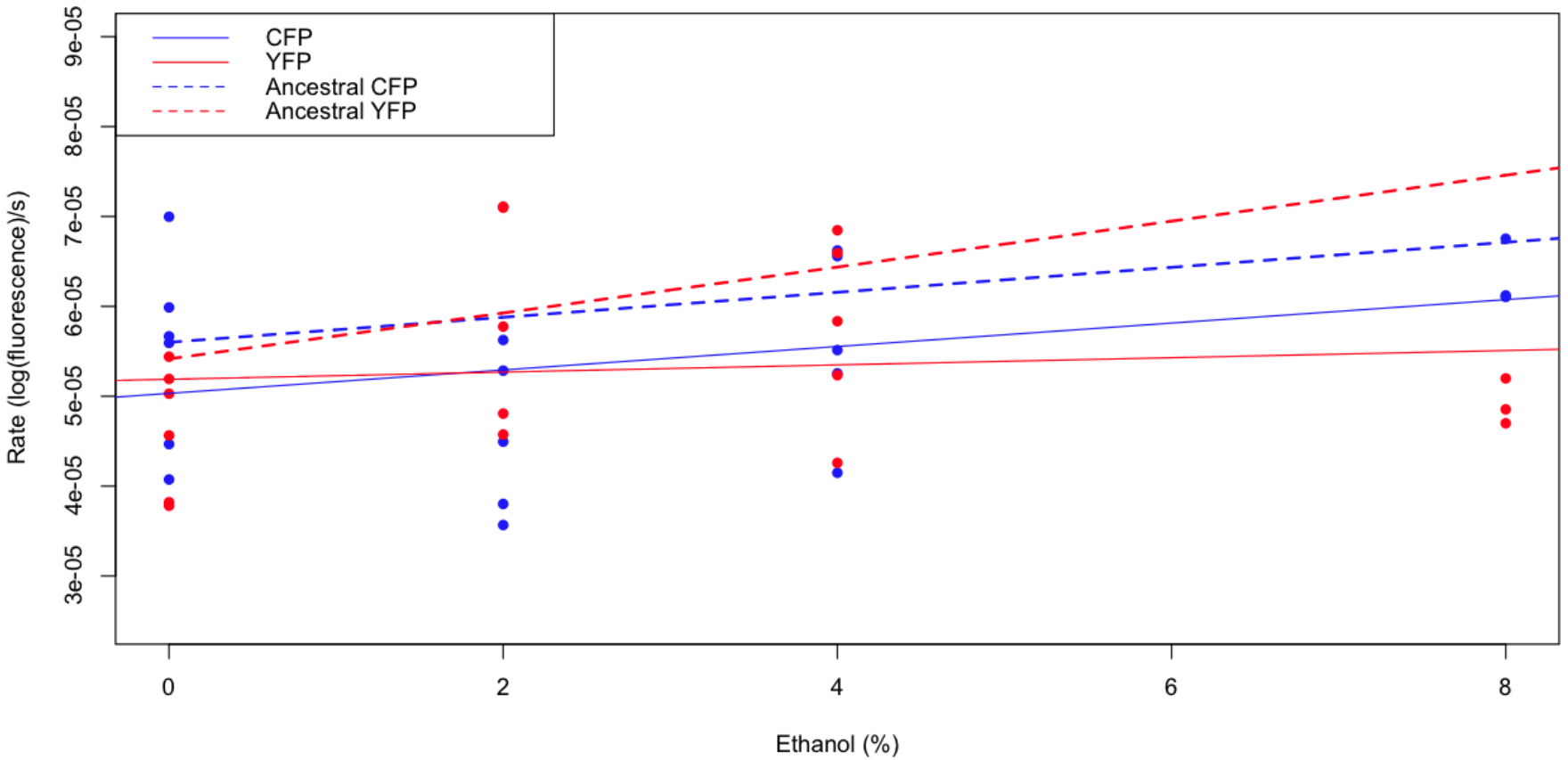


Figure 5.44 A plot of the maximum growth rate during the competition of the evolved strain Evo10_20_F10_CFPe2 and ancestral YFP strain in media with different ethanol concentrations.
The points in blue are for the Evo10_20_F10_CFPe2 and the points in red are for the ancestral YFP strain. The blue and red solid lines represent the linear best fit for the growth rate data for the Evo10_20_F10_CFPe2 and ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo10_20_F10_CFPe_2: fluorescence at 36 hours

Across the range of ethanol concentrations, the fluorescence of Evo10_20_F10_CFPe2 at 36 hours is neither substantially different from its YFP ancestral competitor nor that of the ancestors in competition with each other. Similarly there is a small fall in YFP strain fluorescence at 36 hours as ethanol concentrations increase but it is not significant (see Figure 5.45).

Evo10_20_F10_CFPe_2: summary

Growth by the evolved CFP strain is poorer than the ancestor without an ethanol stress but that a small fitness increment is seen above 4% ethanol.

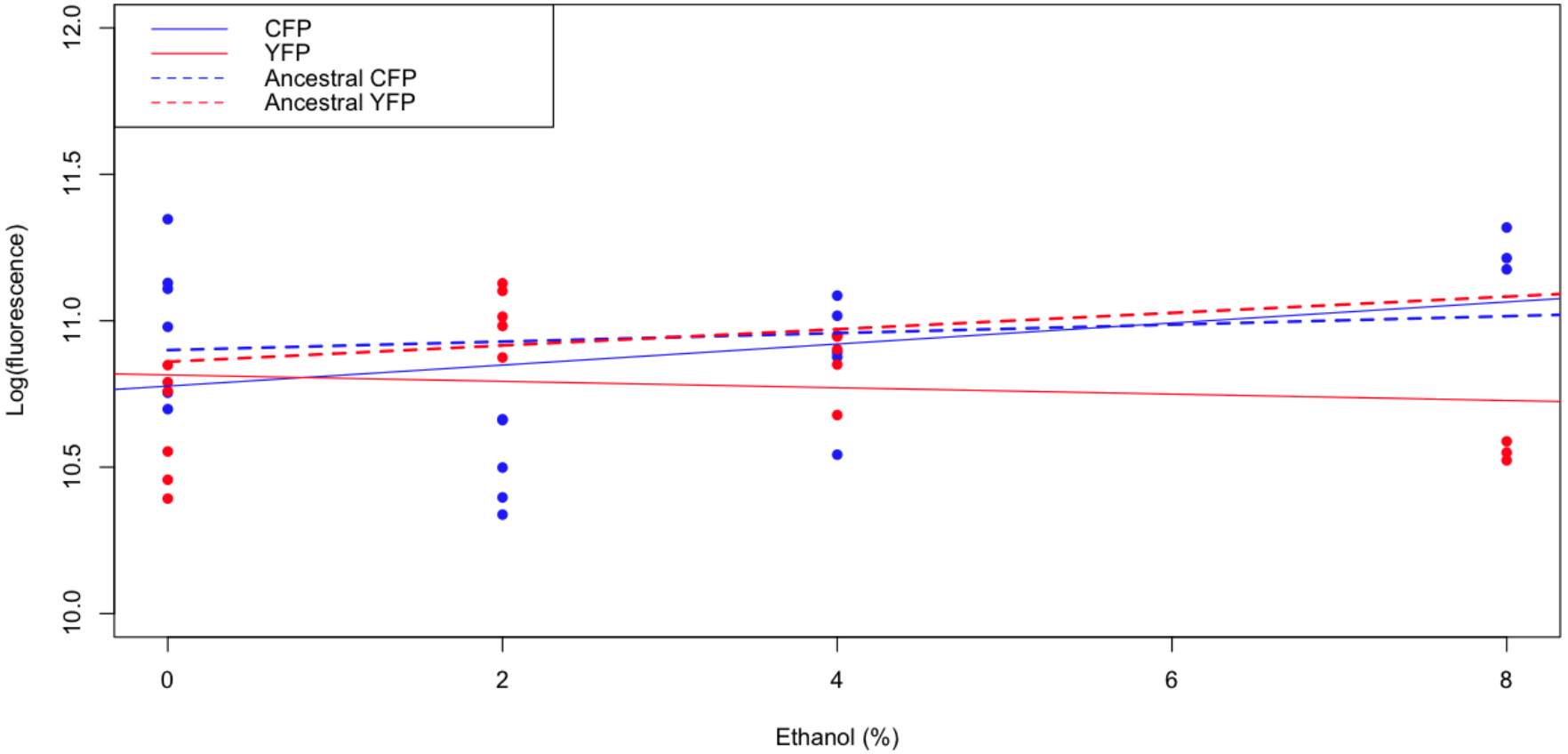


Figure 5.45 A plot of the fluorescence at 36 hours during the competition of the evolved strain Evo10_20_F10_CFPe2 and ancestral YFP strain in media with different ethanol concentrations.
The points in blue are for the Evo10_20_F10_CFPe2 and the points in red are for the ancestral YFP strain. The blue and red solid lines represent the linear best fit for the 36-hour fluorescence for the Evo10_20_F10_CFPe2 and ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo10_20_B03_CFPe_5: growth lag

This strain was assessed using the BMG Fluostar. FACS analysis is consistent with this strain being a euploid triploid (see Chapter 6, section 6.3.8), however it was not sequenced and may be aneuploid. The regression line for the time to detectable emergence is not substantially different across the range of ethanol stresses (see Figure 5.46) for Evo10_20_B03_CFPe_5 and its ancestral YFP tagged competitor. However the time to emergence for the evolved competition is more than four hours earlier (the intercept of the ancestral CFP regression is 52942s against that of Evo10_20_B03_CFPe_5 at 37305s)

Evo10_20_B03_CFPe_5: onset of diauxy

The time to diauxy is correspondingly reduced (see Figure 5.47). The onset of diauxy for Evo10_20_B03_CFPe_5 and the ancestral YFP are close together suggesting that both are triggered by glucose depletion rather than ethanol toxicity.

.

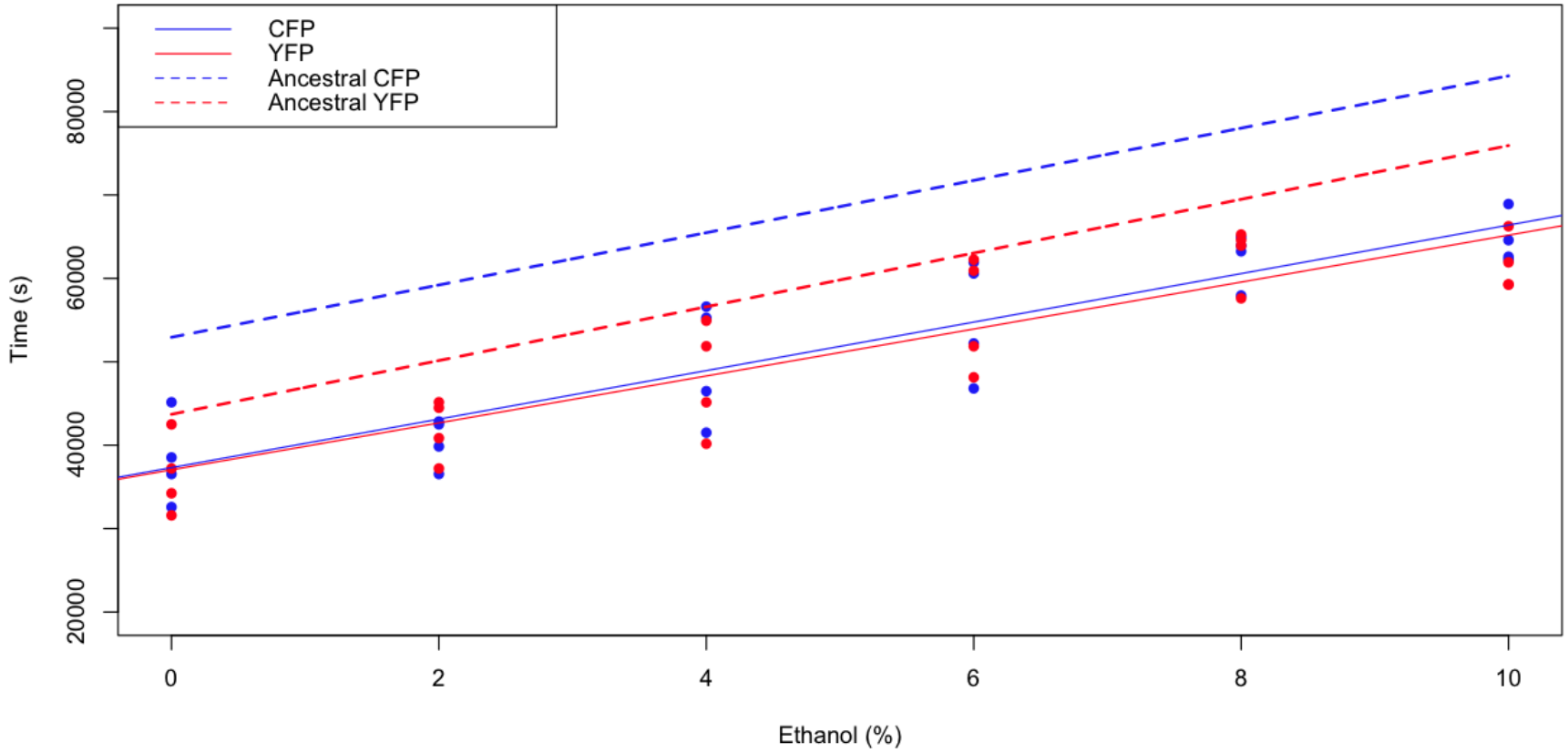


Figure 5.46 A plot of the lag phase before the onset of detectable growth during the competition of the evolved strain Evo_10_20_B03_CFPe_5 and ancestral YFP strains in media with different ethanol concentrations.
The points in blue are for Evo_10_20_B03_CFPe_5 and the points in red are for the ancestral YFP strain. The blue and red solid lines represent the linear best fit for the lag data for the Evo_10_20_B03_CFPe_5 and ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

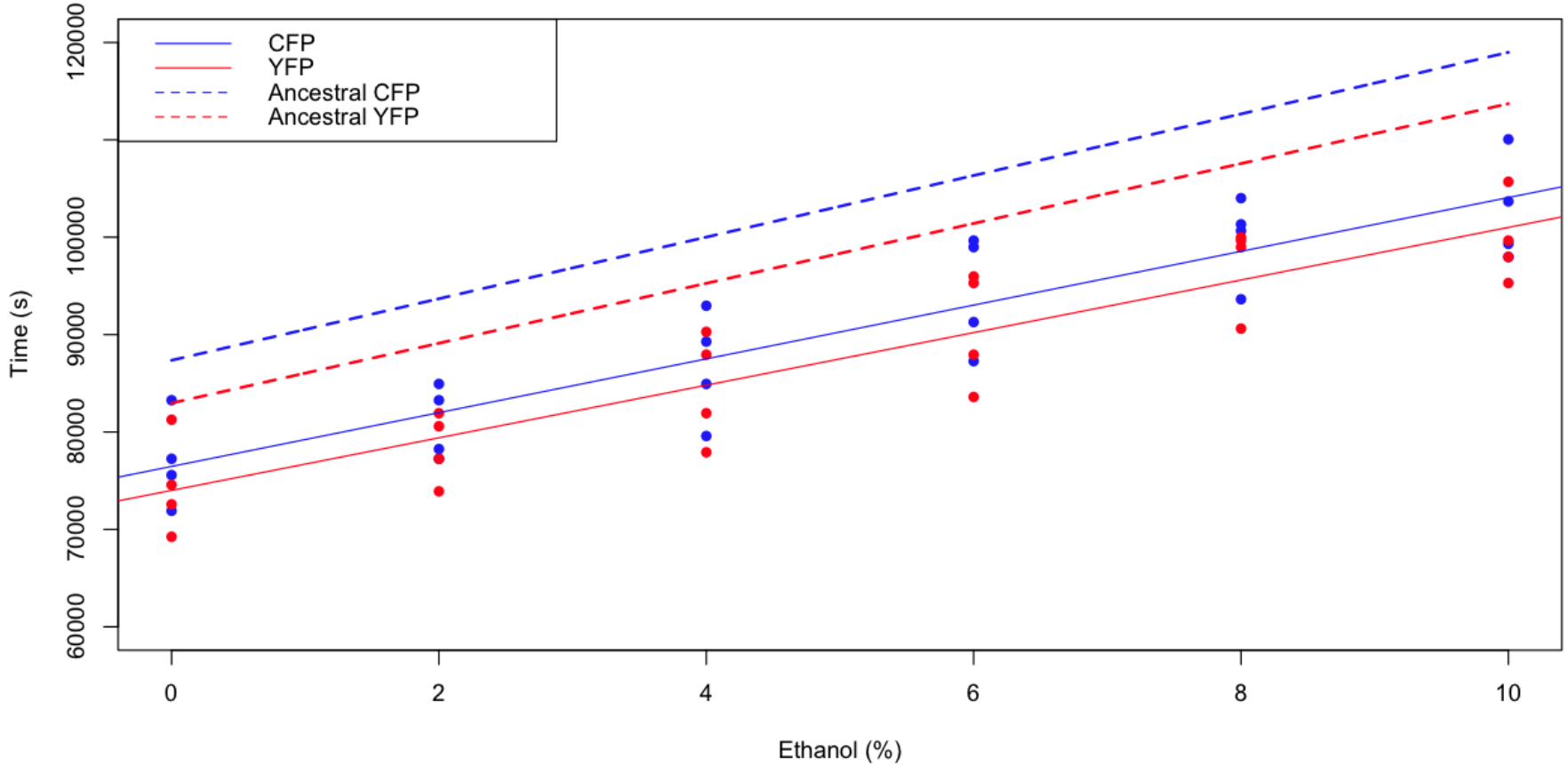


Figure 5.47 A plot of the time of onset of diauxy during the competition of the evolved strain Evo_10_20_B03_CFPe_5 and ancestral YFP strain in media with different ethanol concentrations.
The points in blue are for the Evo_10_20_B03_CFPe_5 and the points in red are for the ancestral YFP strain. The blue and red solid lines represent the linear best fit for the onset of diauxy for the Evo_10_20_B03_CFPe_5 and ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo10_20_B03_CFPe_5: growth duration

The result (see Figure 5.48) is that the time spent in active growth by the evolved strain Evo10_20_B03_CFPe_5 slightly better (~40 minutes) than its YFP competitor and not distinguishable from that of the YFP strain in the ancestral competition. The intercept is 39149s+/-806se., with slope that is not significantly different from zero (-146.5+/-133s.e., p=0.283);see Figure 5.48.

Evo10_20_B03_CFPe_5: maximum growth rate

The maximum rate of growth of the evolved strain is very similar to the CFP strain in the ancestral competition. However, the ancestral YFP competing with Evo10_20_CFPe_B03_5 shows a depressed maximum rate across the range of ethanol stresses (Figure 5.49).

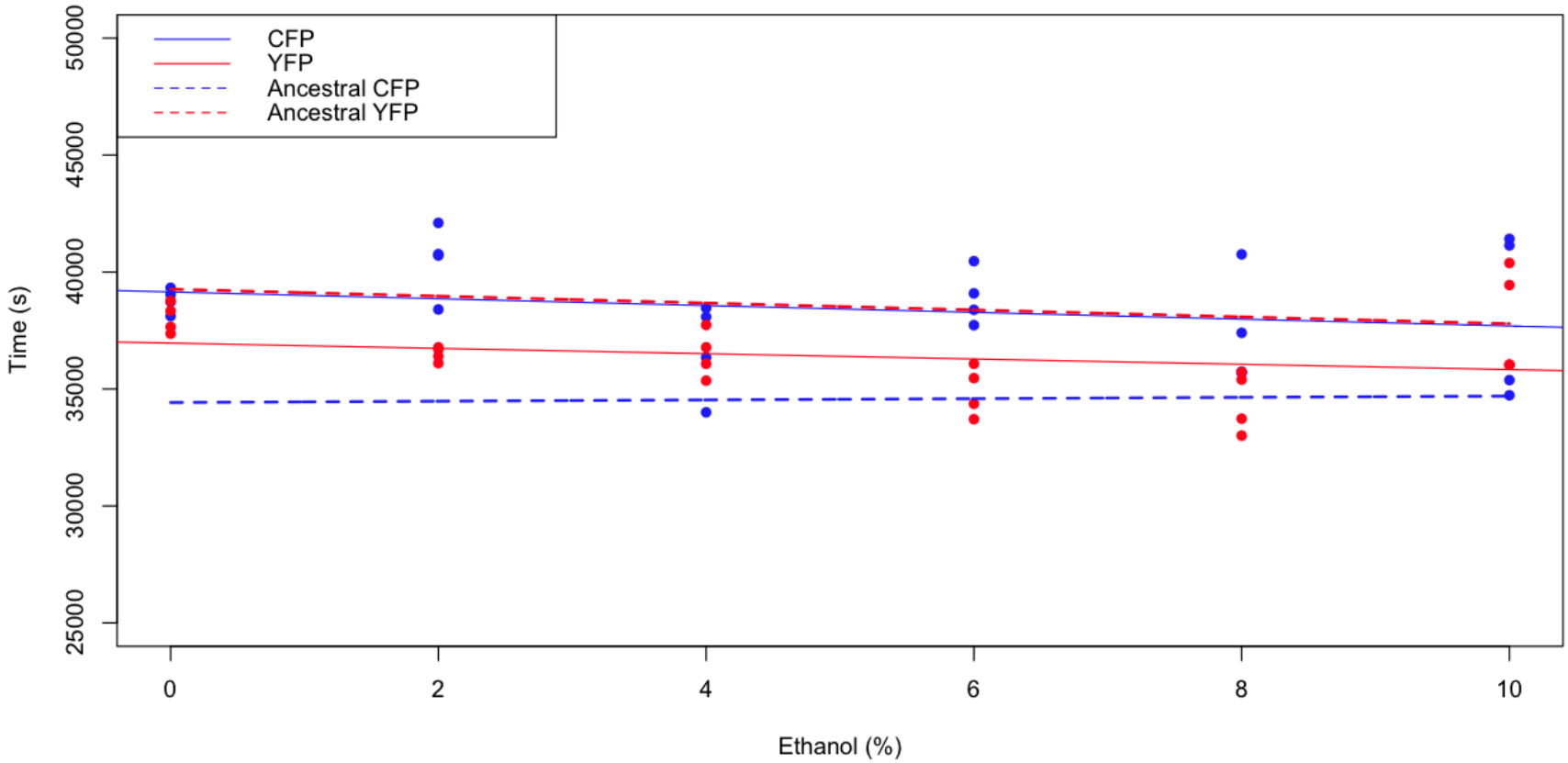


Figure 5.48 A plot of the elapsed time from detectable growth to onset of diauxy during the competition of the evolved strain Evo_10_20_B03_CFPe_5 and ancestral YFP strain in media with different ethanol concentrations.
The points in blue are for the Evo_10_20_B03_CFPe_5 and the points in red are for the ancestral YFP strain. The blue and red solid lines represent the linear best fit for the growth duration data for the Evo_10_20_B03_CFPe_5 and ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

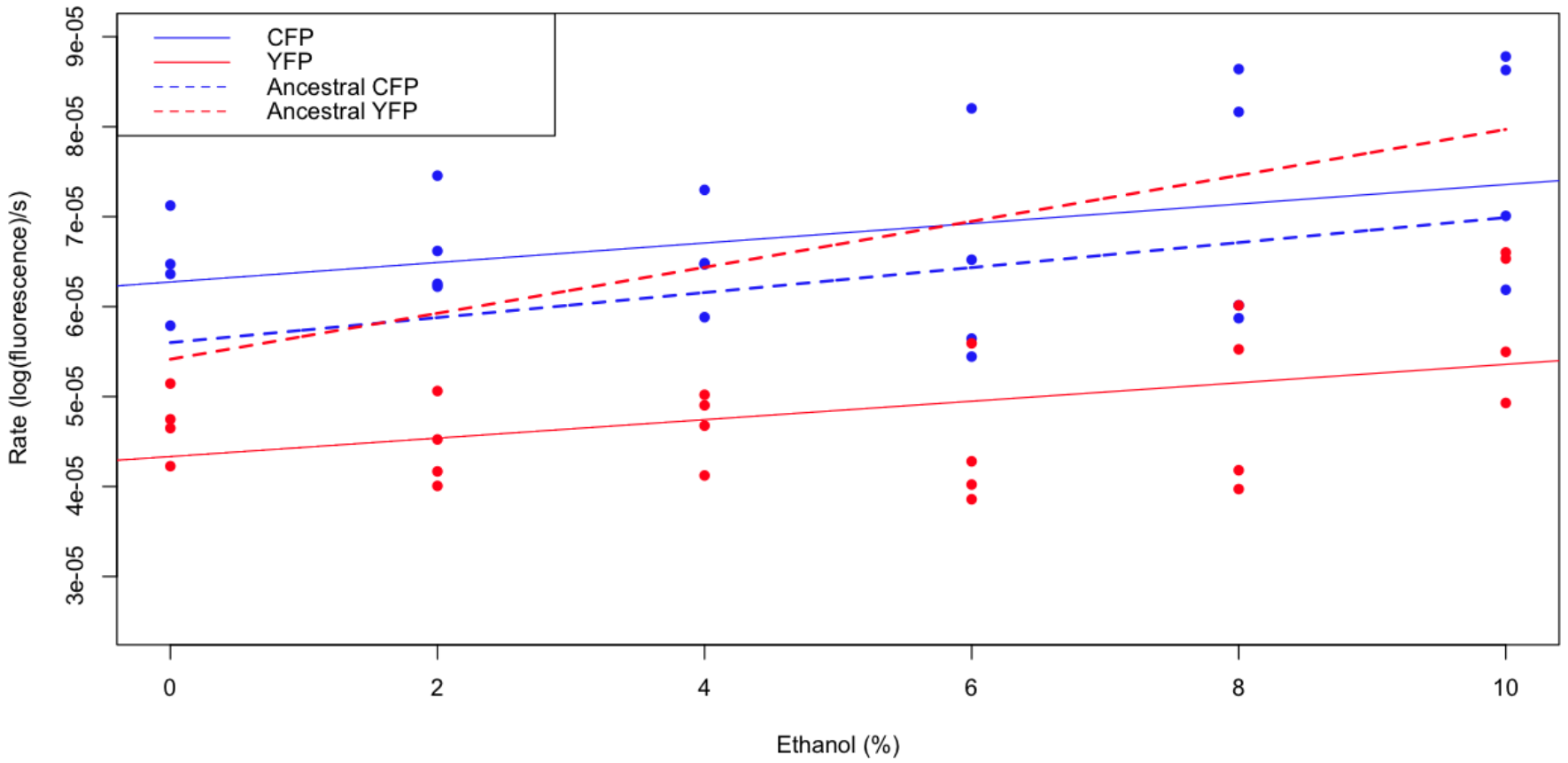


Figure 5.49 A plot of the maximum growth rate during the competition of the evolved strain Evo_10_20_B03_CFPe_5 and ancestral YFP strain in media with different ethanol concentrations.
The points in blue are for the Evo_10_20_B03_CFPe_5 and the points in red are for the ancestral YFP strain. The blue and red solid lines represent the linear best fit for the growth rate data for the Evo_10_20_B03_CFPe_5 and ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo10_20_B03_CFPe_5: fluorescence at 36 hours

The fluorescence at 36 hours is higher for Evo10_20_B03_CFPe_5 than its YFP ancestral competitor (see figure Figure 5.50). The intercept for the evolved strain in competition with ancestral YFP is 11.325+/-0.0517se vs that of the ancestral CFP intercept of 10.901+/-0.0438se (fluorescence values of 82867 vs 54230). The ancestral YFP fluorescence is depressed in competition Evo10_20_B03_CFPe_5. The slope of the evolved strain with ethanol is not significantly different from zero (0.005393+/-0.008544s.e., p=0.534).

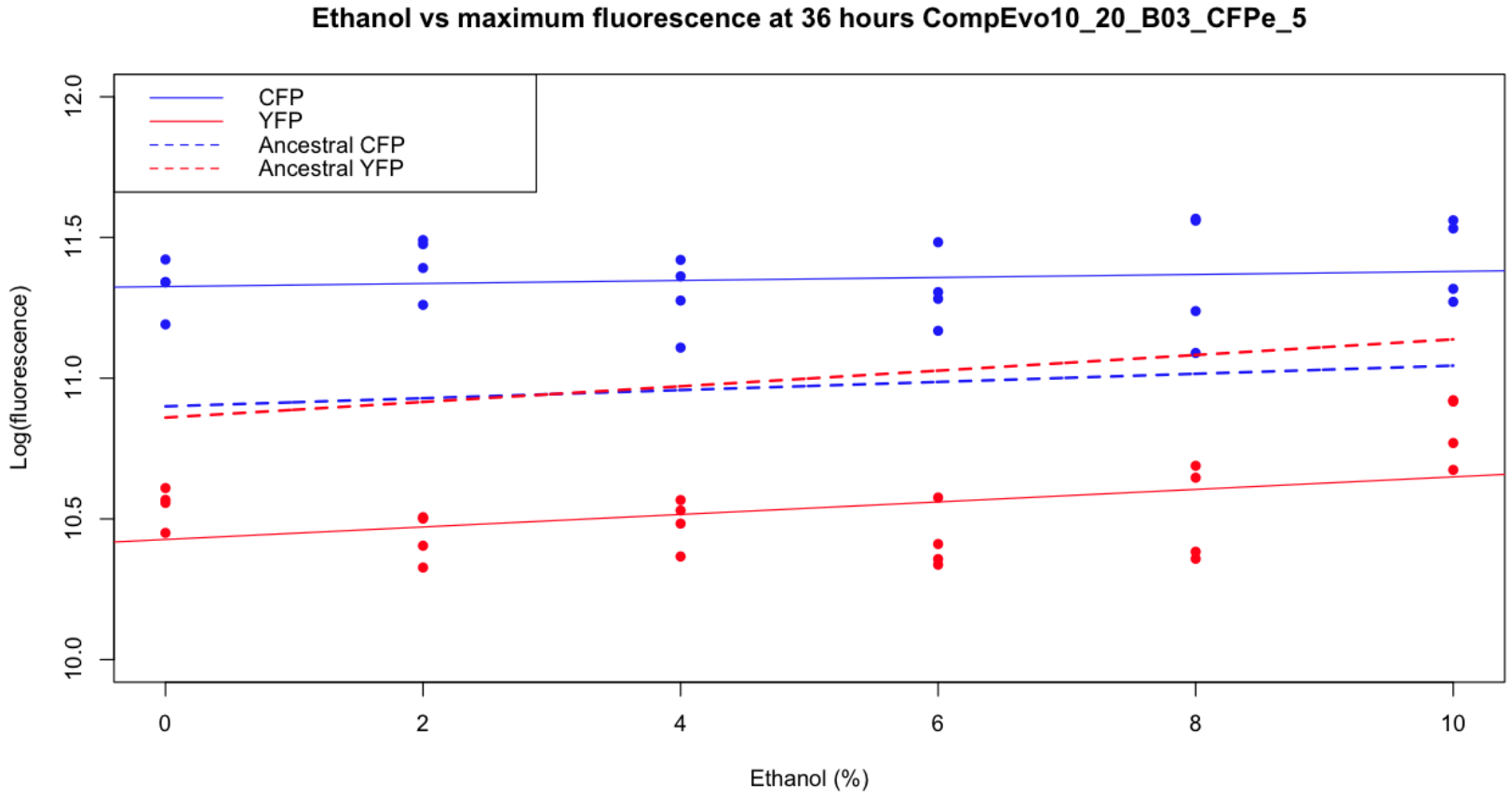


Figure 5.50 A plot of the fluorescence at 36 hours during the competition of the evolved strain Evo_10_20_B03_CFPe_5 and ancestral YFP strain in media with different ethanol concentrations.
The points in blue are for the Evo_10_20_B03_CFPe_5 and the points in red are for the ancestral YFP strain. The blue and red solid lines represent the linear best fit for the 36-hour fluorescence for the Evo_10_20_B03_CFPe_5 and ancestral YFP respectively. The dotted lines represent the linear best fit for the ancestral competitions on this plate. All competitions were run on a single plate on the BMG FluoStar plate reader.

Evo10_20_B03_CFPe_5: summary

In summary, Evo10_20_B03_CFPe_5 out-competes the YFP ancestor by depressing the growth rate of its competitor and having a significantly higher growth rate across the range of ethanol concentrations studied.

Evolution 10 summary

With no ethanol stress, the diploid strains are very similar in fitness (Evo10_20_F10_CFPe_2) or less fit (Evo10_20_G10_YFPe_2) than their ancestral competitors. As ethanol stress increases above ~4% both strains show improved fitness due to increases in growth time and rate. At ethanol stresses of 8%, Evo10_20_G10_YFPe_2 grows much more strongly than its ancestral competitor whilst the fitness increment of Evo10_20_F10_CFPe_2 is much smaller. The 'triploid' strain Evo10_20_B03_CFPe_5 grows only a little longer but at a much faster rate than its ancestral competitor. Its productivity as expressed by the fluorescence at 36 hours is consistent across the range of ethanol stresses from 0% to 10%

5.3.5. Histidine biosynthesis

The evolved strain Evo8_E07_CFPe_ was able to grow on minimal media without histidine as were two diploid strains produced by plasmid mating type switching that were tested, CFPd_17_1 and YFPd_4_1 (the four other plasmid derived diploids were not tested). However, both ancestral strains and all other evolved diploid strains were unable to grow on media requiring histidine biosynthesis.

5.3.6. Change of cell diameter with ethanol

Changing mean cell diameters with ethanol

For the starting CFP and YFP fluorophore strains (see Figure 5.51) there is a small decrease in the post-diauxic mean cell diameter with increasing fixed ethanol stress. However ancova analysis (see Table 5.2) suggests this is not significantly different from zero (gradient =-0.0114, p=0.52). The CFP strain is approximately 0.23 microns larger than the YFP strain (p=0.046*). There is no significant interaction between strain and ethanol (p=0.96).

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.678554	0.092908	39.594	<2e-16 ***
Ethanol	-0.011380	0.017679	-0.644	0.5222
fluoYFP	-0.232995	0.114666	-2.032	0.0466 *
ethanol:fluoYFP	-0.001029	0.022219	-0.046	0.9632

Table 5.2 Analysis of variance for the ancova model for ancestral cell diameter, medium ethanol concentration (diameter~ethanol*ploidy). The intercept is for the CFP ancestral strain. Residual standard error: 0.2464 on 60 degrees of freedom.

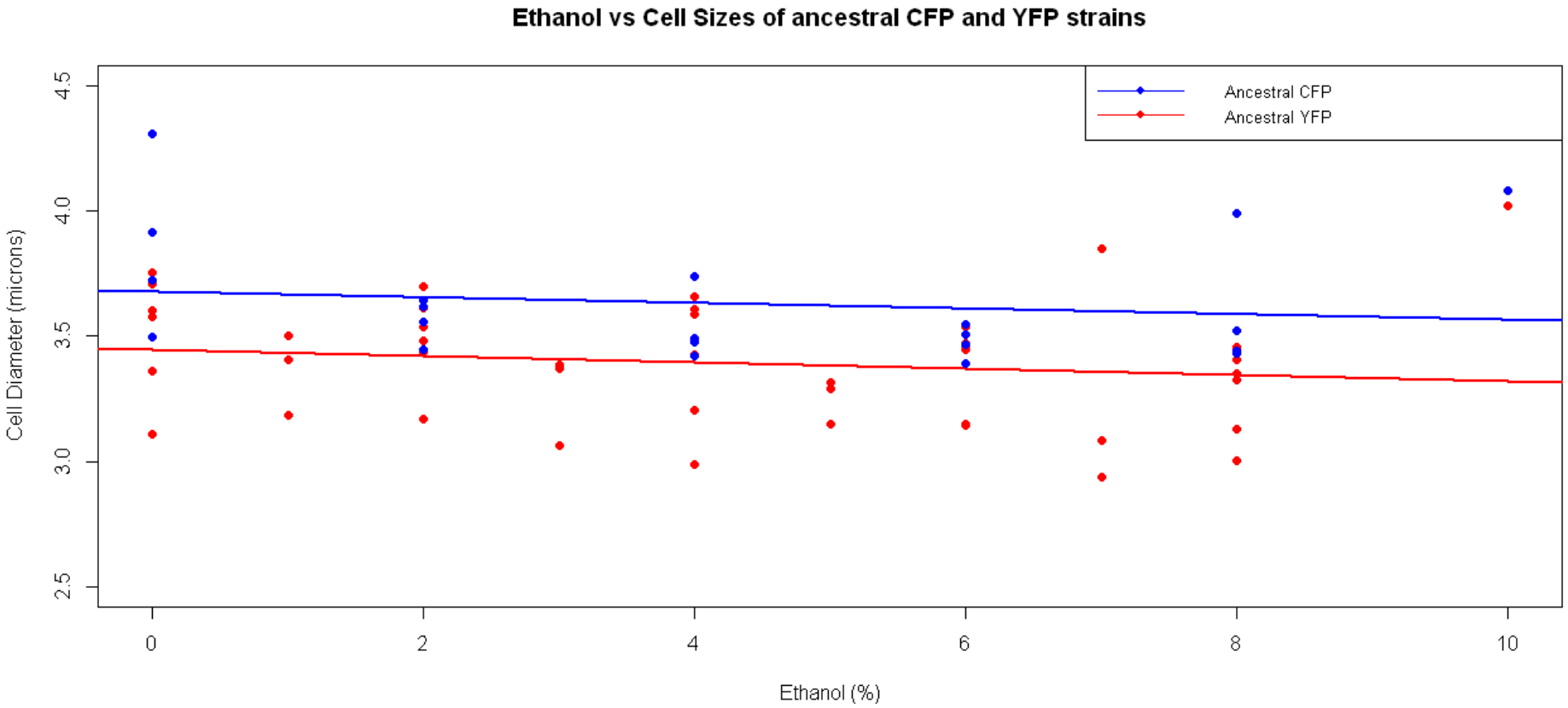


Figure 5.51 Mean cell diameters of ancestral CFP and YFP strains grown in varying ethanol concentrations.
Measurements of the CFP strain are in blue and the YFP strain in red. Linear fits of the cell sizes suggest the CFP strain is approximately 0.25 microns larger than the YFP.

All of the evolved strains that were isolated from evolutions had larger mean cell diameters than their CFP ancestors. All evolved strains had a larger negative gradient with increasing ethanol (see Figure 5.52). The CFP diploid strain DIA_17_1 constructed using plasmid mediated mating (green point) has a mean cell diameter without the plasmid typical of the evolved strains.

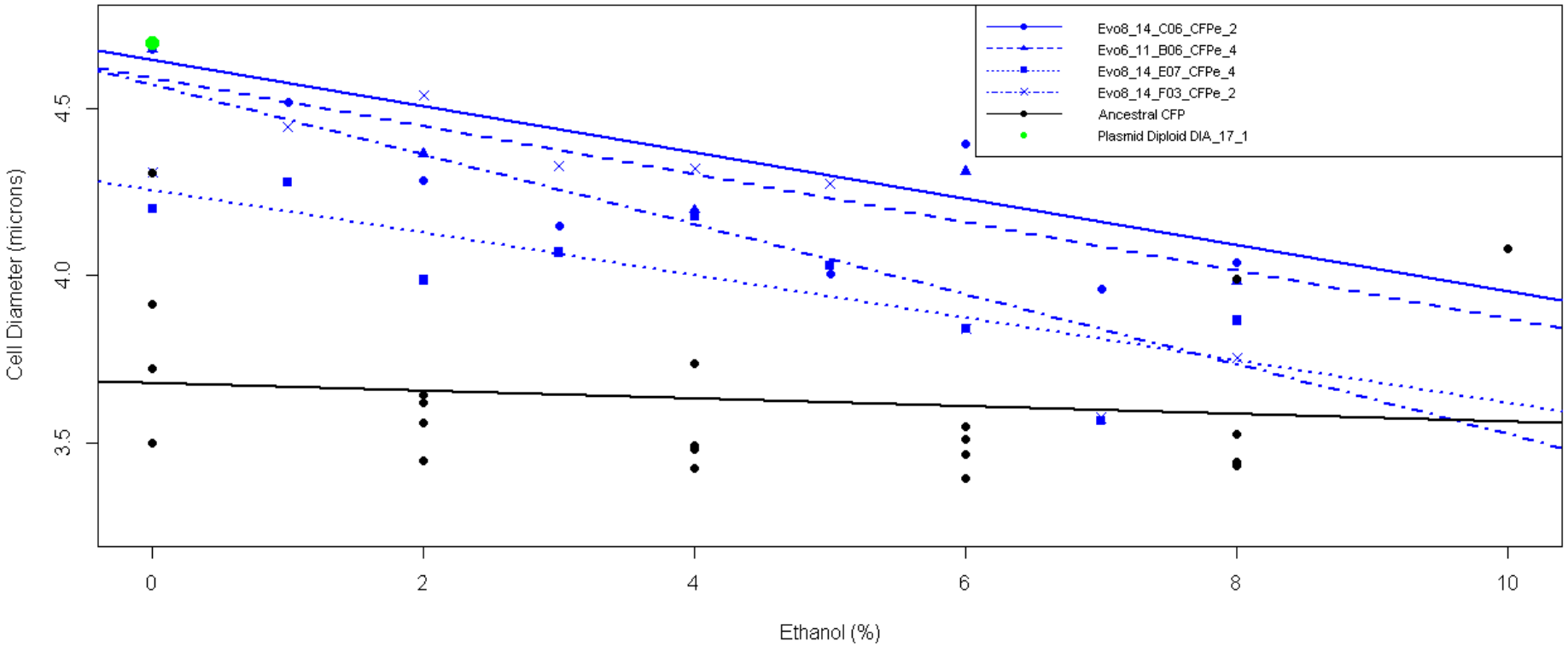


Figure 5.52 The mean strain cell diameters for yeast strains grown in varying ethanol concentrations; the experimentally evolved CFP strains from evolutions 6 and 8 are compared with their ancestral CFP strain.
The ancestral CFP strain (black points and lines) and evolved CFP strains from evolution series 6 and 8 (blue points and lines). The cell size of the diploid CFP strain DIA_17_1 grown in YPD is shown as a single green point.

When the single replicates from evolved CFP are combined and analysis of variance is used to fit the model: cell diameter ~ ethanol*ploidy

The model shows not just that ethanol and ploidy are strongly significant but also that the interaction between them is significant (p=0.023) (see Table 5.3).

.	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Ethanol	1	1.710	1.710	16.310	0.000168 ***
Ploidy	1	5.599	5.599	53.407	1.18e-09 ***
ethanol:ploidy	1	0.573	0.573	5.465	0.023069 *
Residuals	55	5.766	0.105		

Table 5.3 Analysis of variance for the ancova model for CFP cell diameter, medium ethanol concentration and ploidy (diameter~ethanol*ploidy)

When the model is simplified to remove the interaction between ploidy and ethanol concentration then the main effects are similar but the residuals are slightly bigger (see Table 5.4).

.	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Ethanol	1	1.710	1.710	15.11	0.000272 ***
Ploidy	1	5.599	5.599	49.46	3.05e-09 ***
Residuals	56	6.339	0.113		

Table 5.4 Analysis of variance for the ancova model for CFP cell diameter, medium ethanol concentration and ploidy model (diameter~ethanol + ploidy)

Comparison of the two models by anova shows that removing the interaction does significantly affect the fit of the model with p=0.023 (see Table 5.5)

.	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	55	5.7656				
2	56	6.3385	-1	-0.57288	5.4648	0.02307 *

Table 5.5 Anova comparison of ancova models of cell diameter, strain ploidy and ethanol content of the medium. Model 1: diameter ~ ethanol * ploidy; Model 2: diameter ~ ethanol + ploidy

When regression models are fit for both haploids and diploids there are clear differences. The intercept of the haploid regression model is 3.68 microns and the slope is not significantly different from zero (p=0.536), for coefficients see Table 5.6

.	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	3.67855	0.09479	38.806	<2e-16 ***
Ethanol	-0.01138	0.01804	-0.631	0.536

Table 5.6 Regression model diameter~ethanol for haploid CFP cells

Whilst the intercept of the diploid regression model is 4.62 (almost a micron larger) and the slope is significantly different from zero (p=0.00049), see Table 5.7.

.	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	4.61576	0.10157	45.445	< 2e-16 ***
Ethanol	-0.08228	0.02145	-3.836	0.000485 ***

Table 5.7 Regression of CFP diploid strains cell diameter against ethanol content of the media fitting the model diameter~ethanol.

A speculative extrapolation of the best fit lines suggests a convergence of the cell sizes of the ancestral and evolved strains close to a 12% fixed ethanol stress (see Figure 5.53 overleaf). It is possible that there is a physiological lower bound to the population cell size which is close enough to the haploid cell size to skew the distribution of cell sizes.

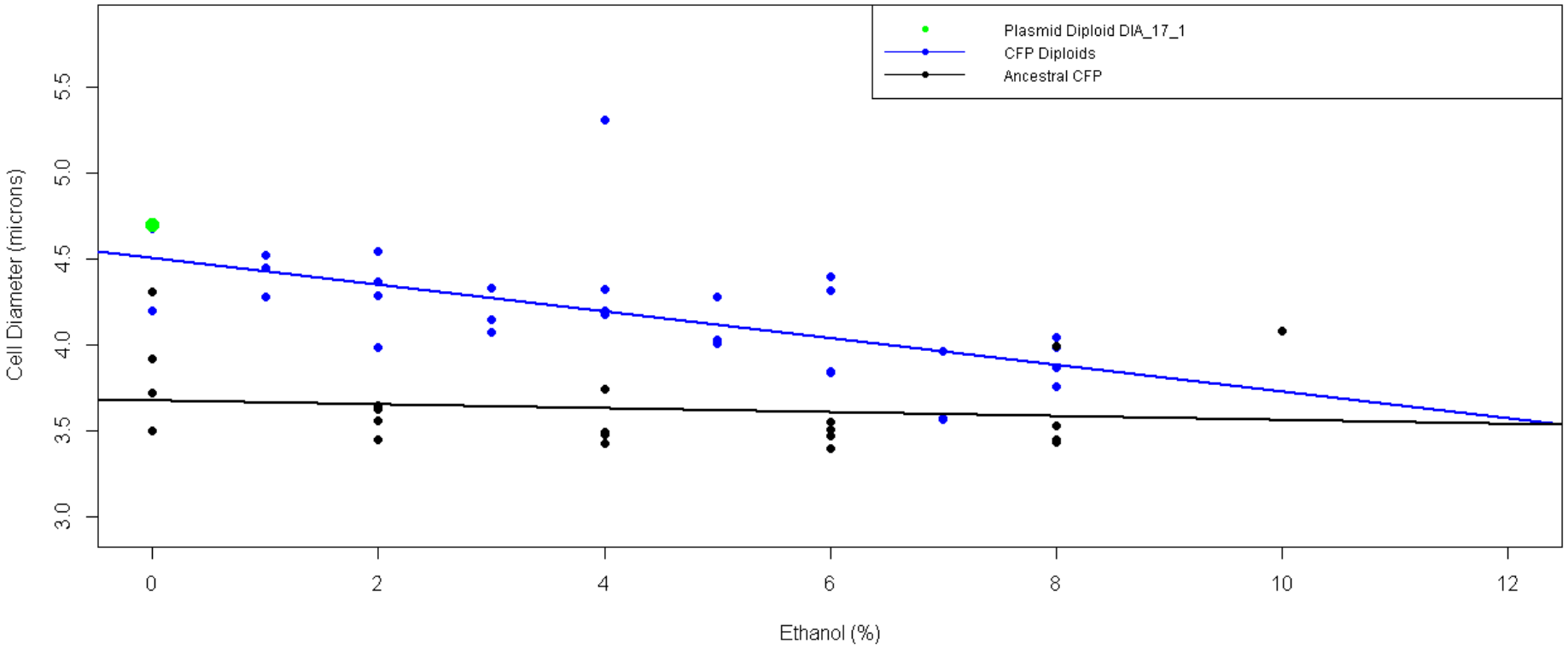


Figure 5.53 Mean cell diameters for CFP strains grown in varying ethanol concentrations, showing a single best straight line fit for all experimentally evolved strains from evolutions 6 and 8.
The cell sizes of ancestral CFP strains are shown as black points and lines and evolved CFP strains from evolution series 6 and 8 (blue points and lines). The cell size of the histidine prototrophic diploid CFP strain DIA_17_1 grown in YPD is shown as a green point.

The evolution 10 CFP strain, Evo_10_20_F10_CFPe_2, appears larger than the evolved strains from evolution 6 and 8 (Figure 5.54) with a mean diameter without a fixed ethanol stress of 5.24 microns. When an analysis of variance is carried out solely on the diploid cell size data, there is a very strong effect of strain (p=3.26x10-5), see Table 5.8

.	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Ethanol	1	1.706	1.7060	26.158	1.55e-05 ***
Strain	4	2.535	0.6338	9.718	3.26e-05 ***
Residuals	31	2.022	0.0652		

Table 5.8 Analysis of variance fitting the model cell diameter~ethanol + strain for all CFP diploids including Evo_10_20_F10_CFPe_2

However, when the Evo_10_20_F10_CFPe_2 data is removed the effect is much smaller, though still significant at p=0.02 (see Table 5.9).

.	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Ethanol	1	1.3372	1.3372	21.246	8.7e-05 ***
Strain	3	0.6927	0.2309	3.668	0.0245 *
Residuals	27	1.6994	0.0629		

Table 5.9 Analysis of variance Diploids fitting the model: cell diameter~ethanol + strain for CFP diploids excluding Evo_10_20_F10_CFPe_2

Table 5.10 shows the ancova output for the model complete analysis of covariance for the model: cell diameter ~ ethanol * strain. The largest intercept is for the Evo_10_20_F10_CFPe_2 strain.

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	5.239810	0.208004	25.191	< 2e-16 ***
ethanol	-0.097724	0.042459	-2.302	0.029303 *
strainEvo6_11_B06_CFPe_4	-0.645190	0.294162	-2.193	0.037076 *
strainEvo8_14_C06_CFPe_2	-0.591199	0.265531	-2.226	0.034519 *
strainEvo8_14_E07_CFPe_4	-0.983788	0.265531	-3.705	0.000961 ***
strainEvo8_14_F03_CFPe_2	-0.665941	0.265531	-2.508	0.018456 *
ethanol:strainEvo6_11_B06_CFPe_4	0.025578	0.060046	0.426	0.673500
ethanol:strainEvo8_14_C06_CFPe_2	0.028408	0.054814	0.518	0.608496
ethanol:strainEvo8_14_E07_CFPe_4	0.034137	0.054814	0.623	0.538648
ethanol:strainEvo8_14_F03_CFPe_2	-0.007142	0.054814	-0.130	0.897293

Table 5.10 Ancova analysis of CFP diplods fit to the model: cell diameter~ethanol*strain.
The maximum intercept is that for the strain Evo_10_20_F10_CFPe_2. Residual standard error 0.2685 on 27 degrees of freedom

There is no significant interaction between strain and ethanol; the slopes do not significantly vary with strain. Removing this interaction term improves the significance of the intercepts (Table 5.11).

.	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	5.17293	0.13061	39.607	< 2e-16 ***
Ethanol	-0.08100	0.01584	-5.114	1.55e-05 ***
strainEvo6_11_B06_CFPe_4	-0.54288	0.16152	-3.361	0.00207 **
strainEvo8_14_C06_CFPe_2	-0.47757	0.14245	-3.353	0.00212 **
strainEvo8_14_E07_CFPe_4	-0.84724	0.14245	-5.948	1.42e-06 ***
strainEvo8_14_F03_CFPe_2	-0.69451	0.14245	-4.876	3.07e-05 ***

Table 5.11 Ancova analysis of CFP diplods fit to the model: cell diameter~ethanol + strain.
The maximum intercept is that for the strain Evo_10_20_F10_CFPe_2. Residual standard error: 0.2554 on 31 degrees of freedom.

The intercept for the Evo_10_20_F10_CFPe_2 in this minimal model is 5.17µm, and all the other CFP strains evaluated have significantly smaller intercepts with the most significant reduction (0.85µm, $p=1.42x10^{6}$) being that for the strain isolated from the evolution 8 control with no added ethanol Evo8_14_E07_CFPe_4. Larger still than Evo_10_20_F10_CFPe_2 is the evolution 10 triploid strain, Evo10_20_B03_CFPe_5, which without an added ethanol stress has a mean cell diameter of 5.81microns (see Figure 5.54 overleaf). The ancestral cell size of the CFP strain is shown as black points with a best-fit line. The blue circular points and best fit line shows data from evolved CFP strains from evolution series 6 and 8. The squared points and dotted best fit line are from strain Evo_10_20_F10_CFPe_2 and the triangular points are from strain Evo10_20_B03_CFPe_5. The histidine prototrophic diploid CFP strain DIA_17_1 grown in YPD is shown as a green point.

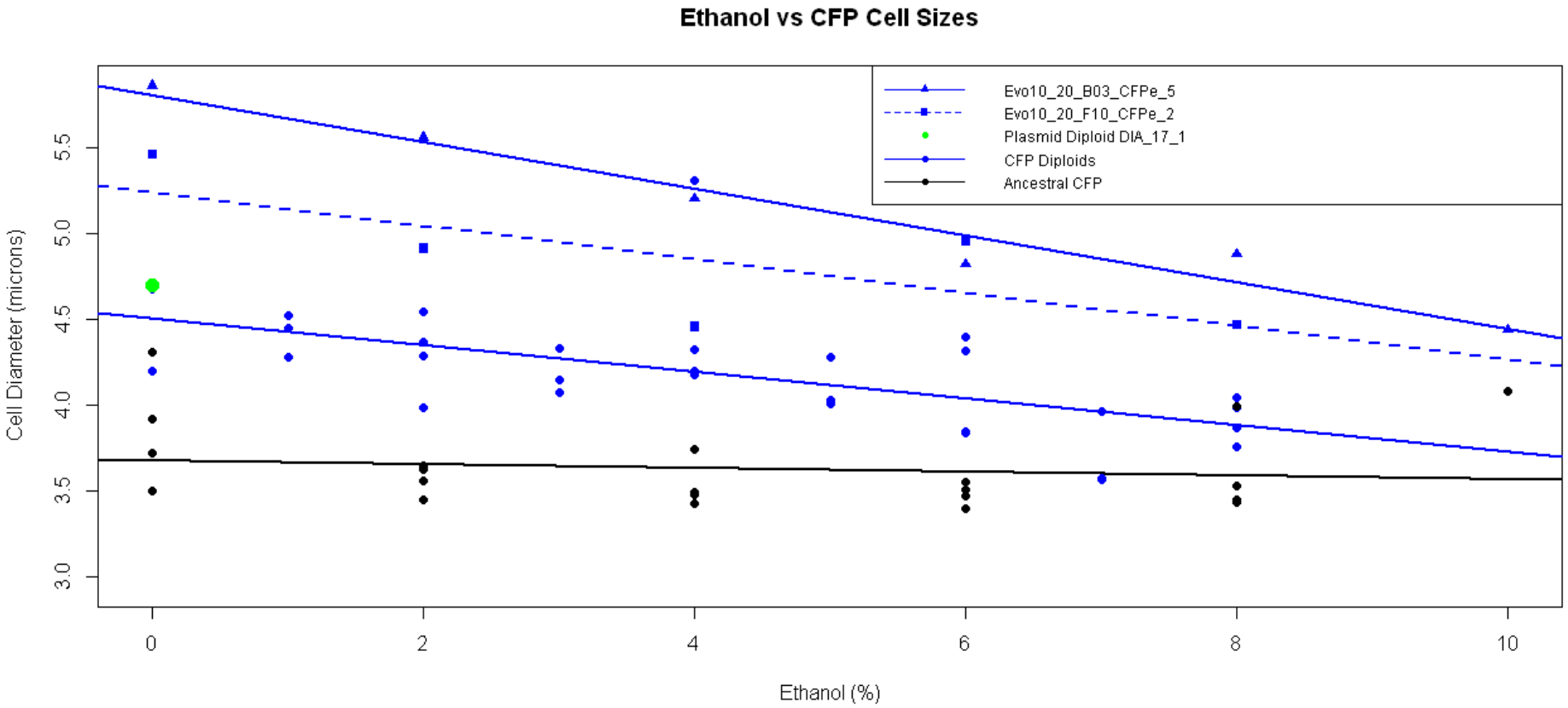


Figure 5.54 Mean cell diameters against ethanol concentrations for the experimentally evolved CFP strains Evo_10_20_F10_CFPe_2 and Evo10_20_B03_CFPe_5 compared with the ancestral CFP and other evolved CFP strains.

The evolved YFP strain from evolution 6, Evo6_11_D06_YFPe_4 is very similar to the evolved CFP strains from evolution 6 and 8 (see Figure 5.55 overleaf). The two isolates of the plasmid derived YFP strain DIA_4_1 are close to the evolved YFP strain (4.59 and 4.95 microns) without an added ethanol stress.

Again the mean diameter of the evolution 10 YFP strain, Evo10_20_G10_YFPe_2, is somewhat larger than the evolution 6 and evolution 8 evolved strains without added ethanol at 5.54 microns and similar to Evo10_20_F10_CFPe_2 (5.46 microns). However, whilst the slope of the decrease in cell size is apparently steeper (see Figure 5.56 overleaf), the different in slope is not significant (p=0.2).

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	5.23981	0.20340	25.762	2.26e-07 ***
Ethanol	-0.09772	0.04152	-2.354	0.0568
strainEvo10_20_G10_YFPe_2	0.11562	0.28765	0.402	0.7017
ethanol:strainEvo10_20_G10_YFPe_2	-0.07667	0.05872	-1.306	0.2395

Table 5.12 Ancova analysis of evolution 10 diploids strains fit to the model: cell diameter~ethanol * strain.
The maximum intercept is that for the strain Evo_10_20_F10_CFPe_2. Residual standard error: 0.2626 on 6 degrees of freedom

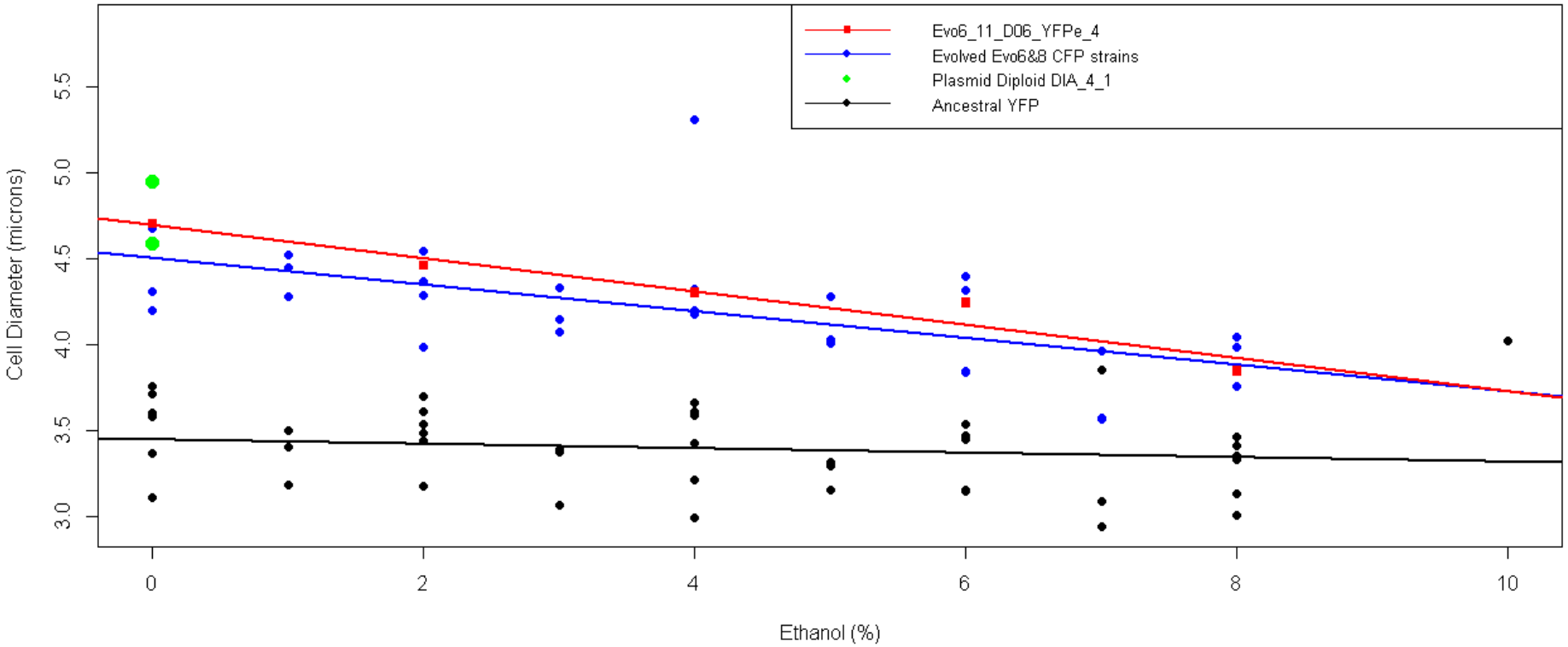


Figure 5.55 Mean cell diameters for YFP strain Evo6_11_D06_YFPe_4 and evolved CFP strains from evolutions 6 and 8 grown in varying ethanol concentrations. The ancestral YFP strain (black points and line), Evo6_11_D06_YFPe_4 (red squares and line), and evolved CFP strains from evolution series 6 and 8 (blue points and line) and the his+ diploid CFP strain DIA_4_1 (green points).

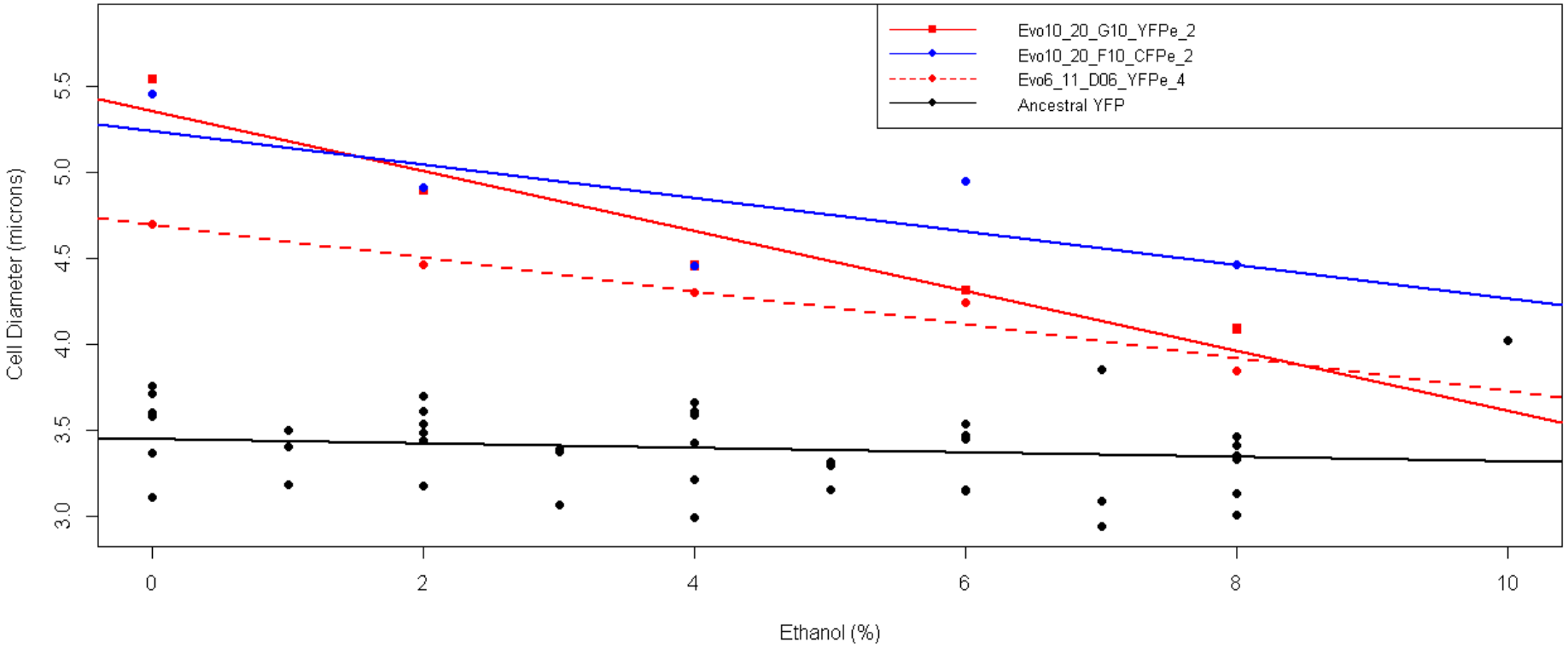


Figure 5.56 Ethanol vs mean cell diameter in microns for ancestral and evolved YFP strains
Plot shows the ancestral strains (black points), the evolved YFP strain Evo10_20_G10_YFPe_2 (red squares), evolved CFP strain Evo10_20_F10_CFPe_2 (blue points), Evo6_11_D06_YFPe_4 (red points).

5.3.7. Cell size summary

Slow growth (post-dauxic shift) cell populations have different mean sizes and have different responses to ethanol. The ancestral CFP strains are slightly but significantly (~0.23 microns) bigger than the YFP fluorophore ancestors. While neither haploid ancestral strain shows any significant change in cell size as the ethanol content of the medium increases, the experimentally evolved strains characterised, which are all diploid (see Chapter 6, section 6.4.2) are larger without ethanol added to the growth medium but do have a negative gradient; mean cell size shrinks with ethanol stress. The two strains isolated and characterised from evolution 10 are larger than those from evolution 6 and 8 across the ethanol concentration range tested.

Changing cell diameter distributions with ethanol

Post-diauxic parental CFP and YFP strains show marginal changes in their cell size distribution when they are grown with added ethanol. The ancestral YFP strain shows a small reduction in larger cell sizes while the CFP strain shows a slight increase (see Figure 5.57 overleaf). These strains are haploid

The evolution 6 strains, Evo6_11_B06_CFPe_4 and Evo6_11_D06_YFPe, both show substantial reductions in the proportion of larger cells in the mean cell diameter distribution (see Figure 5.58 overleaf).

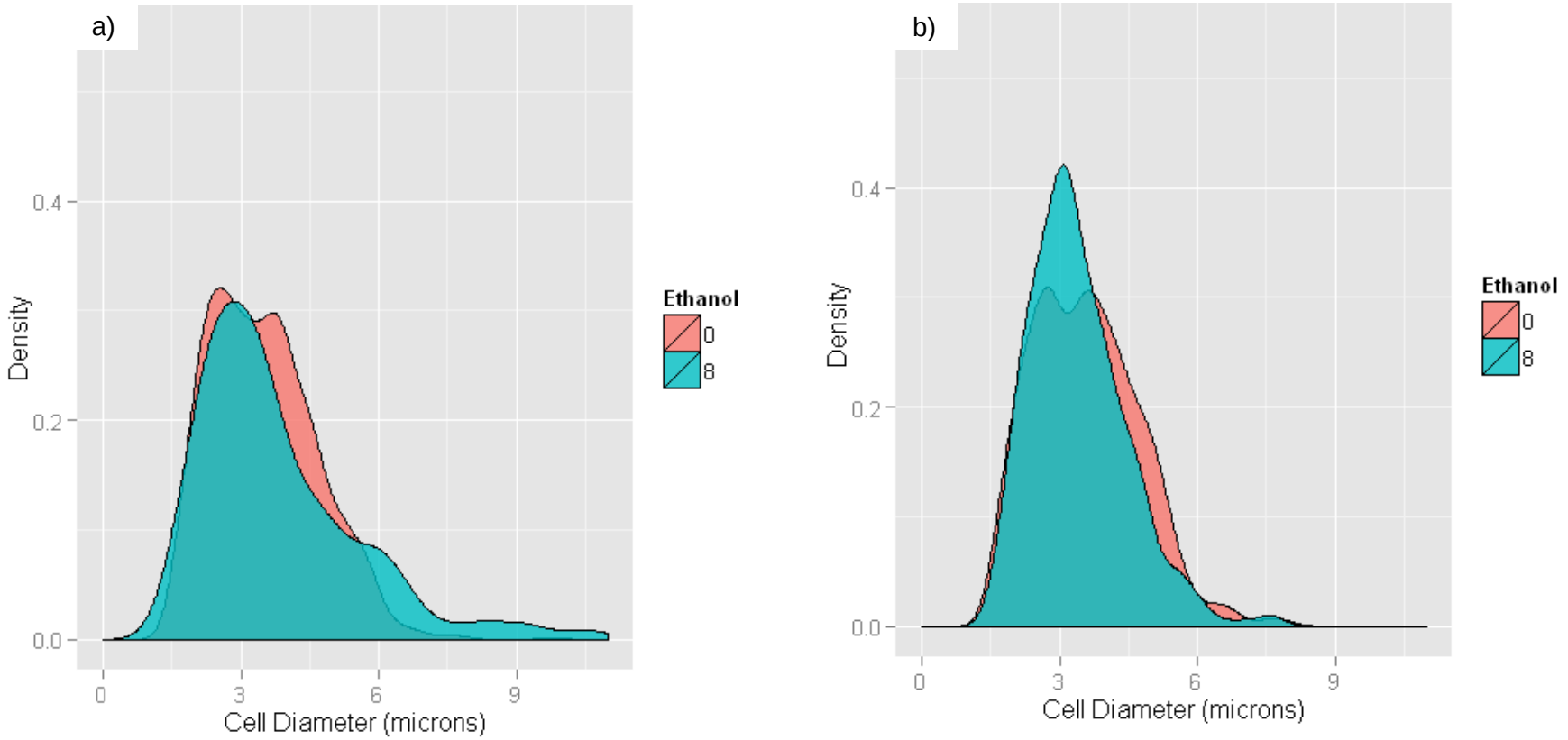


Figure 5.57 Cell diameter distribution of post-diauxy ancestral CFP and YFP strains grown in YPD with 0% or 8% added ethanol.
Figure a) shows the CFP starting strain and figure b) the YFP starting strain. For both strains, the salmon coloured distribution shows growth in YPD with no added ethanol while that coloured teal is that with growth in 8%. The CFP strain size distribution shows a slight increase in the largest cell sizes with 8% ethanol while the YFP strain is largely unchanged.

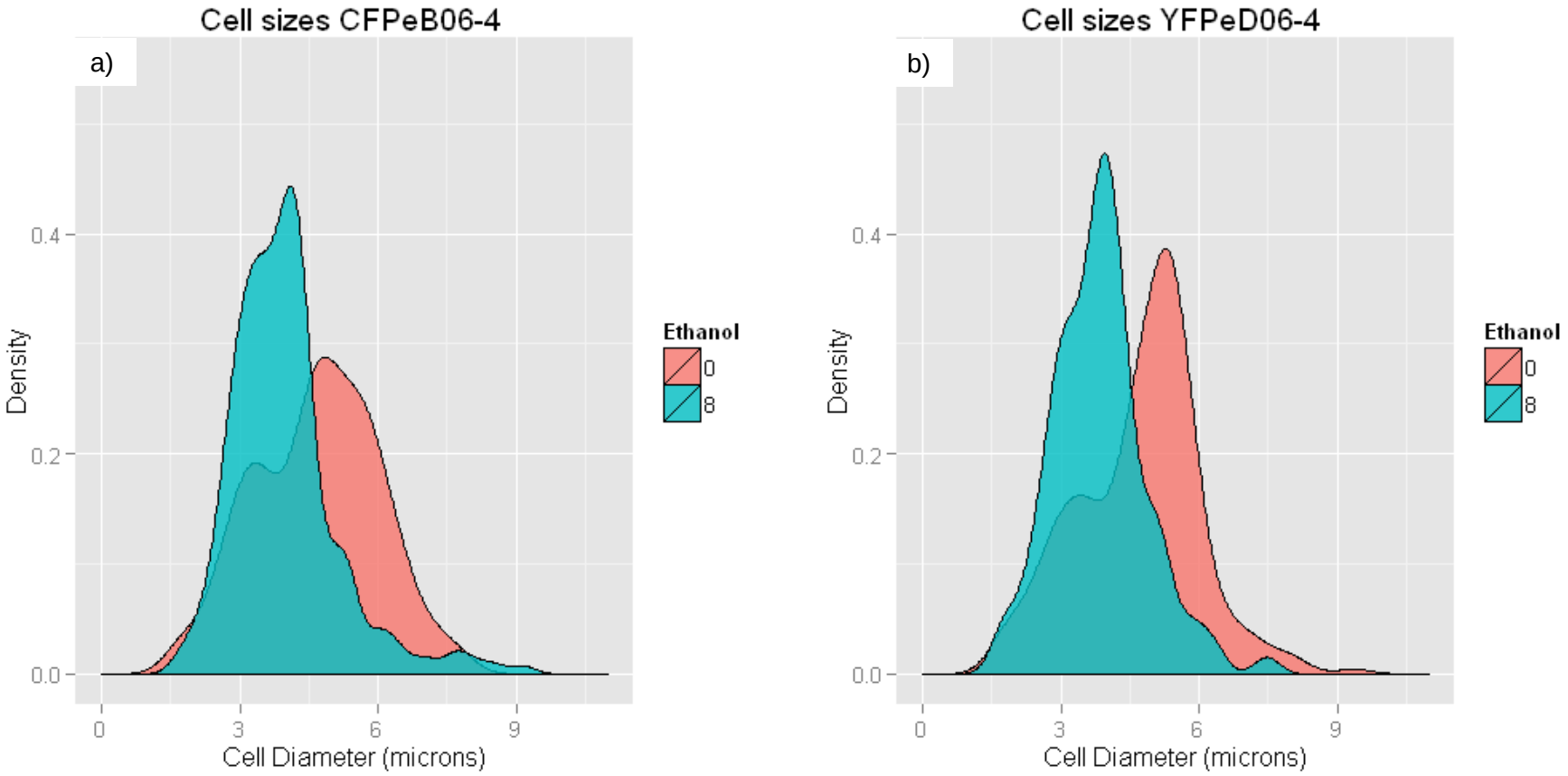


Figure 5.58 The cell size distribution of the experimentally evolved strains from evolution 6 (Evo6_11_B06_CFPe_4 and Evo6_11_D06_YFPe) grown in YPD containing 0% or 8% ethanol.

For both strains, the salmon coloured distribution shows growth in YPD with no added ethanol while that coloured teal is that with growth in 8%. Both strains (evolved in 4.5% ethanol) show a substantial reduction in the proportion of cells of larger size. Figure a) shows the mean cell diameter distribution for Evo6_11_B06_CFPe_4 and figure b) shows that for Evo6_11_D06_YFPe.

This reduction in large cell sizes with ethanol is very similar for the evolution 8 strains Evo8_14_C06_CFPe_2 and Evo8_14_F03_CFPe_2. Both these strains were experimentally evolved in the same 4.5% ethanol treatment as evolution 6 (see Figure 5.59).

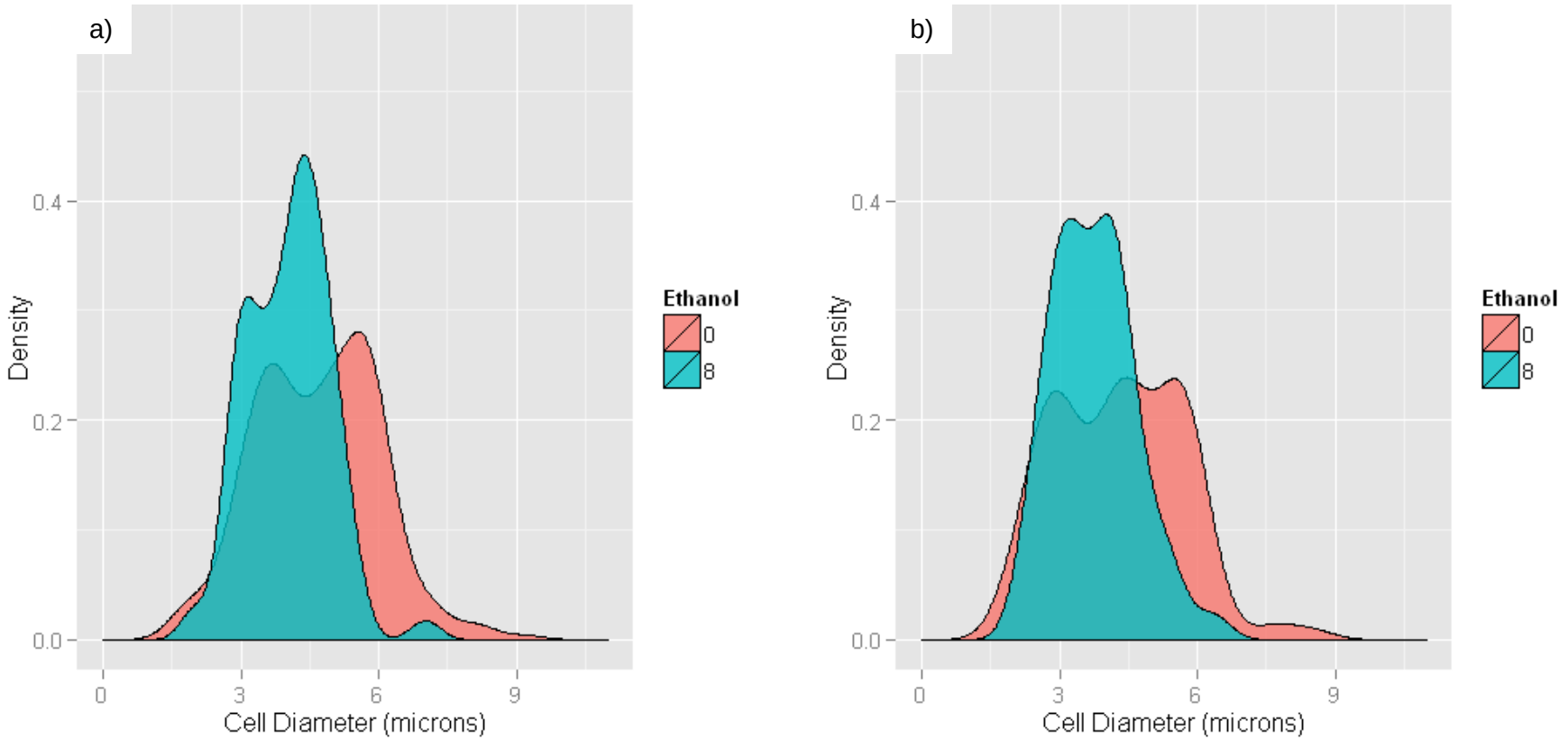


Figure 5.59 The cell size distribution of the experimentally evolved strains from evolution 8 (Evo8_14_C06_CFPe_2 and Evo8_14_F03_CFPe_2) grown in YPD containing 0% or 8% ethanol.
For both strains, the salmon coloured distribution shows growth in YPD with no added ethanol whilst that coloured teal is growth with 8% ethanol. Both strains (evolved in 4.5% ethanol) show a substantial reduction in the proportion of cells of larger size. Figure a) shows the mean cell diameter distribution for Evo8_14_C06_CFPe_2 and figure b) shows that for Evo8_14_F03_CFPe_2.

The evolution 8 CFP strain Evo8_14_E07_CFPe_4 (isolated from the 0% ethanol control treatment) shows the reduction in larger cells with 8% ethanol typical of the other evolved strains. Additionally however, the starting mean cell diameter distribution appears bimodal, with peaks in the distribution close to 3 and 5.8microns. This initial peak almost disappears as ethanol increases to 4%. As ethanol increases from 4% to 8% the larger peak shifts in the direction of smaller cell sizes.

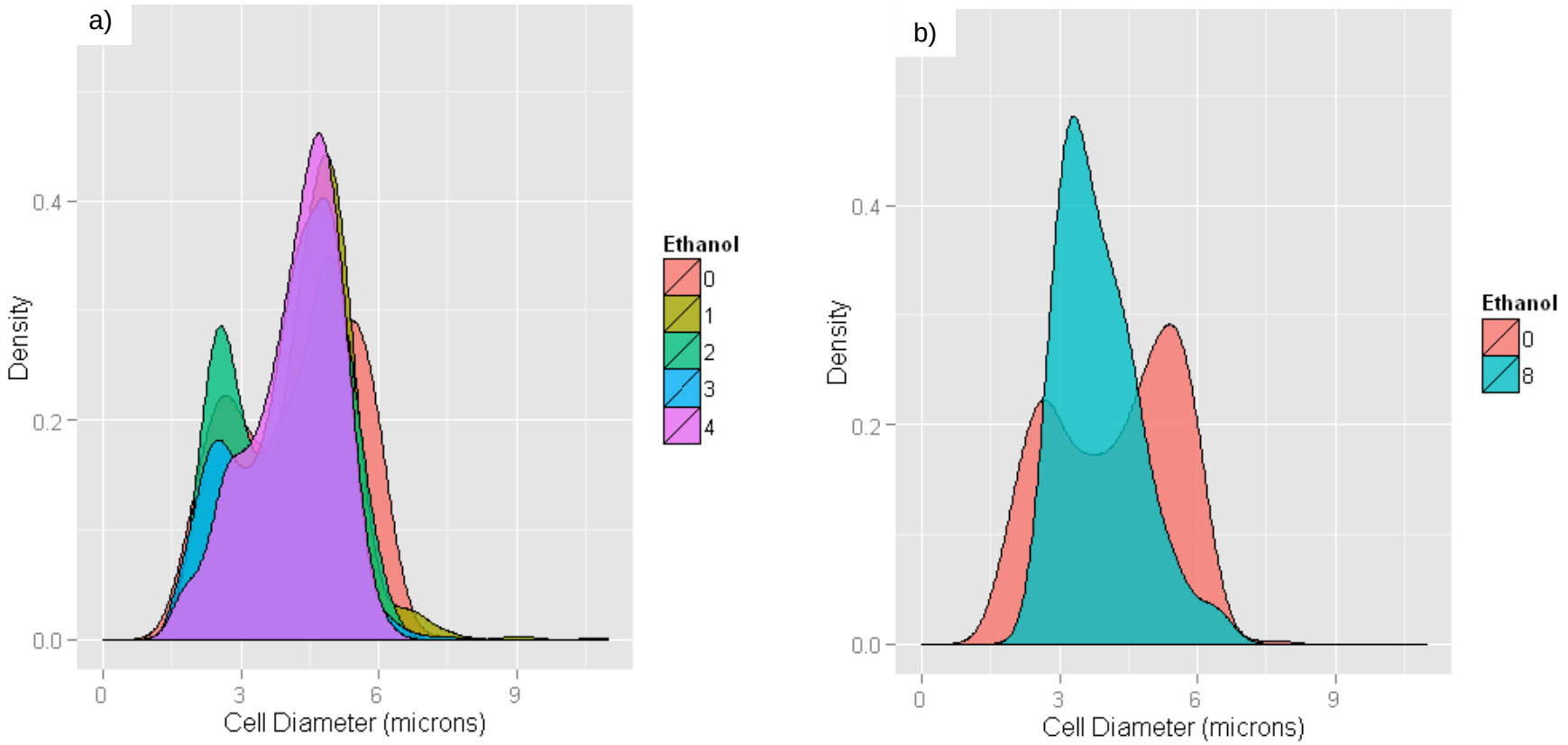


Figure 5.60 The mean cell diameter distribution of the experimentally evolved strain from evolution 8, Evo8_14_E07_CFPe_4, grown in YPD containing varying ethanol concentrations.
For figure a) the size distributions are shown for cells grown in YPD with 0%,1%,2%,3% or 4% ethanol; in salmon, olive green, light green, blue and purple respectively. Figure a) shows the bimodal distribution without ethanol increasing in size as it becomes unimodal. For figure b) the salmon coloured distribution shows growth in YPD with no added ethanol while that coloured teal is that with growth in 8%. Overall, increasing ethanol from 0 to 8% reduces the range of the cell diameter distribution so that the smallest and largest cells become much rarer (e.g. 6 micron cells) or absent (0.75 micron cells).

Without ethanol, the evolution 10 strains are larger than those from the other evolutions, 6 and 8. With no added ethanol, only a minority of the evolved strains from evolutions 6 and 8 are larger than 6 microns but a substantial proportion of the evolution 10 strains are larger (see Figure 5.58, Figure 5.59, Figure 5.60 for evolutions 6 and 8 and Figure 5.61 and Figure 5.62 for evolution 10, see also Figure 5.54 and Figure 5.56 which show that evolution 10 strains have the largest mean cell diameter values). Figure 5.61 shows the cell diameter distributions for the evolution 10 strains, Evo10_20_F10_CFPe_2 (Figure 5.61.a) Evo10_20_G10_YFPe_2 (Figure 5.61.b); both show the typical reduction in cell size with ethanol in the growth medium.

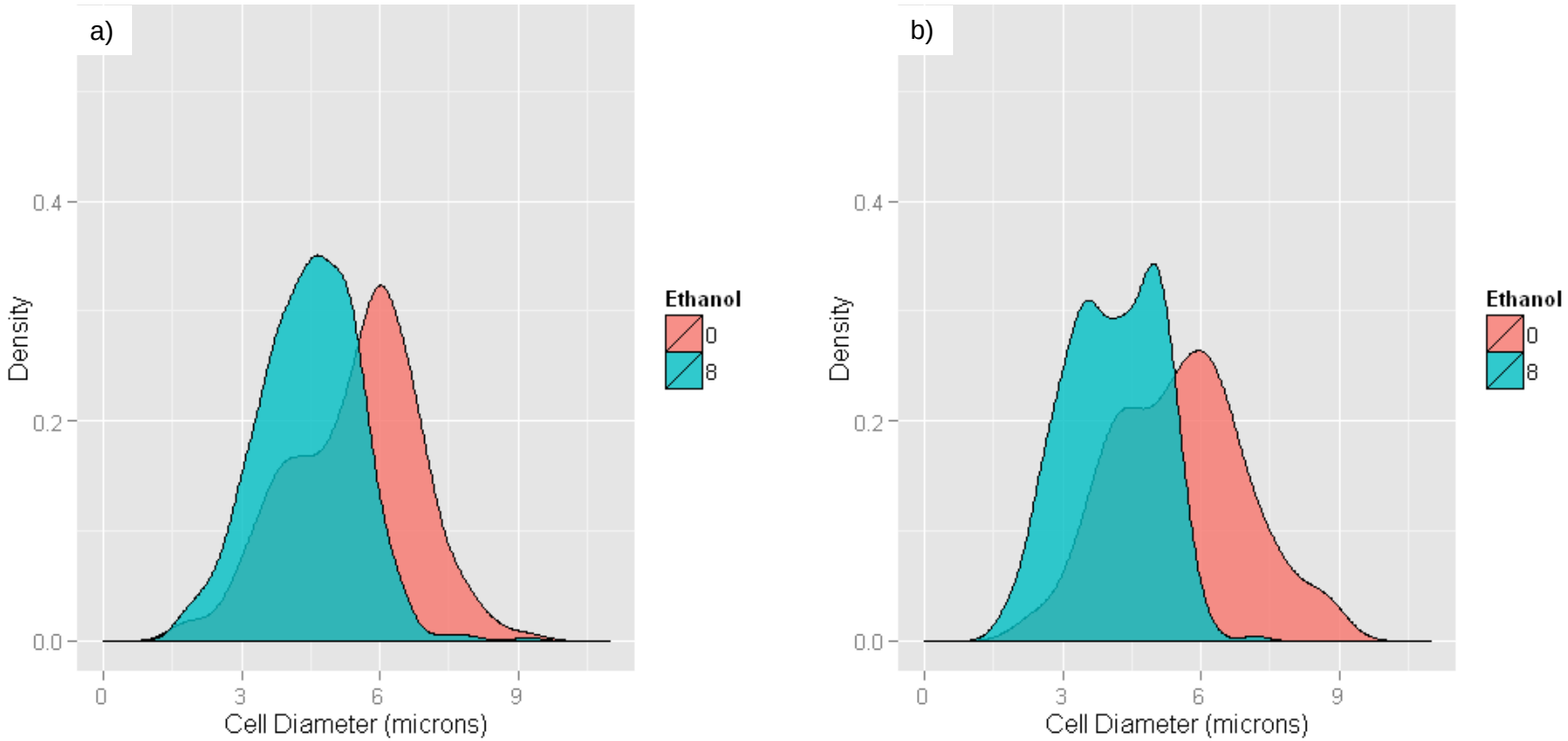


Figure 5.61 The cell size distribution of the experimentally evolved strains from evolution 10 (Evo10_20_F10_CFPe_2 and Evo10_20_G10_YFPe_2) grown in YPD containing 0% or 8% ethanol.
For both strains, the salmon coloured distribution shows growth in YPD with no added ethanol whilst that coloured teal has 8% ethanol. Figure a) shows the mean cell diameter distribution for Evo10_20_F10_CFPe_2 and figure b) shows that for Evo10_20_G10_YFPe_2. Both strains show a reduction in cell size with 8% added ethanol.

Figure 5.62 (overleaf) shows the cell size distribution for the evolution 10 strain Evo10_20_B03_CFPe_5 which also reduces significantly with ethanol stress. Subsequent FACS analysis suggests that this strain is broadly triploid in DNA content (see chapter 6, section 6.3.8).

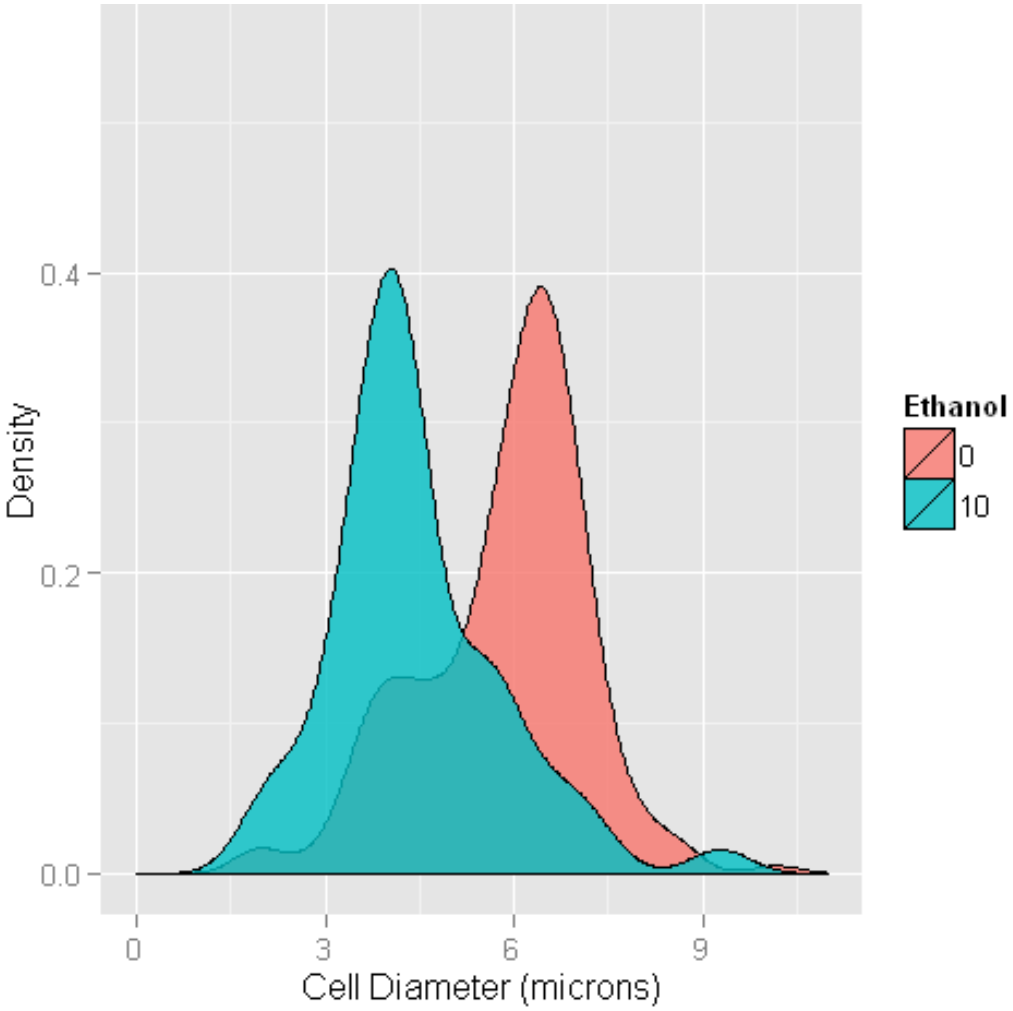


Figure 5.62 The cell size distribution of the experimentally evolved strains from evolution 10 strain Evo_10_20_B03_CFPe_5 grown in YPD containing 0% or 10% ethanol.
The salmon coloured distribution shows growth in YPD with no added ethanol whilst that coloured teal is growth in 10%.

5.3.8. Change in size profile throughout fermentation

The mean cell diameters during log phase growth are smaller than during the slow growth phase for both ancestral YFP strain and the evolved strain Evo10_F10_CFPe_2 (see Figure 5.63). However, the absolute reduction in cell size Cytometer slide 1 in figure 11.60 is the size profile of the starting culture while slides 2 and 3 are in the log phase (at 4hrs 50min and 5hrs 50min respectively). For each strain the log phase growth profiles are very similar but shifted to smaller sizes. The mean starting cell diameter in slide 1 for Evo10_F10_CFPe_2 is 6.43 microns but in slides 2 and 3 the log phase values are 4.08 and 3.89 microns respectively, a 37-40% reduction in mean diameter. The mean starting diameter of the starting YFP strain was 3.45 microns which reduced to 3.10 and 3.25 microns, a reduction of 10-6% respectively. Despite the larger absolute and percentage decrease in size for the evolved strain, the mean and mode of the evolved CFP strain is larger than the YFP ancestor.

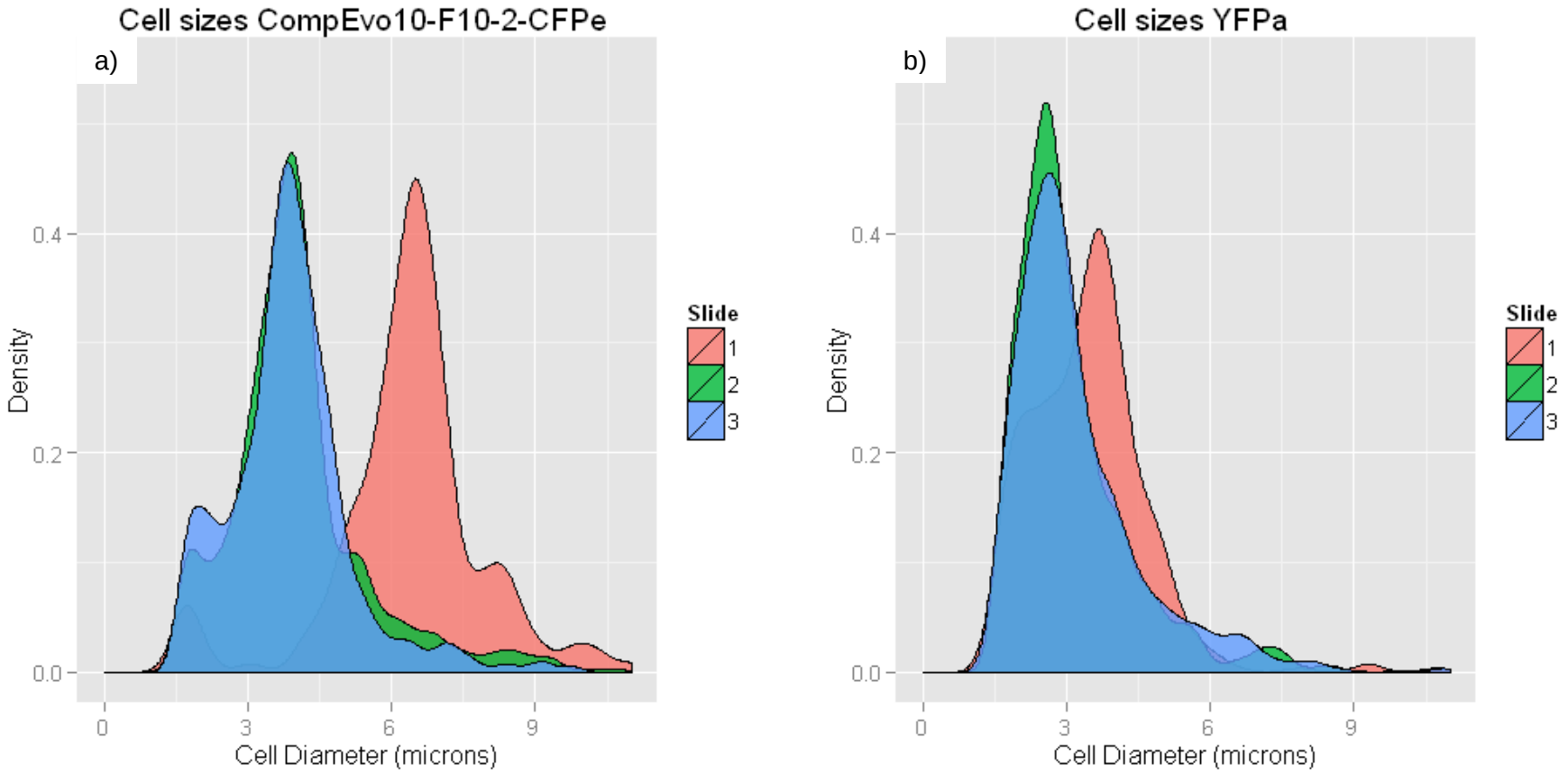


Figure 5.63 Change in mean cell diameters (microns) for Evo10_20_F10_CFPe_2 and the ancestral YFP strains during a fermentation in YPD medium without added ethanol.

Figure a) shows the cell size distribution for Evo10_20_F10_CFPe_2 and figure b) shows the ancestral YFP distribution. The salmon colour denotes the starting cell culture (slide 1), the distribution in green is from 4h 50min (slide 2) and that in in blue is from 5hrs 50min (slide 3). The later distributions were both in the log phase of growth.

5.3.9. Biolog GenIII plate phenotyping - overview

Most of the GenIII plate treatment wells had little growth and in the best growing wells the OD_{600} change was generally 0.2-0.4 from an initial mean of 0.14. Each Gen III plate has an empty well 'negative control' and a 'positive control'. Biolog provide a description of the well contents (see appendix II, section 9.2) but the concentrations and exact contents of each well have not been disclosed.

Figure 5.64 shows the wells that grew best and the controls in all plates. Initially the data was blanked using the 'growth' in the negative control. However, whilst changes within it were small, the OD_{600} of the negative control usually falls, possibly due to cell lysis, by an amount that varies between strains and replicates, making this a suspect blank. For every strain, growth in the 'positive control' of each replicate was similar to the antibiotic well 'nalidixic acid'. The four wells 'pH5' and 'pH6' and the antibiotic wells 'aztreonam' and 'troleandomycin' all grow comparably well, better than the positive control. However, growth was best overall, and most consistent between strains and replicates, in the well containing the oxidising agent sodium bromate.

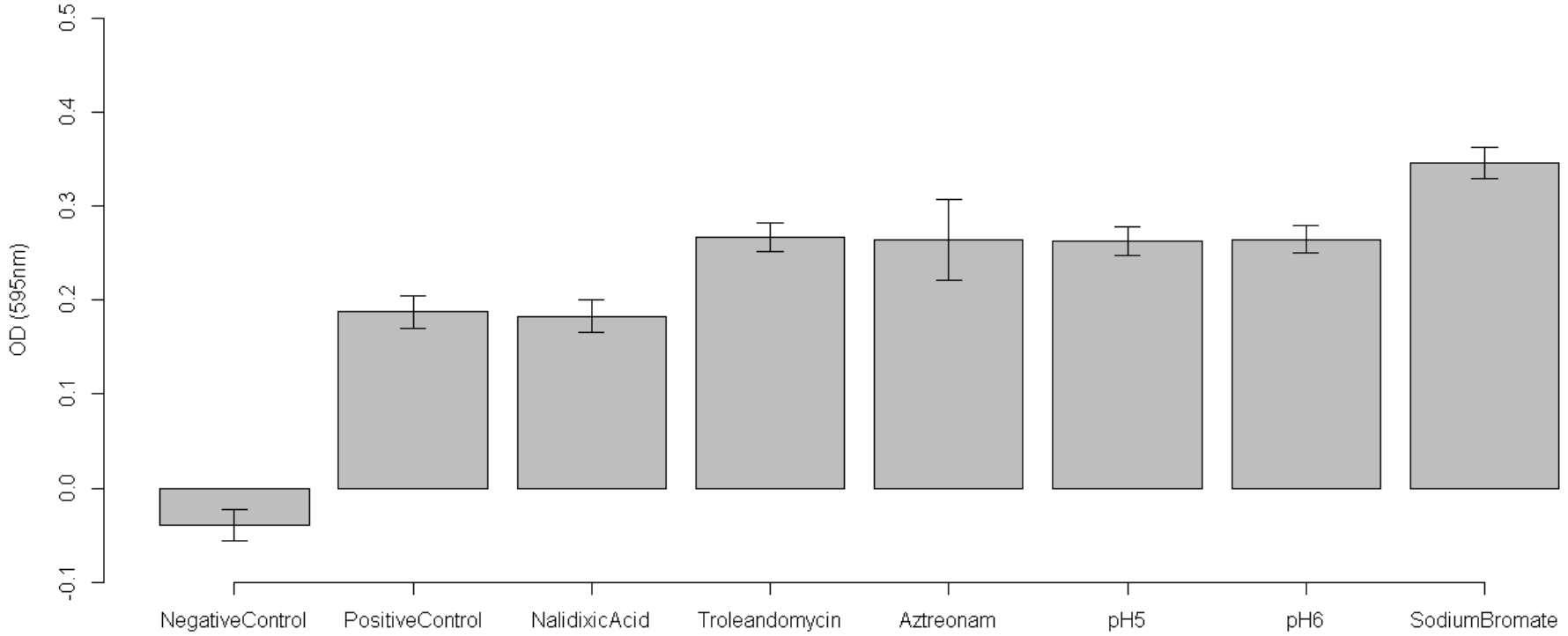


Figure 5.64 Biolog GenIII mean increase in OD_{600} (+/- s.e.)
The 'negative control' well showed a small fall in OD_{600} while the 'positive control' showed a rise comparable to nalidixic acid, slightly less than the antibiotic wells 'troleandomycin' and 'aztreonam' and the 'pH5' and 'pH6' wells. The best growth was in the ''Sodium Bromate well.

A maximal analysis of variance model of the increase in OD_{600} in R (version 2.15.1) with the explanatory variables: ploidy (haploid or diploid), well contents and strain, shows no significant effect of ploidy; see Table 5.13.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Ploidy	1	0.014	0.01377	1.467	0.22604
Ethanol	94	18.050	0.19203	20.464	< 2e-16 ***
Strain	8	0.154	0.01920	2.046	0.03852*
ploidy:ethanol	94	0.997	0.01061	1.131	0.19501
ethanol:strain	752	8.485	0.01128	1.202	0.00353 **
Residuals	970	9.102	0.00938		

Table 5.13 Maximal analysis of variance model for increase in OD_{600}

Removing the effect of ploidy leaves the significant explanatory variables largely unchanged, there is the expected strong effect of ethanol but also a weakly significant effect of strain (p=0.038) and a more significant interaction between ethanol and strain (p=0.0037), see Table 5.14.

.	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Contents	94	18.050	0.19203	20.464	< 2e-16 ***
Strain	9	0.167	0.01860	1.982	0.03835*
contents:strain	846	9.482	0.01121	1.194	0.00373 **
Residuals	970	9.102	0.00938		

Table 5.14 Minimal analysis of variance model for increase in OD_{600}.

5.3.10. Biolog GenIII plate genotyping – ancestor analysis

Plots of the two replicates for each strain (see Figure 5.65) suggested a guideline variation in OD_{600} of +/-0.15 above and below parity (indicated by the red and green lines on the plot).

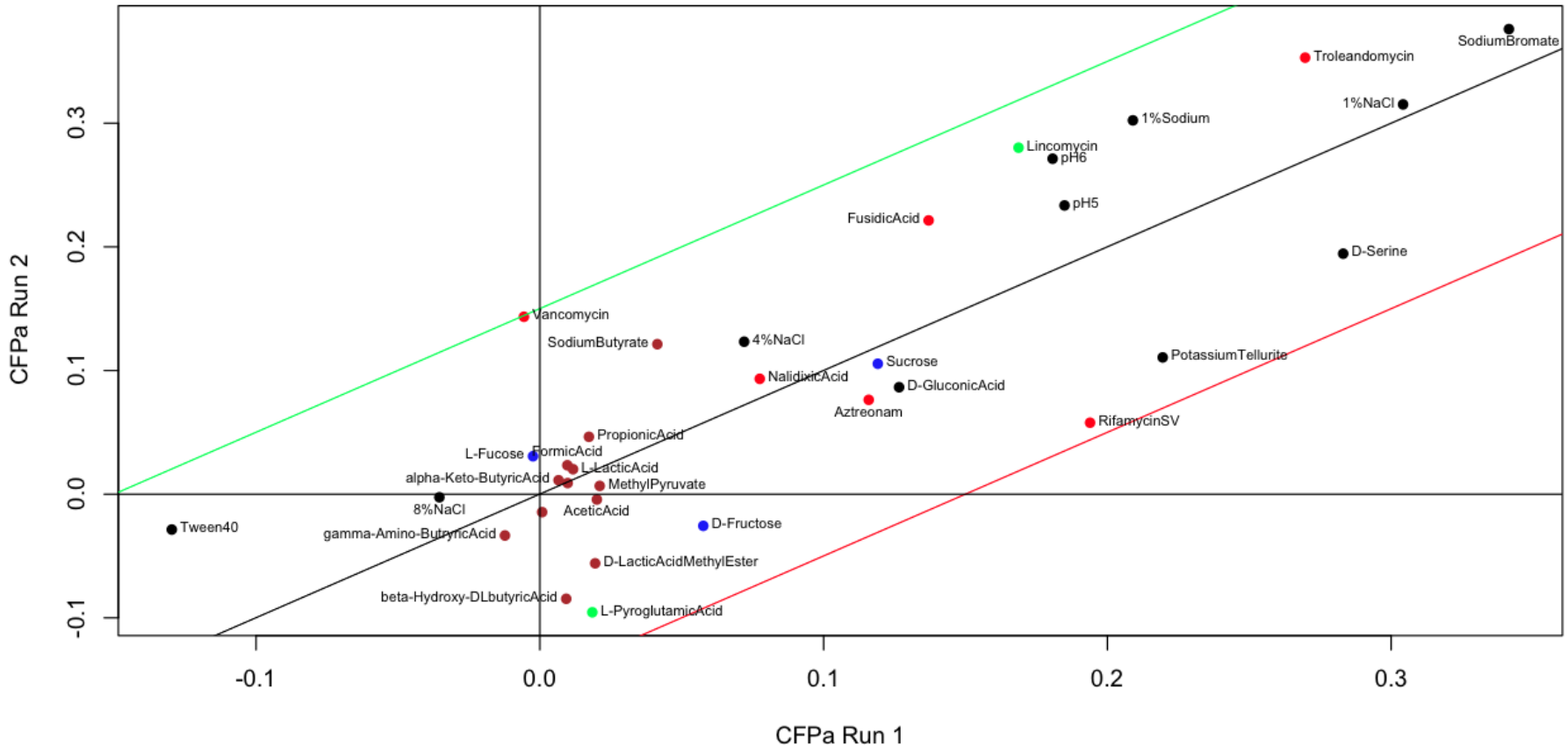


Figure 5.65 Growth of ancestral strain replicates as Biolog GenIII well contents differ.
The units on each axis are the change in OD_{600} after 24 hours of incubation. Wells that failed to grow in any strain have been omitted for clarity. Growth of each fluorescently tagged ancestor was very similar; the differences observed were minor and found in only one replicate (see Figure 5.66).

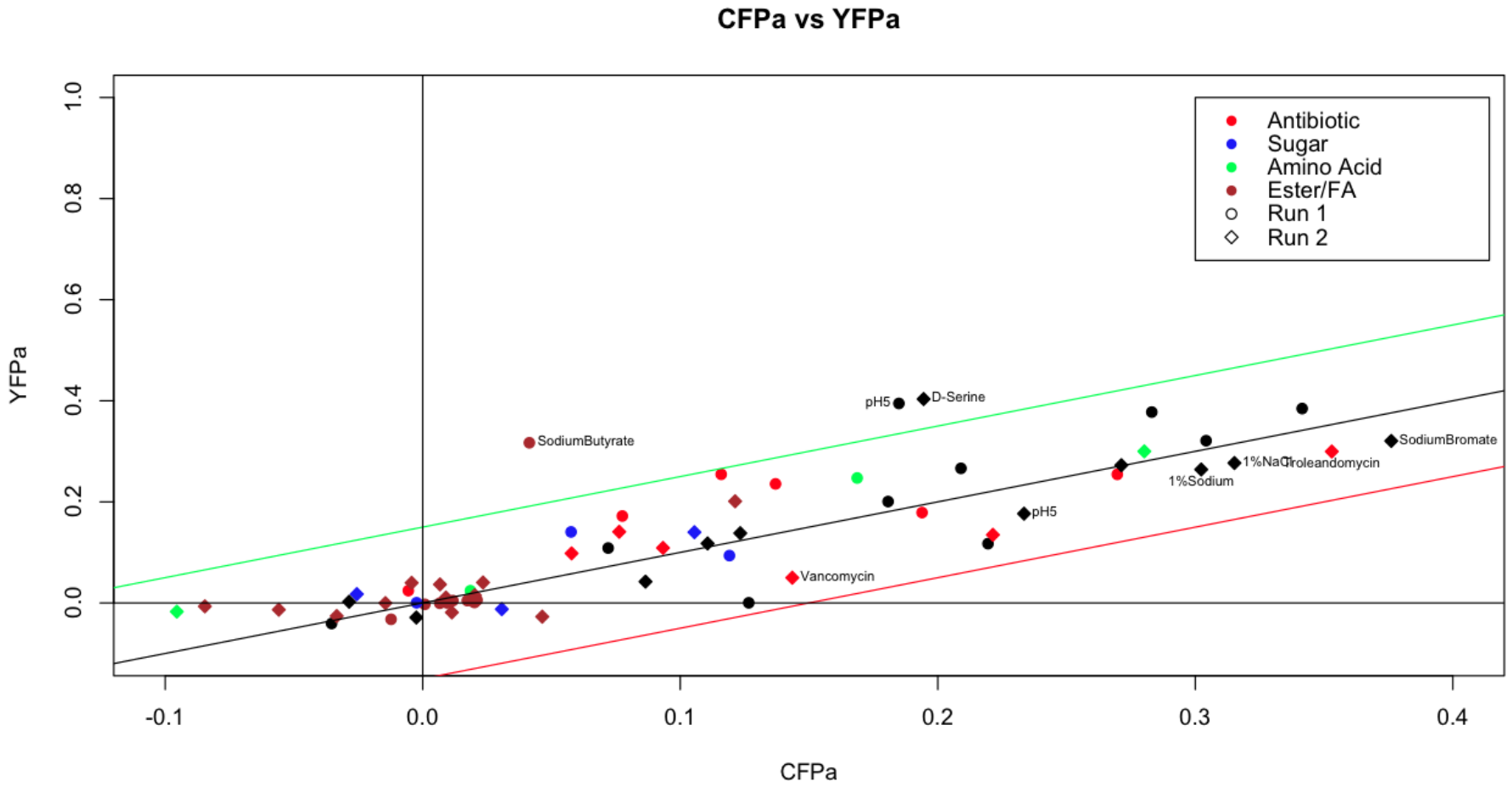


Figure 5.66 Growth of ancestral CFP and YFP tagged strains with differing Biolog GenIII well contents
The units on each axis are the change in OD_{600} after 24 hours of incubation.

5.3.11. Biolog GenIII plate genotyping – evolved strains

The growth of evolved strains was compared against the growth of their ancestor in each replicate. Five of the eight evolved strains evaluated (Evo6_11_D06_YFPe_4, Evo8_14_E07_CFPe_1, Evo8_14_D03_YFPe_1, Evo8_14_G04_YFPe and Evo10_20_G10_YFPe) showed very little difference in growth from their ancestor; any differences were small and inconsistent between the replicates. However, three evolved strains, Evo6_11_B06_CFPe_4, Evo8_14_F03_CFPe_2 and Evo10_20_F10_CFPe_2 showed different growth patterns from their ancestors.

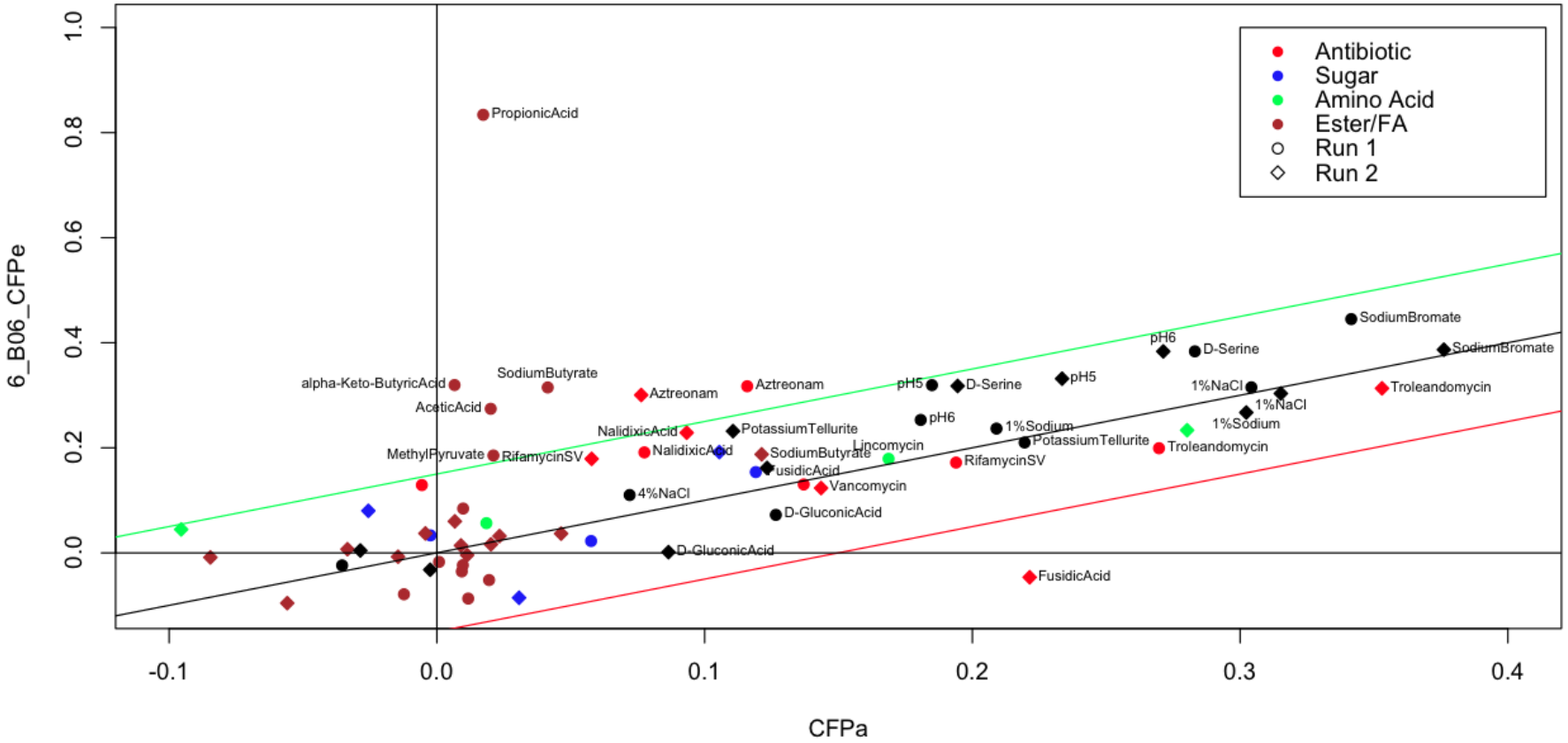


Figure 5.67 A comparison of the growth of Evo6_11_B06_CFPe_4 and its ancestor on Biolog Gen III plates.
The units on each axis are the change in OD_{600} after 24 hours of incubation.

Growth of Evo6_11_B06_CFPe_4 on wells containing acetic acid, propionic acid, alpha-keto-butyrate or sodium butyrate was better than the ancestor for the first replicate but not the second. Propionic acid growth for replicate 1 was particularly good.

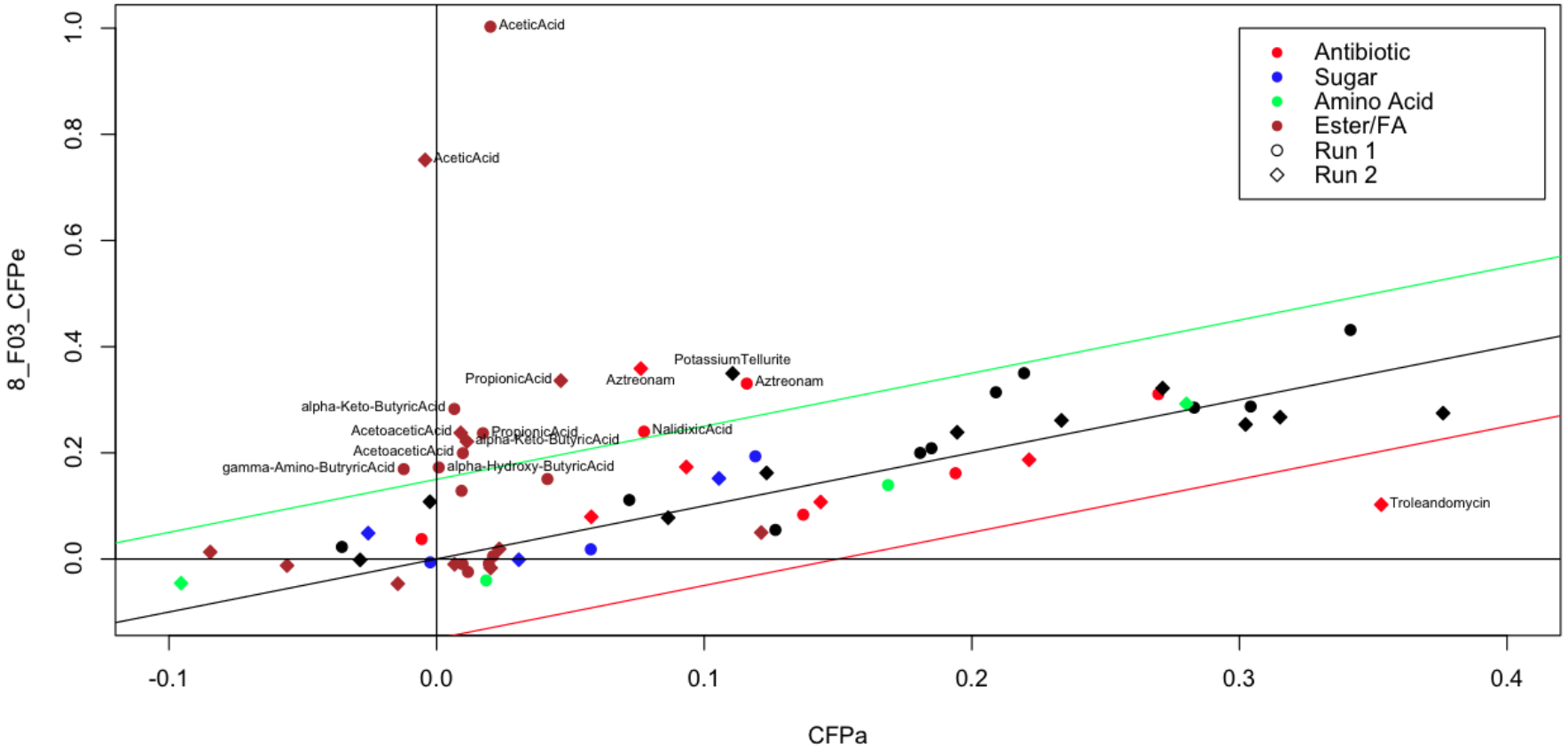


Figure 5.68 A comparison of the growth of Evo8_14_F03_CFPe_2 and its ancestor on Biolog Gen III plates.
The units on each axis are the change in OD_{600} after 24 hours of incubation.

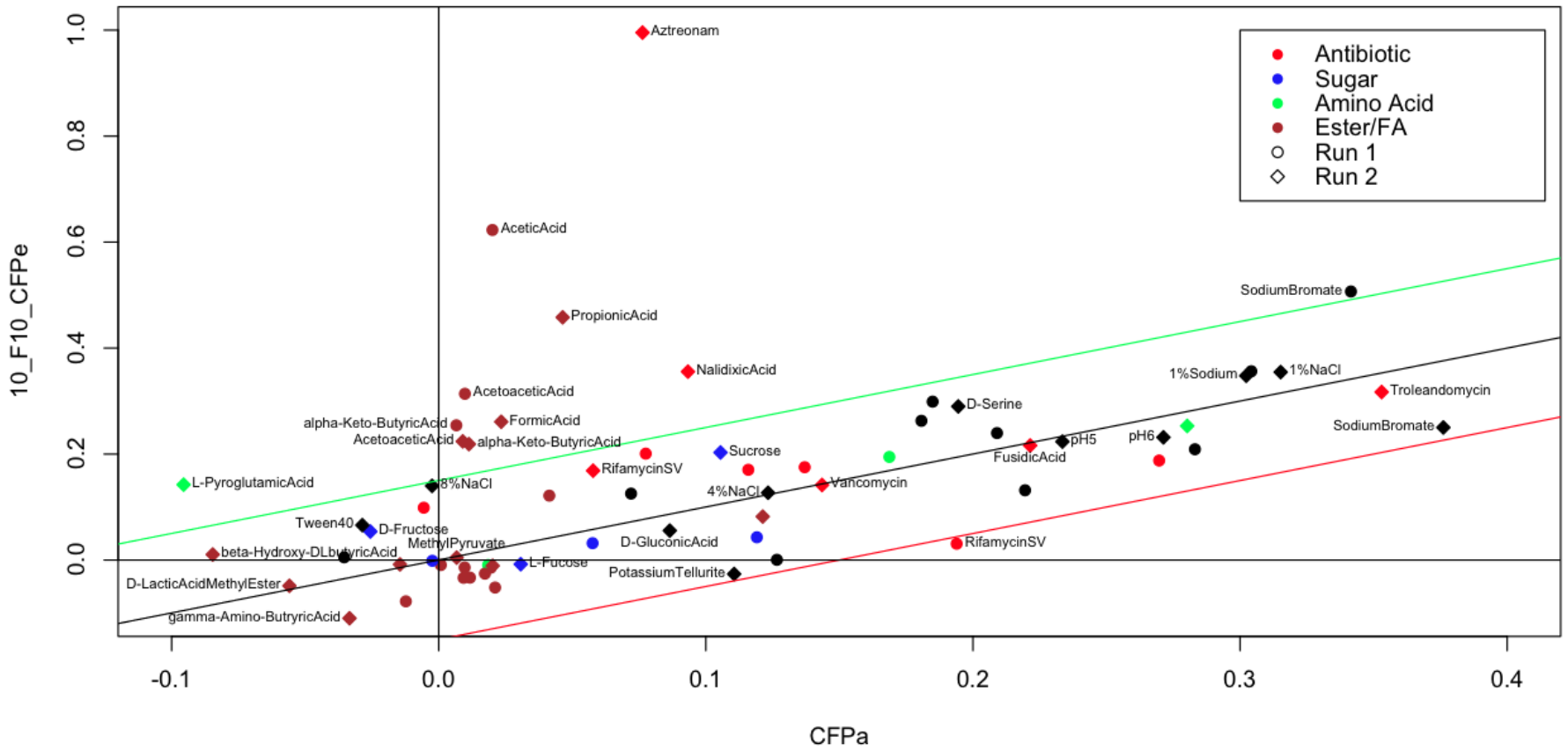


Figure 5.69 A comparison of the growth of Evo10_20_F10_CFPe_2 and its ancestor on Biolog Gen III plates.
The units on each axis are the change in OD_{600} after 24 hours of incubation.

This differential growth was on the organic acids: acetic acid; propionic acid, acetoacetic acid, and α-keto-butyric acid (see Figure 5.70, Figure 5.71, Figure 5.72, and Figure 5.73). Growth by the other strains is very poor in all four of these organic acids but the three evolved strains grew comparably to the positive control on propionic acid, acetoacetic acid and α-keto-butyric acid. On acetic acid Evo8_14_F03_CFPe_2, Evo10_20_F10_CFPe_2, grew much better than in any other well (see Figure 5.70). The first replicate of Evo6_11_B06_CFPe_4 grew better in organic acids but the second replicate grew comparatively poorly.

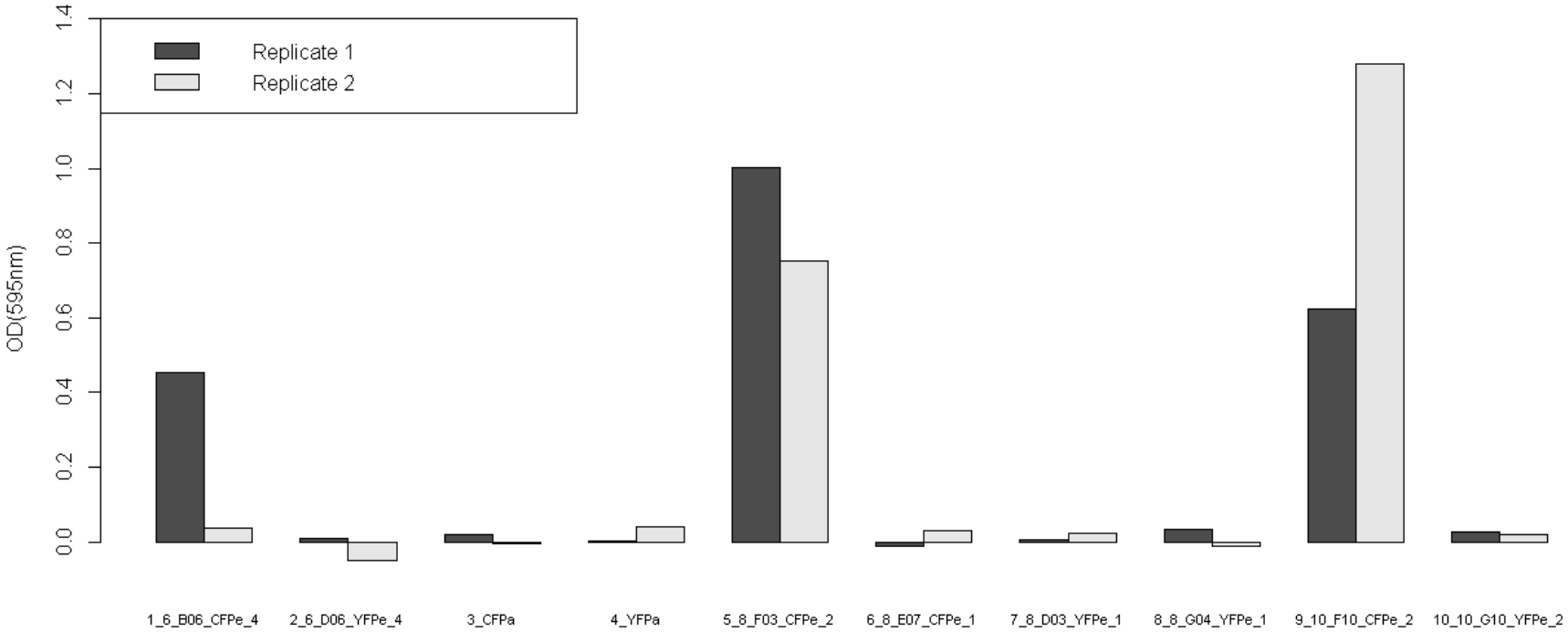


Figure 5.70 Growth by strain in the 'acetic acid' well of Biolog GenIII plates.
Bars show change in OD_{595}

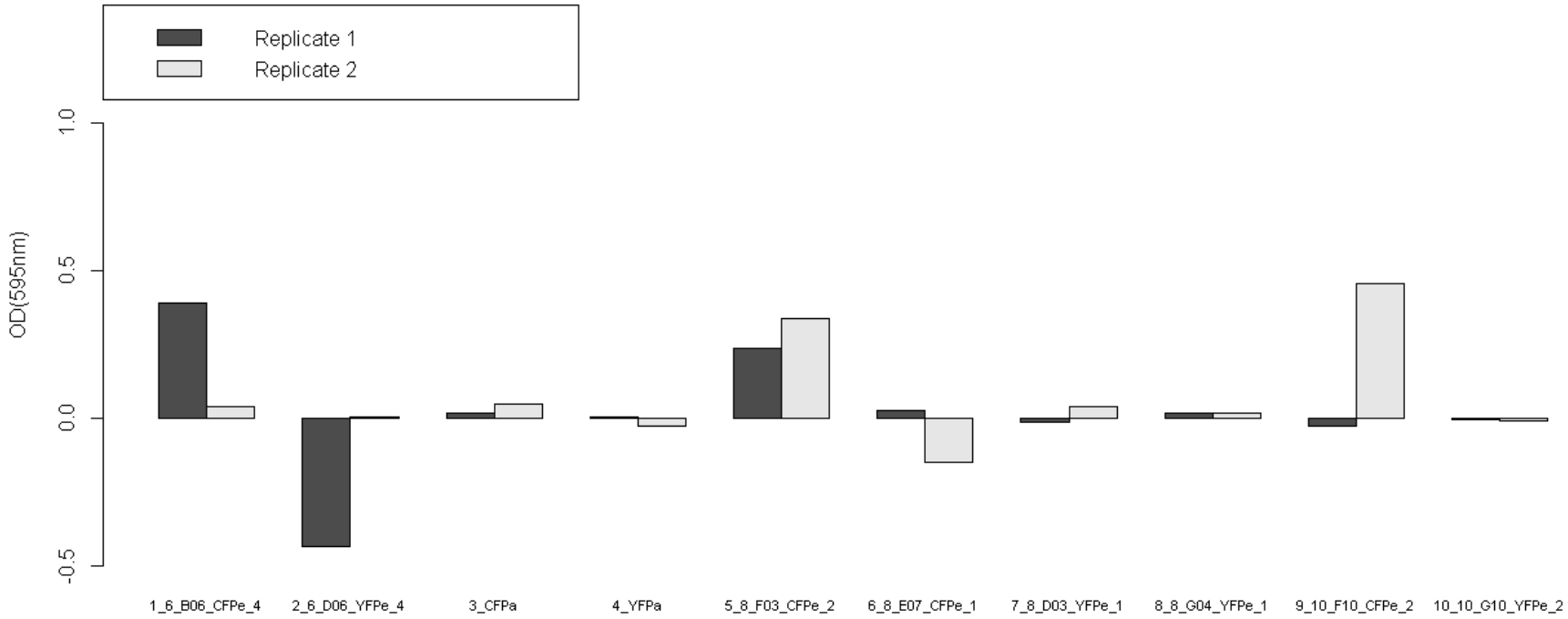


Figure 5.71 Growth by strain in the 'propionic acid' well of Biolog GenIII plates
Bars show change in OD_{595}

In the ‘propionic acid’ well the same three evolved CFP strains are the only ones to show significant growth (see Figure 5.71), but the addition of a single carbon to the organic acid reduces growth markedly. Although growth is consistent between replicates for Evo8_14_F03_CFPe_2, only one of the two replicates grows for Evo6_11_B06_CFPe_4, and Evo10_20_F10_CFPe_2.

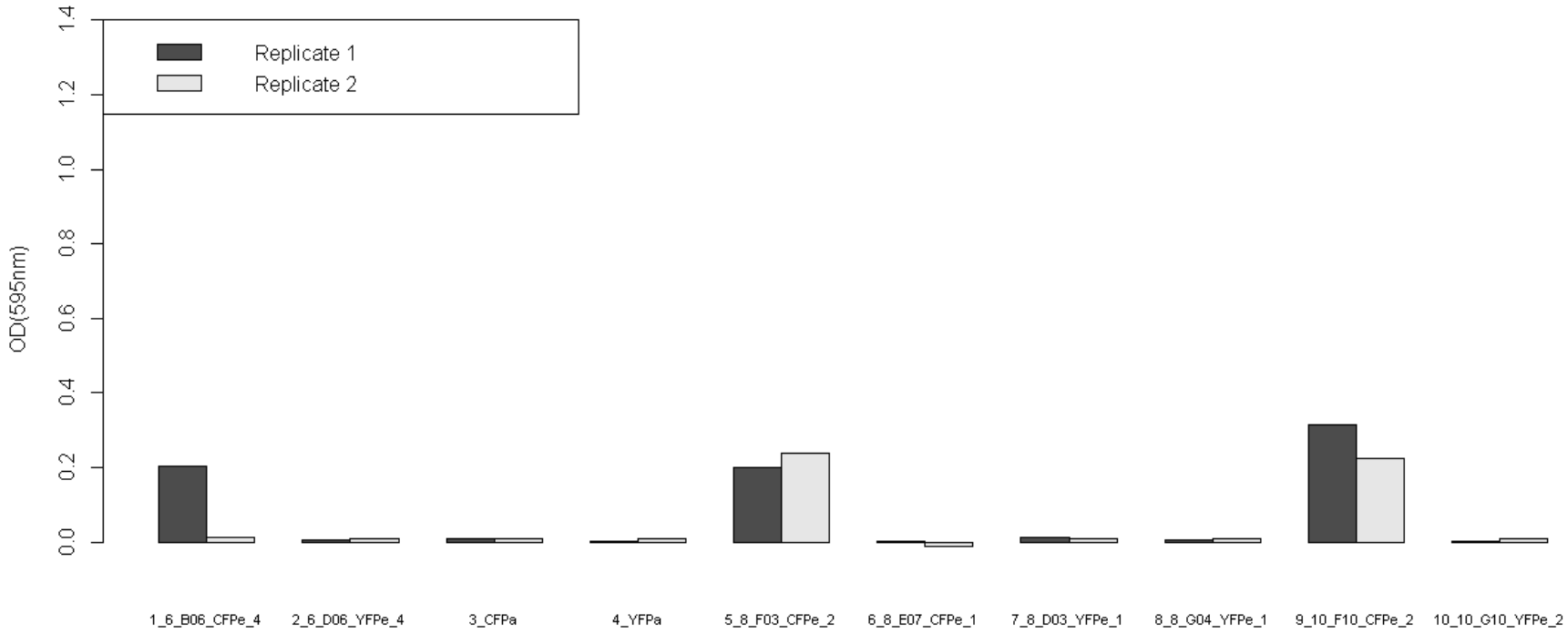


Figure 5.72 Growth by strain in the 'acetoacetic acid' well of Biolog GenIII plates
Bars show change in OD_{595}

As for the ‘acetic acid’ and ‘propionic acid’ wells, the same three evolved CFP strains (Evo6_11_B06_CFPe_4, Evo8_14_F03_CFPe_2, and Evo10_20_F10_CFPe_2) are the only ones to grow, although that growth is poorer still. Only one replicate of Evo6_11_B06_CFPe_4 shows measurable growth. Growth in the ‘α-keto-butyric acid’ well is very similar to that for ‘acetoacetic acid’.

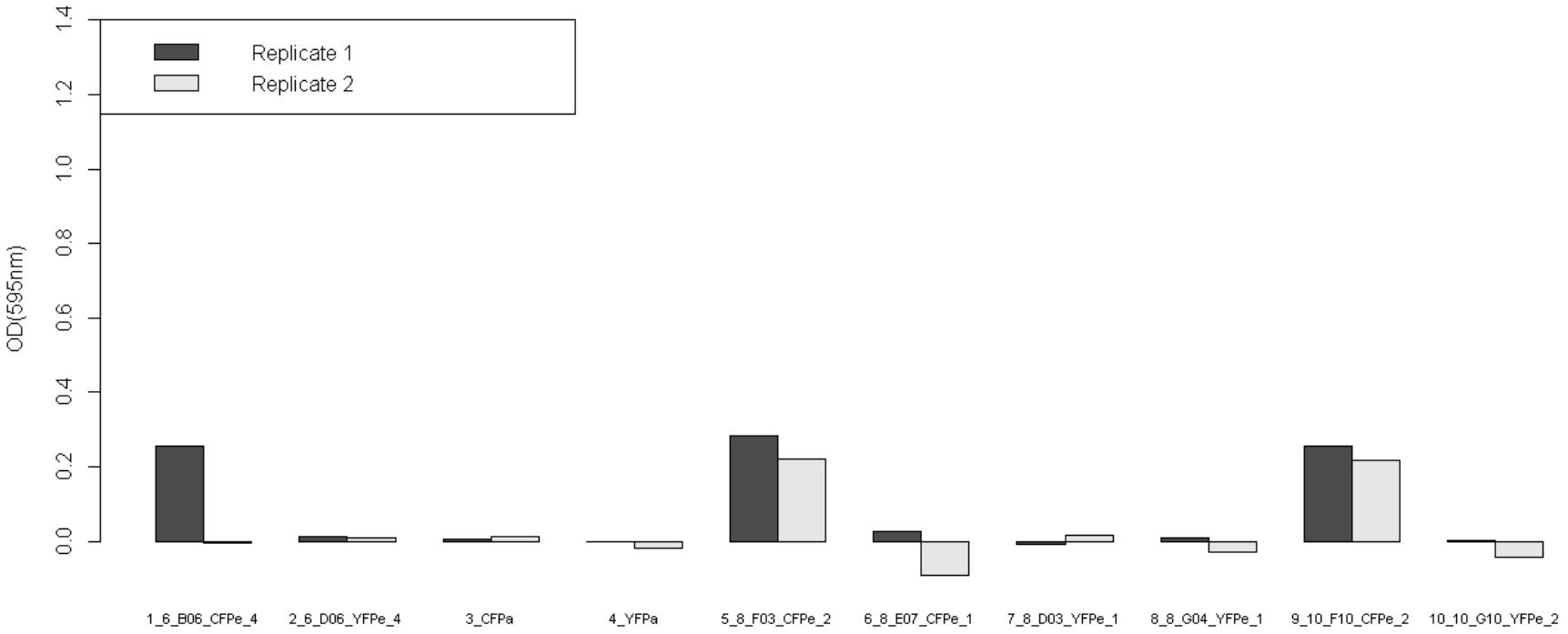


Figure 5.73 Growth by strain in the 'α-keto-butyric acid' well of Biolog GenIII plates
Bars show change in OD_{595}

5.4 Discussion

5.4.1. Summary

A research goal (see Section 2.7 “Research Aims and Goals”) was to phenotypically characterise those strains isolated from experimental evolutions identified as containing adaptively evolving populations using the methods discussed in Chapter 4. This chapter has demonstrated that strains were isolated that out-compete the ancestral strains. Some patterns of changed growth in competition emerged, such as the depression of the maximum growth rate of the ancestral competitors in evolutions 6 and 8. There is also a suggestion of antagonistic pleiotropy in evolution 10, at ethanol concentrations below approximately 3% the ancestral strains outcompete the experimentally evolved strains but the fitness advantage of those evolved strains increases with ethanol concentration up to the maximum value tested of 8%.

Those strains that were demonstrably able to out-compete their ancestors were all larger than those ancestors. Subsequent analysis of this phenotypic change in both isolated strains and evolving populations revealed a greater cellular DNA content consistent with an increase in ploidy; this is discussed further in in Chapter 6.

Some experimentally evolved CFP strains have gained a capacity to grow on organic acids that is absent in the ancestral CFP strain and any of the YFP strains tested.

The strains isolated from Evolution 10 were both larger than those from the other experiments and they were also the only ones to show a fitness effect that changed both magnitude and sign with ethanol concentration; they were either no fitter or were outcompeted by their ancestors without added ethanol but became progressively fitter than their ancestors at higher ethanol concentrations.

5.4.2. Strain growth analysis

The two ancestral fluorophore strains (DeLuna et al. 2008) were intended to be isogenic with identical patterns of growth prior to experimental evolution. The ancestral strains have very similar, but not identical, fitness in competition. Across the range of added ethanol from 0-8%, the ancestral YFP emerges from lag 2.5 hours earlier and enters diauxy 1.2 hours earlier than CFP but with similar fluorescence at 36 hours. Modelling work by Davey et al (Davey et al. 2012) suggests that strains can outcompete others over multiple rounds of batch growth

solely by having a shorter lag, something that might not be apparent in a single competition.

However, of the experimental evolutions only the 6.5% ethanol treatment in evolution 8 showed a drift in the in the fluorophore ratio, likely due to a beneficial mutation that arose in the overnight YFP culture. The other experimental evolutions showed initial plates with generally stable fluorophore ratios suggesting that there was no appreciable difference in initial fitness between the ancestral strains. Neither was there a significant difference in the number of strains that favoured one ancestral strain or the other (see Chapter 4, section 4.3). Accordingly it does not appear that the difference confers a substantial fitness benefit on the YFP strain.

As measured by the Nexcelom Cellometer, the ancestral CFP strain is approximately 0.25 microns larger than the YFP ancestor across the ethanol range from 0-8% (see Figure 5.51). Whole genome sequencing shows a 30bp CFP-specific deletion in the *EGT2* (YNL327W) gene which encodes a glucanase (see Chapter 6, section 6.4). The phenotypic effects of this deletion may differ, but deleting the whole EGT2 gene does not enlarge cells, it decreases competitive fitness (Breslow et al. 2008) which tends to reduce bud size with growth rate, but delays bud separation, during which buds continue to grow (Kovacech et al. 1996). It is possible that such adherent buds may be measured as larger single cells by the Nexcelom Cellometer software particularly if the cells are oriented to obscure the bud neck. However, a review of the Nexcelom Cellometer images of both the ancestral and evolved strains used in these experiments failed to reveal the clusters of four cells typical of the *EGT2* deletant phenotype (see Figure 5.74). Some chains of three or four cells were observed, similar to that typical of a wild-type haploid at the bottom left of Figure 5.74.a), but these were relatively rare even in populations that had not been sonicated, and occurred with both CFP and YFP strains.

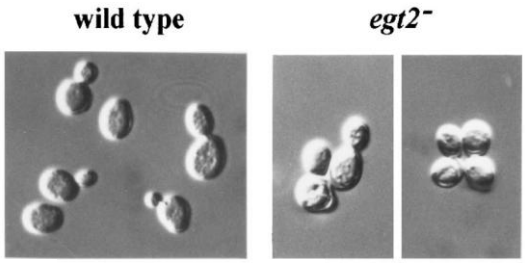


Figure 5.74 Micrographs of: a) haploid, and b) *EGT2* deletant *S. cerevisiae* strains from Kovacech et al, figure 2 (Kovacech et al. 1996).
According to Kovacech et al. the cells in this figure were from exponentially growing cultures and were 'mildly' sonicated. Three of the haploid wild-type cells in a) are in pairs, mother cells with adherent buds and towards the bottom left of the image there are a chain of three cells. The EGT2 deletant cells in b) are in groups of four adherent cells.

Although the evolved strains show diverse patterns of duration and rate of growth, there are strong similarities in each evolution series. Both strains in evolution 6 emerge from lag earlier than their competitor and grow for longer, particularly Evo6_11_D06_YFPe_4, unlike the strains in evolution 8 and 10 which grow for a duration similar to their competitors.

In evolutions 6 and 8 there are clear increases in the maximum rate of growth of the evolved strains over that of the ancestral strain in direct competition but the maximum growth rates seen are lower than those observed when the ancestral strains compete. It appears that, irrespective of ethanol concentration, the haploid ancestors have a higher maximum rate of growth but they are unable to realise their potentially greater growth rate when competing with evolved diploids in these experiments. The evolved diploid strains have consistently invaded the haploid starting strains. Hence it is possible to speculate that there may be a trade-off between better overall growth throughout every phase, from lag to slow growth, in the evolved strains against a higher rate of instantaneous growth in the ancestors. The smaller surface area to volume ratio of diploids could limit their maximum growth rate whilst improving their ethanol tolerance and hence their overall growth.

The experimentally evolved strains in evolution 6 were exposed to an added ethanol stress of 4.5%. The evolution 6 strains characterised (Evo6_11_B06_CFPe_4, Evo6_11_D06_YFPe_4) have two advantageous competitive characteristics; they both have a shorter lag (their growth starting ~4.5 hours before their ancestral competitors) and they depress the maximum ancestral competitor growth rate. Those evolution 8 strains that were also grown with 4.5% added ethanol (Evo8_14_F03_CFPe_2, Evo8_14_C06_CFPe_2) also depress the maximum growth rate of their competitors but do not show an earlier start to growth. This

implies that strains evolved in 4.5% ethanol can access genetic changes showing a range of competitive phenotypes. All the evolved strains have become (apparently euploid) diploids (see Section 6.3.8) but these additional differences in evolution 6 strains suggest they may have different genetic sequence changes from the 4.5% evolution 8 strains. Despite the relatively short duration of these evolutions (less than 150 generations for these experiments) variation in competing phenotypes may be due to more than one genetic change in some strains. For example, depression of the competitors' maximum growth may be an effect of ploidy whilst there may be additional genetic changes that cause a faster start to growth in the evolution 6 evolved strains. However, apart from a synonymous, heterozygous SNP in the SSL2 gene of Evo6_11_B06_CFPe_4, no other sequence changes were detected in evolution 6 (see Section 6.3.4).

Both the 4.5% stress and 6.5% stress have produced experimentally evolved strains that are diploid. However, it is not possible firmly to draw further conclusions about whether the adaptive response of this experimental system with a 6.5% ethanol stress differs greatly from a 4.5% stress. When the Dirac equation is fitted to the fluorophore ratio changes of the 6.5% treatments (See Fig. 4.26) only four wells (D03, D04, G04, G05) out of the 16 in the 6.5% ethanol treatment had estimates for the tau (onset) parameter substantially greater than zero (although these were less than 5) and most estimates were negative, albeit with very large error values. Additionally both of the strains sequenced from this treatment (in addition to diploidy) had identical mutations in the fourth intron of the mitochondrial gene *COB1* (see Section 6.3.5). This strongly indicated that an adaptive change swept through the overnight culture (YFP) population before the beginning of the experiment. It is not clear whether similar adaptive change would arise and sweep through the large population in overnight cultures with 6.5% added ethanol if this was repeated or whether this was an exceptional event. Regardless there appear not to be 16 independent trials with 6.5% ethanol to draw conclusions from.

Populations may adapt in experimental evolutions not just to the overt stresses considered by the experimenter but also to unrecognised problematic characteristics of the experimental system e.g. the reactive oxygen load of illuminating fluorophores, evaporation of media causing osmotic stress etc. Similarly adaptive changes may enhance fitness by ameliorating the deleterious effects of genetic lesions such as multiple auxotrophies rather than responding directly to an applied stress. Conversely however, even if it is not added to the starting medium, ethanol is added to these experimental evolutions by fermentation. Accordingly it is not possible to

conclude with certainty that the adaptive changes seen in the control wells without ethanol in the starting media are not due to ethanol stress. In order to demonstrate with confidence that adaptive changes confer enhanced ethanol tolerance rather than being solely of benefit in a particular experimental system, it would be necessary to test that experimentally evolved strains grew better with ethanol stresses in different growth conditions e.g. by growing better on ethanol stressed solid as well as liquid media. Adaptive changes that solely confer a benefit in the conditions in which they evolved but not in other ethanol stressed environments may have arisen to confer fitness benefits in particular response to environmental stresses other than ethanol.

Greater relative rate of growth appears to be the principal determinant of increased fluorescence in the evolved strains. However, the diploid strains in evolution 10 appear do not grow better over the whole range of ethanol stresses; Evo10_20_F10_CFPe_2 grew no better without added ethanol stress and Evo10_20_G10_YFPe_2 grew substantially worse. The fitness increment of the evolved strain Evo10_20_F10_CFPe_2 was modest at 8% but Evo10_20_G10_YFPe_2 grows substantially better than its ancestor or competitor under this stress.

In evolution 10 the two strains phenotypically characterised appear to have made adaptive change(s) which show antagonistic pleiotropy; the evolved strains are less fit than their ancestors without added ethanol but they both out-compete the ancestors with ethanol stress, their fitness benefit increases with ethanol stress up to 8%, the maximum concentration tested. None of the phenotypically characterised strains evolved with constant ethanol stresses showed such a pattern of fitness trade-off between low and high ethanol. Both the maximum growth rate and duration of growth prior to diauxy improve in the evolution 10 strains, with the lag for Evo10_20_F10_CFPe_2 being particularly poorer than the ancestor without ethanol added to the competition.

With only a small number of strains characterised, it may solely be chance that the evolution 10 strains differ from those isolated in other experiments. However, the evolution 10 culture conditions may also have elicited a different response from the competing populations. The ethanol stress was both ramped (increasing at 0.5% per plate) and ultimately rising to higher levels than the other experiments (9.5%) over a larger number of generations. Higher added ethanol stresses (which greatly inhibit the growth of ancestral strains) may allow mutations to sweep that have bigger

beneficial fitness effects under intense ethanol stress. Such large effect mutations may be more likely to be deleterious in the absence of ethanol stress and at lower ethanol concentrations the high stress benefits may never outweigh the deleterious effects in the absence of ethanol. Having a ramped stress with low ethanol selection pressure for a larger number of generations may also have allowed greater genetic diversity to develop. Evo10_20_F10_CFPe_2 also showed strong growth in organic acids, particularly acetic acid (see section 5.3.11).

The diploid evolution 10 strains were significantly larger than the other evolved strains (see Section 5.3.6), with the exception of the triploid strain Evo10_20_B03_CFPe_5 which was also isolated from evolution 10. All the evolved strains that out-competed the ancestors were larger than the haploid starting strains, and the evolution 10 strains were largest of all. A general characteristic of the evolved diploids is that although they have a lower maximum growth rate than their ancestor, they depress the maximum rate of growth of a haploid competitor much further. It is conceivable that larger cells are better able to maintain an active transport gradient under ethanol stress (see Section 2.1.3) allowing them to absorb nutrients more efficiently than the smaller haploids although these large cells cannot increase their growth rate as rapidly. It is further possible to speculate that the larger diploids from evolution 10 (Evo10_20_F10_CFPe_2, Evo10_20_G10_YFPe_2) are so large that their growth is actually poorer than their competing haploids without ethanol stress. However, this cannot explain all observations, the triploid strain Evo10_20_B03_CFPe_5 is larger than the evolution 10 diploids and shows both a faster rate of growth and greater overall growth in all ethanol concentrations. It is possible that this (unsequenced) strain has genetic changes in addition to ploidy which confer a fitness benefit.

The fact that there are some strains which show different durations of growth and others which display antagonistic pleiotropy suggests that there is more than one route to enhanced fitness in this system. It also strongly suggests that diploidy alone does not confer a common pattern of fitness increase and that sequence changes modulate the effect of ploidy. Accordingly the high fitness of the likely triploid Evo10_20_B03_CFPe_5 at all ethanol levels may be due to increased ploidy or changed sequence but is likely a combination of both.

A low level of ethanol stress can be used to evolve strains with tolerance at higher levels; Evo6_11_D06_YFPe_4 evolved in 4.5% ethanol and had high fitness at 8%.

Conversely strains isolated using a stress ramped to 9% (Evo10_20_F10_CFPe_2, Evo10_20_G10_YFPe_2) grew poorly without added ethanol.

The bulk of the population growth occurs during log phase fermentation. The protocols used for evolution (and strain evaluation) in this experiment have a limited period (less than a day) before fresh nutrition is supplied. This limited period in slow-growth phase is insufficient for the population to complete the transition to the stationary phase (approximately a week depending on strain) and the experimental conditions may favour remaining in a fermentative rather than respiratory state and continuing to grow (albeit at a low rate) rather than entering quiescence.

Despite suggestions that the ethanol concentrations of media are stable between 0-8% in aerobic culture (Walker-Caprioglio et al. 1985), a concern of evaluating strains using small volumes of culture in lidded, but unsealed, 96 well plates is that evaporative loss of ethanol may lead to an over-estimation of ethanol stress. Yeast ethanol tolerance with standard media in sealed plates may be impaired by lack of ergosterol synthesis in anaerobic conditions (Andreasen & Stier 1953). The linear increase in lag with added ethanol in all the strain evaluations suggests that evaporative loss does not wholly vitiate the imposed stress, at least at the onset of fermentation. However, further work to assess the stability of ethanol concentrations in 96 well plates would be desirable.

The times of onset of detectable growth (15 hours plus) are substantially longer than published values. For example (Medawar et al. 2003) cite lag values without ethanol of 20 to 40 minutes. However, these values were obtained with an inoculum that is 10% of a previous stationary culture and these experiments were carried out with ~10^4 cells. Accordingly approximately four to five rounds of replication are required before detectable increases in fluorescence or optical density are apparent. There is no apparent discontinuity in the lag time with ethanol – i.e. no additional time required for an extra tolerance step over a certain threshold.

5.4.3. Biolog GenIII plate phenotyping

The good general growth of the yeast strains in the ‘ph5’ and ‘ph6’ wells is consistent with known characteristics of *S. cerevisiae* (see Chapter 2, section 2.3). The wells containing antibiotics ineffective against yeast may have other contents (such as an acidic pH buffer) that make them effective yeast growth media. Aztreonam is a beta-lactam antibiotic effective against gram negative bacteria and troleandomycin is solely used against bacterial infections. There are reports that nalidixic acid (0.5

mg/ml) induces a transient cessation of DNA, RNA and protein synthesis in *S. cerevisiae* (Sobieski & Brewer 1976) and promotes petite formation (Gross & Smith 1972) which would suggest that the antibiotic is moderately impairing growth while an acidic buffer, the antibiotic's own acidity, or other components of the growth medium in this well might be promoting it. The potassium bromate salt (E924) has been widely used as a flour improver, thought to work by breaking disulphide bonds in the dough protein, gluten (Wilderjans et al. 2010).

There is no recognisable increment or changed pattern in growth common to all the evolved diploids (the 'triploid' strain Evo10_20_B03_CFPe_5 was isolated after this work was complete). Five of the eight evolved diploid strains did not grow discernibly differently from their haploid ancestors on Biolog GenIII plates; diploidy alone appears insufficient to generate different growth profiles. However, the evolved strains Evo6_11_B06_CFPe_4, Evo8_14_F03_CFPe_2, Evo10_20_F10_CFPe_2 show good growth in wells containing organic acids as a carbon source in which all the other strains grow negligibly despite the good general tolerance of *S. cerevisiae* to acid growth conditions (see section 2.3). This suggests that these three strains have sequence changes not present in the other five that could: change the pH tolerance of the strains, reduce organic ion toxicity, or enable the metabolism of the organic acids.

It is not statistically significant that the three strains with good organic acid tolerance all have the CFP fluorophore rather than YFP. However, there is a deletion in the EGT2 gene found exclusively in the CFP ancestral strain (see Chapter 6, section 6.3.3) and a null mutation in this gene increases tolerance of acetic, lactic and hydrochloric acids (Kawahata et al. 2006). The ancestral genotype cannot wholly explain the pattern of organic acid growth; the ancestral CFP strain does not grow in the Biolog GenIII organic acid wells and neither does the CFP diploid Evo8_14_E07_CFPe_4.

Despite their good growth on organic acids, no changes from the starting genotype in genetic sequence, copy number or structural genetic rearrangements were found in the Evo6_11_B06_CFPe_4 or Evo10_20_F10_CFPe_2 strains; see Chapter 6, section 6.4, for further discussion of this discrepancy. Strains with changed patterns of growth on organic acids arose in all three evolution series suggesting that this phenotype is a common, but not obligatory, corollary of adaptation in these experiments.

Organic acid food preservatives generally do not kill microorganisms but greatly extend the growth lag (Fernandes et al. 2005). Accordingly, extending the incubation time of the Biolog GenIII plates might have enabled the other strains to grow in the organic acid wells.

6. Analysis of genotype including ploidy change and its phenotypic effects in evolved strains

6.1 Introduction

A research goal (see Section 2.7 “Research Aims and Goals”) was to genetically characterise experimentally evolved strains. Next generation, whole genome (Illumina) sequencing was undertaken to look for genetic sequence changes. Repeating patterns of change in parallel experimental evolutions were sought that might be causing the increases in fitness observed (see Section 2.4.7 “Drawing conclusions from experimental evolution”). The strains chosen had been isolated from populations that exhibited fluorophore ratio change indicative of a change in fitness (see Chapter 4) and those strains had demonstrated that they could out-compete their ancestors not just in the conditions in which they evolved but also over a range of ethanol concentrations(see Chapter 5). Some strains also displayed the novel phenotype of organic acid tolerance (see Section 5.4.3).

Illumina next generation genome sequencing can find not just genetic sequence changes such as single nucleotide polymorphisms (SNPs) and insertions or deletions (indels) of varying sizes, but can also identify some variations in copy number and structural changes such as translocations and inversions of portions of chromosomes. The falling cost of, and increase in data generated by, sequencing enables the genetic analysis of many *S. cerevisiae* strains in parallel, potentially enabling repeating patterns of adaptive evolution to be identified.

Cell enlargement was a feature of all experimentally evolved strains that out-competed their ancestors (see chapter 5). This can be caused by changes in the sequence of individual genes, but can also be caused by the global increment in copy number that accompanies increasing ploidy. Ploidy was assessed by fluorescence cytometry (a procedure commonly referred to as FACS) of fixed cells stains with DNA binding dye and that was correlated with cell size; it was important to determine whether changes in cellular DNA content accompany changes in cell size. In addition to the characterisation of isolated strains, these ploidy assessments were carried out on populations at different stages of an experimental evolution to look at the speed and extent of any ploidy changes. Ploidy changes, particularly increases in ploidy, have been noted in response to other stresses (see Section 2.2.4).

6.2 Methods

6.2.1. DNA extraction

Whole genome sequencing required >10µg of DNA at a concentration of 50ng/µl for each sample. The Promega Wizard Genomic DNA purification kit (Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711 USA) was used with some changes to the protocol necessitated by the DNA quantity required. A 100ml of a dense overnight YPD culture was pelleted by centrifugation at 13,000g for 2 minutes and the supernatant discarded. Residual liquid in the pellet was sufficient to produce a dense slurry when stirred that could be frozen and ground in a mortar and pestle using liquid nitrogen. This ground residue was washed from the mortar with 1.2ml of 50mM EDTA, resuspended by vortexing, and split into four equal aliquots in 1.6ml micro centrifuge tubes. Initial extractions were mixed with 7.5µl of 20mg/ml lyticase and incubated for an hour at 37°C but this proved not to contribute materially to the yield and the step was later omitted. The suspension was spun at 13,000g for 2 minutes, and the supernatant discarded. Thereafter the protocol of the Promega kit was followed; 300µl of Promega 'Nuclei Lysis' solution and 100µl of 'Protein Precipitation' solution was added to each aliquot before incubation on ice for 5 minutes. Following centrifugation at 13,000g for 3 minutes, the supernatant was transferred to fresh micro-centrifuge tubes containing 300µl of room temperature isopropanol and gently mixed by inversion. After centrifugation at 13,000g for 2 minutes the supernatant was discarded and 300µl of 70% ethanol at room temperature added. After centrifugation at 13,000g for 2 minutes the ethanol was aspirated and the pellet air-dried. Promega 'DNA Rehydration' solution was added and each aliquot was incubated at 37°C for 15 minutes with Roche DNase-free RNase (Roche Diagnostics Ltd., Charles Avenue, Burgess Hill, West Sussex, RH15 9RY) before rehydration overnight at 4°C and a final 2 minute spin at 13,000g before combining the micro-centrifuge supernatants.

DNA content and integrity was assayed by running samples on 0.8% agarose gels and by Qubit fluorimetry (Life Technologies Ltd, 3 Fountain Drive, Inchinnan Business Park, Paisley PA4 9RF).

6.2.2. DNA Sequencing and Analysis Pipeline

Illumina HiSeq 2000 paired end sequencing was carried out by GenePool (Ashworth Laboratories, The King's Buildings, University of Edinburgh, Edinburgh, EH9 3JT, Scotland) who returned fastQ format files containing sequence and quality data.

Initial quality control of the fastQ format raw reads was carried out using FastQC (Babraham Bioinformatics, http://www.bioinformatics.babraham.ac.uk/projects/fastqc/); a suite of read quality tools that can identify over-represented sequences (such as library primers), and plot disparities between complementary bases and declining quality against read position. The read trimmer TrimmomaticPE (Lohse et al. 2012) was used to cut the leading 5 bases from the 5' end of each read as well as performing a 3' crop to a length where base quality falls below a Phred score of 20 in a sliding window of four bases. The Phred score (so called because it was originally produced by a program of that name), a statistical indicator of read quality

$$Q = -10*\log_{10}P$$

Equation 6.1 The relationship between Q, the Phred quality score, and P the base calling error probability.

The Phred score is determined from experimental DNA sequence traces using software calibrated with known sequence data having sequence traces of varying clarity of peak shape and resolution. A Phred score of 20 indicates a probability of an incorrect base call of 1 in a 100.

The (BWA) Burrows-Wheeler Aligner (Li & Durbin 2010) was used for paired-end mapping to the R64-1-1/SaCer3 S288c reference (released 3/2/2011) from the Saccharomyces Genome Database (http://www.yeastgenome.org/). SAMTools (H. Li et al. 2009) were used to convert the mapped sequence alignment/map (SAM) text file to the compressed binary alignment/map (BAM) format and to sort and index the BAM file. SAMTools was used to remove duplicates of sequences with identical mapping locations; such duplication is likely to arise in PCR. The Broad (GATK) Genome Analysis Toolkit (McKenna et al. 2010) was used to re-align and re-index the BAM files to minimise mismatches in reads due to short insertions or deletions. The CollectInsertSizeMetrics tool in Picard (http://picard.sourceforge.net/) was used to collect the insert size metrics of the paired end reads and the SAMTools view produced mapping statistics.

6.2.3. Identifying changes to DNA sequence and copy number

The reliable identification of single nucleotide variation requires the evaluation of a series of DNA sequence metrics. Firstly sequence reads must be of good quality (see Section 6.2.2. above), secondly the reads must be aligned to the reference with

a good mapping quality score. This mapping score estimates the probability of a particular read being wrongly aligned to the reference derived from a sum of the base quality scores of mismatches at the best alignment and at any other competing alignments (Li et al. 2008); the more good alternative alignments there are the smaller the map quality score will be. Thirdly a variant caller must produce an estimate of how good a predicted variation is. Mpileup in SAMTools was used for identifying small nuclear polymorphisms (SNPs) and small insertions or deletions (indels which in this context is a group of SNPs). SAMTools produces an indicator of confidence in the called variation based on the mapping quality of the reads and the read depth; the larger the value of this "SNPQ" quality the more confident SAMTools is that there is a variant. SAMTools classifies these variant calls into "Strong" (SNPQ greater than 100) "Medium" (SNPQ greater than 20 and less than 100) or "Weak" (SNPQ less than 20). Park et al. evaluated the accuracy of variant calls of differing quality by SAMTools using Sanger sequencing; they found that the positive validation of the variant was positively correlated with SNPQ but showed no correlation with total read depth (M. Park et al. 2014). Variants with a SAMTools SNPQ value of 80 or greater had a validation rate using Sanger sequencing in excess of 80% and they consider this a good cutoff for a reliable variant prediction by SAMTools (Park et al. 2014). This is just below the SNPQ cutoff of 100 for a SAMTools "strong" prediction.

Pindel was used to search for larger indels (Ye et al. 2009). Sequence differences between starting and derived strains and between starting strains and the S288c reference were noted. The (IGV) Integrative Genomics Viewer (Robinson et al. 2011) was used for initial visualisation of sequence changes since it allows a sequence to be viewed in parallel with that of both its ancestor and other derived strains. The UCSC Genome Browser (Kent et al. 2002) was used for its comprehensive annotation. CNVseq (Xie & Tammi 2009) was used to look for signals of copy number variation during experimental evolution by comparing the mapped sequences of a derived strain with its ancestor.

BreakDancer (Chen et al. 2009) was used to check mapped reads for discontinuities, read pairs that map to different locations in the reference or with changed orientation that suggest structural rearrangements in the genome have occurred, large deletions inversions or translocations. Candidate structural variations were visualised in the IGV viewer, which highlights reads with read pairs that do not map to proximal locations.

6.2.4. Cell DNA content using flow cytometry

A stationary phase culture of yeast cells was diluted x50 into a 1.5ml microcentrifuge tube and 20µl used for cell count and size analysis using the Nexcelom Cellometer in the usual manner (see Chapter 3, section 3.2). Density plots of the cell sizes were made using default settings in the "geom_density" function of the ggplot2 library (version 0.9.3.1 in R version 2.15.3).

The cell counts were used to calculate the undiluted volume (typically ~50 µl) of cell suspension containing $I0^7$ cells. This volume of culture was added to 1.5ml of water in a 15ml screw-cap centrifuge tube. 3.5ml of 95% ethanol was slowly added, followed by gentle mixing by inversion. The cells were left to fix overnight to fix at 4°C.

The fixed cells were centrifuged (15s at 10,000g) and the supernatant discarded. The cell pellet was washed by resuspension in 1ml water followed by a second 15s centrifugation at 10,000g and the supernatant discarded. The cells were then resuspended and RNA degraded using 0.5ml of fresh Sigma (Sigma-Aldrich, Fancy Road, Poole, Dorset, BH12 4QH) RNase solution (2mg/ml chromatographically purified DNase-free RNase A in 50mM TrisCl, pH 8.0 and 15 mM NaCl). The cells were incubated for 2 hours at 37°C.

The cells were centrifuged (15 sec at 10,000 g) and the supernatant discarded to remove the bulk of the RNase, then resuspended in 0.2 ml of fresh protease solution (5 mg/ml Sigma porcine pepsin in water containing 4.5µl HCl/ml) and incubated for 30min at 37°C.

Cells were centrifuged (15s at 10,000g) and the supernatant discarded prior to resuspension in 50mM Tris-Cl. pH 8.0. 50µl of the cell suspension was transferred to a microcentrifuge tube containing 1ml of fresh 1µM SYTOX Green (Rieseberg et al. 2001) DNA staining solution (Molecular Probes, Eugene, Oregon) in 50mM Tris-Cl, pH 8.0.

Cell clumps were broken up by sonicating for 20s on low power in a water bath sonicator and cells were analysed using a flow cytometer with 488-nm excitation. The green fluorescence emission at 523 nm was collected using a Beckman Coulter Cyan ADP fluorescence cytometer (Beckman Coulter (UK) Ltd, Oakley Court, Kingsmead Business Park, London Road, High Wycombe HP11 1JU) with the flow rate on the lowest setting. Debris was gated by setting the trigger to green fluorescence. Control of the cytometer, data import and analysis was via the

Beckman Coulter software, Summit v4.3 (build 2451). Settings for each set of fluorescence cytometry samples were imported from the first run and these included a range of fluorescence intensity from 0 to 256 arbitrary intensity units. A plain text export from the Summit software was used to generate histograms of the cytometry run in R.

Haploid ancestral strains and their descendants were used to partition the data into ploidy bands and these bands were applied to all data for that run. The first band (R0) contained low SYTOX Green intensity (values 0 -~20) cells that had not been gated out. The second band contained the main haploid peak (H). The lower boundary of this peak is where the cell count falls to the baseline, the upper boundary is at the minimum before the cell count begins to increase for the diploid (haploid post-replication) peak. The third band contains the post-synthesis diploid peak (D), the upper bound of which is determined by the fall to a minimum of the diploid peak in the haploid ancestors and diploid evolved strains. This forms the lower bound of the fourth the polyploidy peak (P) which contained any cells with a greater intensity up to the maximum. The Summit software was used to export plain text counts of the number of cells in each band and these were analysed using ternary plots in R

Fluorescence cytometry data was visualised, and different runs compared, using ternary plots. Such plots require the cell count data to be partitioned between three bins, in this case between the three ploidy states: haploid, diploid and polyploid. The DNA intensity boundaries for ternary plotting in each run were determined from the histograms of haploid and diploid strains. The data was plotted using the ‘ternaryplot’ routine in the ‘vcd’, Visualising Categorical Data, package (version 1.2-13) in R (version 2.15.2).

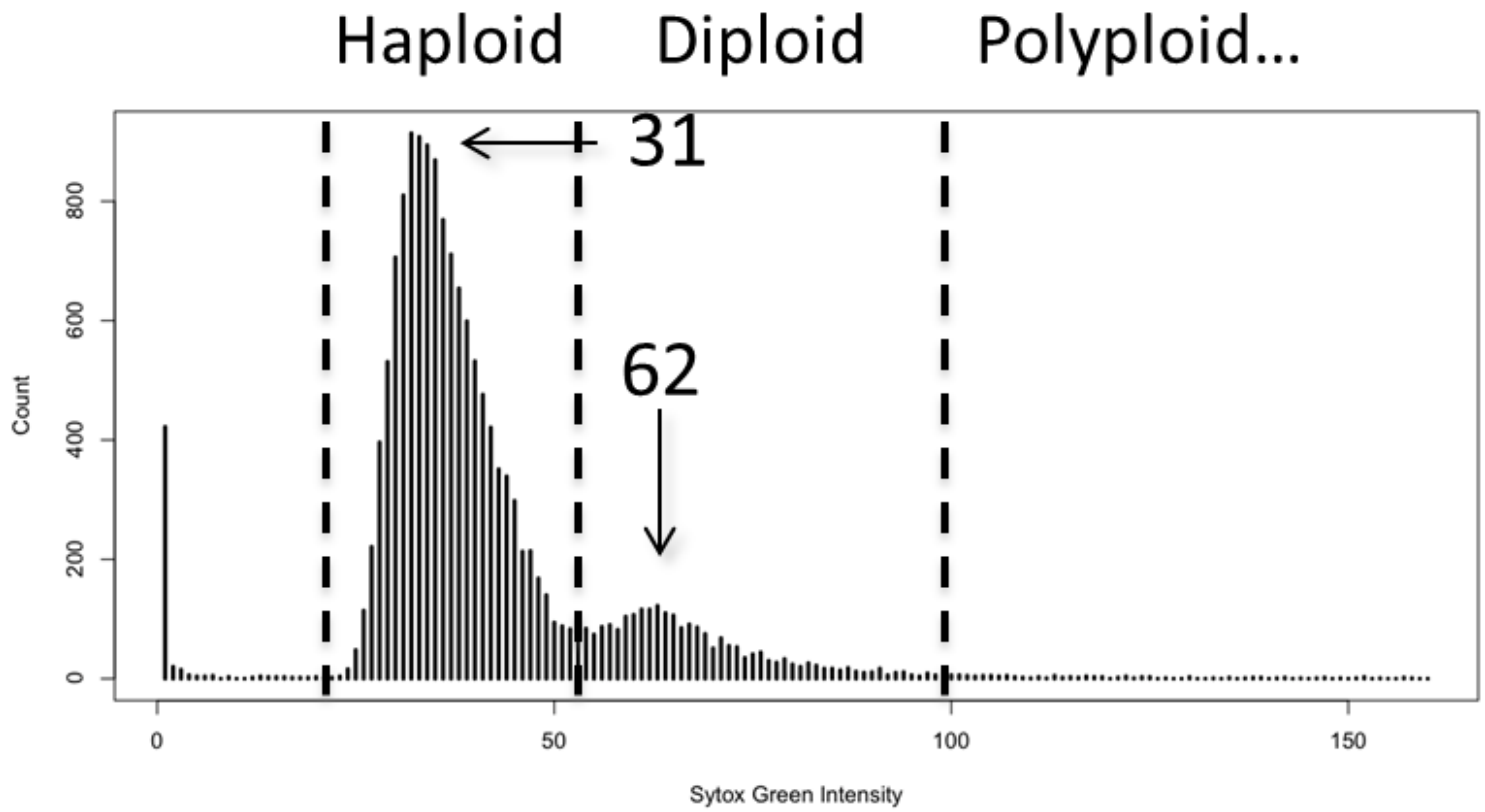


Figure 6.1 Partitioning a histogram of a Sytox Green DNA intensity into ploidy compartments. This histogram of a haploid strain shows a histogram of a haploid strain that has been used, with others to determine the haploid / diploid boundary. Other strains were used to determine the diploid / polyploid boundary.

6.3 Results

6.3.1. DNA sequence quality

FastQC analysis showed variation in the ratio of bases within the first eight bases at start of each read (see Figure 6.2). The same ratio pattern, with high adenine and low thymine was present in all strains in both forward and reverse read directions. A similar but smaller disparity is seen in the percentage of cytosine and guanine.

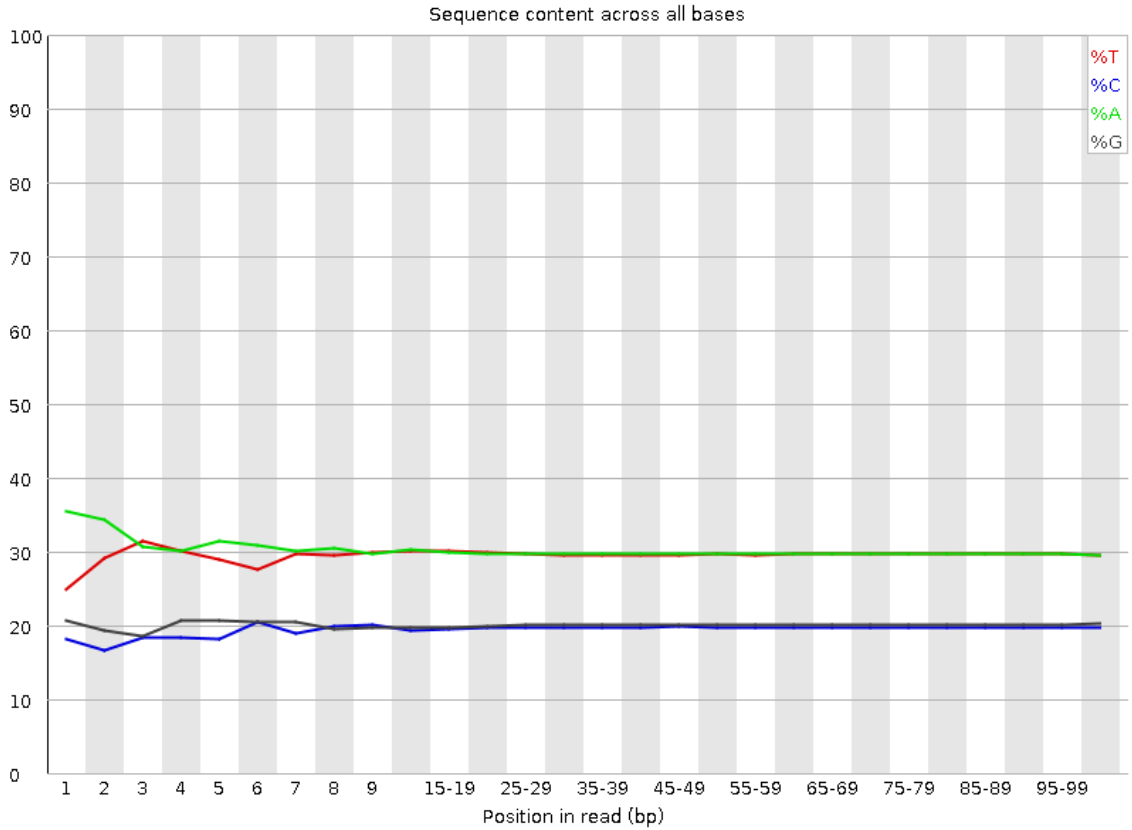


Figure 6.2 FastQC output: a plot of base ratio against position in read for whole genome sequence of Evo6_11_B06_CFPe_4 (reverse).
The percentage composition of the bases adenine (green), and thymine (red), cytosine (blue), and guanine (black) are shown against read position. The plot shows a substantial deviation from parity for complementary bases, particularly adenine and thymine which are higher and lower than parity respectively, in the first 8 bases. A similar but smaller disparity is seen in the percentage of cytosine and guanine.

There were no highly represented kmers in the initial bases and searches for pattern matches identified no known Illumina adaptor sequences.

The read length (most reads between 100-102 bases) and the quality of the sequences were both good. Figure 6.3 is a typical FastQC plot showing the mean read quality (blue line) only dipping below 28 above a length of 99bp.

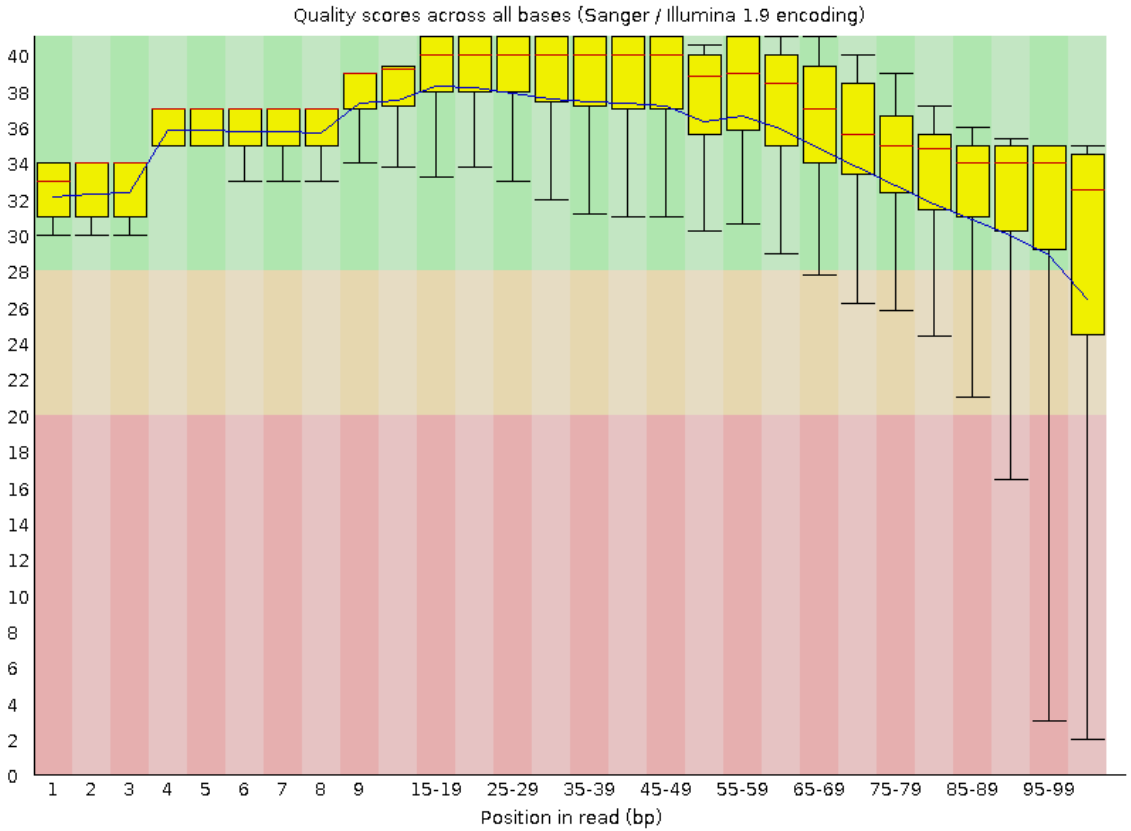


Figure 6.3 FastQC plot of quality against position in reads for whole genome sequence of Evo6_11_B06_CFPe_4 (reverse).
In this boxplot, the yellow box represents the inter-quartile range (25-75%) with the median marked with a red line and the 10% and 90% points marked by whiskers. The blue line represents the mean quality. The quality profile of these reads is typical and is generally good with the median read quality above 28 at the end of the read. However, there are stepped increases in read quality at the beginning of the read with transitions at bases 3-4 and 8-9 that are not typical of Illumina output.

All of the genome libraries generated either quality control warnings or failures because the fraction of duplicated sequences was greater than either 20% or 50% respectively (see Figure 6.4 for the FastQC plot of read duplicates for the reverse read of Evo6_11_B06_CFPe). The five strains that failed were: the ancestral YFP, Evo6_11_D06_YFP, Evo8_14_E07_CFP_4, Evo8_14_F03_CFPe_2, and Evo8_14_D03_YFPe_1.

However, the read depth of the sequence data was extremely high and the removal of duplicates did not compromise general coverage. Figure 6.5 shows a typical histogram of read depth at every uniquely mapped locus for the strain Evo8_14_E07_CFP_4, a sequence that failed quality control because of excessive duplication; with duplicates removed, 94% of the mapped genome had a read depth in excess of x100, with the mode being approximately x150

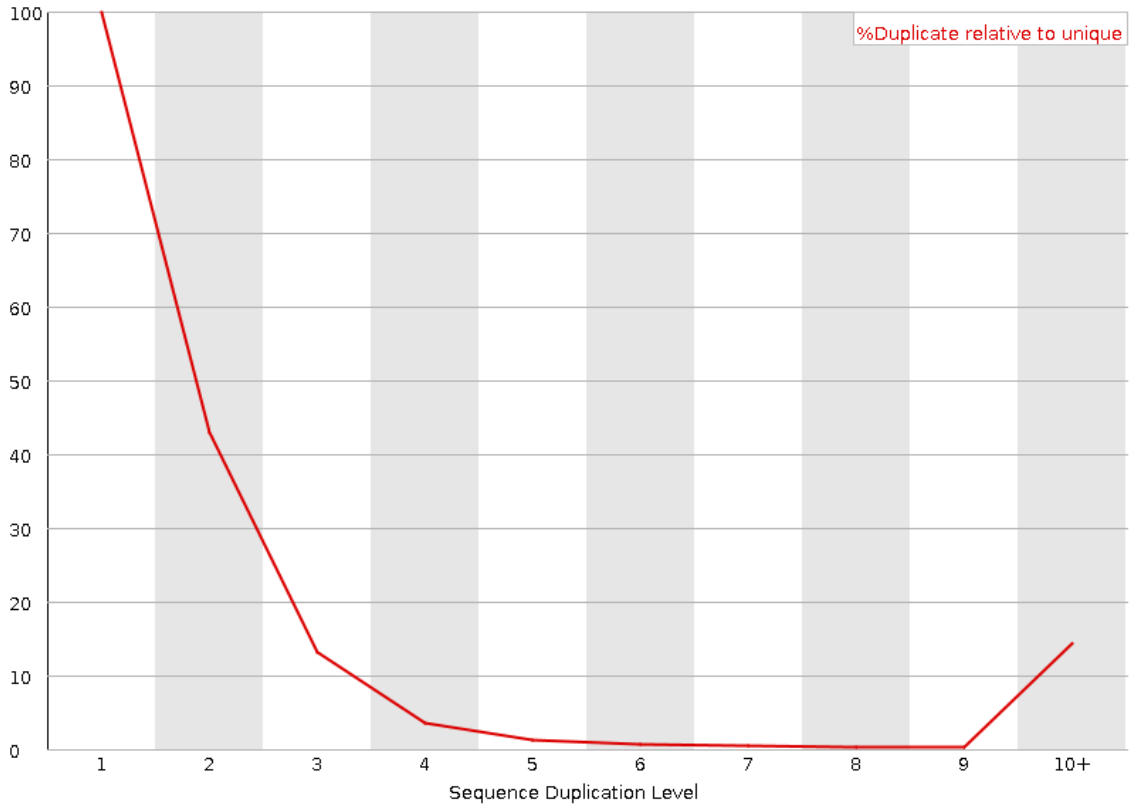


Figure 6.4 Sequence duplication in reads for whole genome sequence of Evo6_11_B06_CFPe_4 (reverse).
Analysis of the mapped reads shows substantial duplication; >=45.97% of reads are duplicated and a minority have ten or more duplicates.

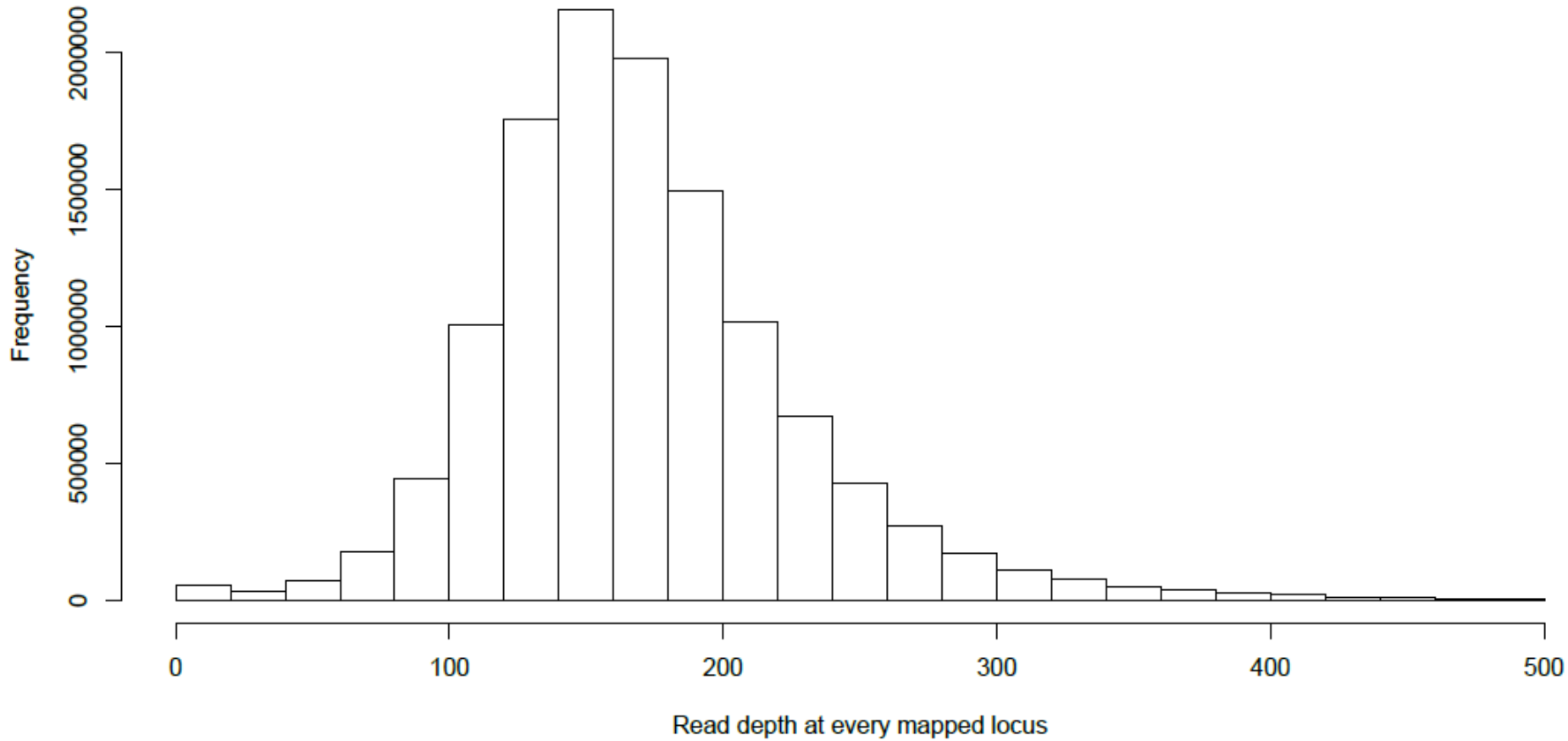


Figure 6.5 The frequency of read depths at every mapped locus for the whole genome sequence Evo8_14_E07_CFP_4.
The genome sequence reads for this strain failed quality control because more than 50% were duplicated. With duplicates removed, the mode read depth remains ~150x. About 0.46% of the mapped loci were omitted because they had depths in excess of 500 and 94% of the mapped loci had read depths of x100 or greater.

Analysis of the sizes of the inserts between the paired reads appears to show a bimodal distribution with two populations of inserts with sizes distributed around peaks at ~180bp and ~300bp. The ratio of the insert populations varies between strains, sometimes manifesting itself as a skew, see the insert size distribution of Evo8_14_E07_CFPe_4 in Figure 6.6.a), and sometimes a distribution with two peaks, see Figure 6.6.b) and c) for the distribution of Evo10_20_F10_CFPe_2 and the ancestral YFP (sequence 1) respectively.

The quality of the mapped reads was consistent for each strain, 83-86% of the reads for each strain were mapped with a quality of 30 or greater. Figure 6.7 shows a typical chart of mapping quality, for strain Evo10_20_F10_CFPe_2.

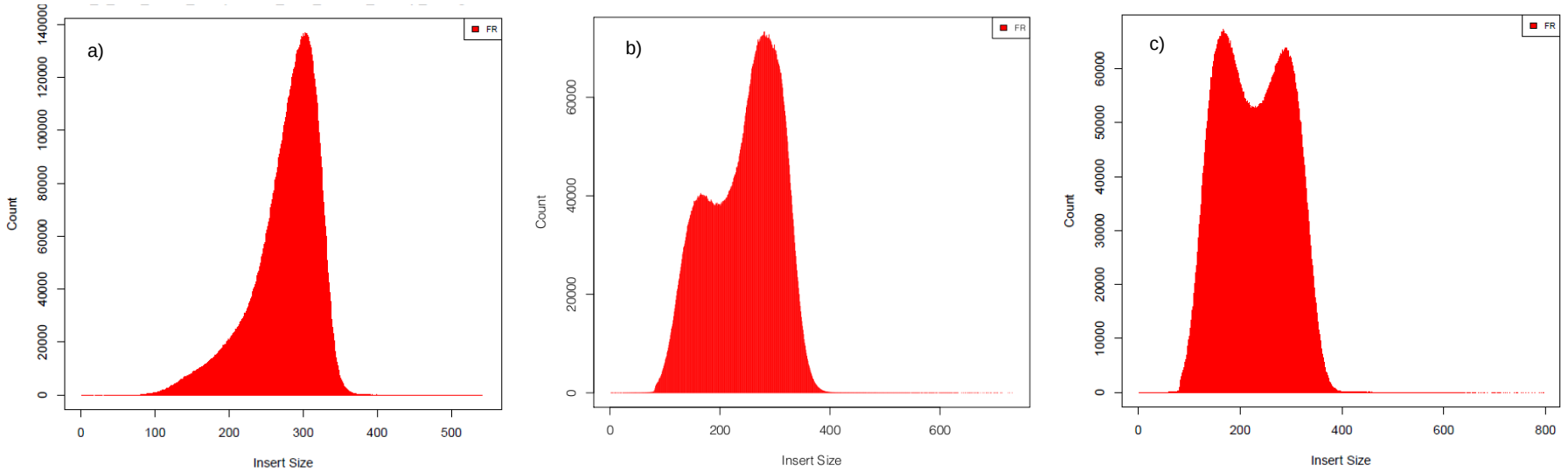


Figure 6.6 A histogram of the paired end insert sizes for three mapped sequences: a) Evo8_14_E07_CFPe_4, b) Evo10_20_F10_CFPe_2 c) ancestral YFP (first sample).

None of the insert size histograms are normally distributed. There appear to be two populations of inserts with sizes distributed around peaks at ~180bp and ~300bp. The population of short inserts is variable and manifests itself as a skew in Figure a) Evo8_14_E07_CFPe_4, a secondary peak in Figure b) Evo10_20_F10_CFPe_2, and a major peak in c) ancestral YFP (sequence 1). The data has had duplicates removed.

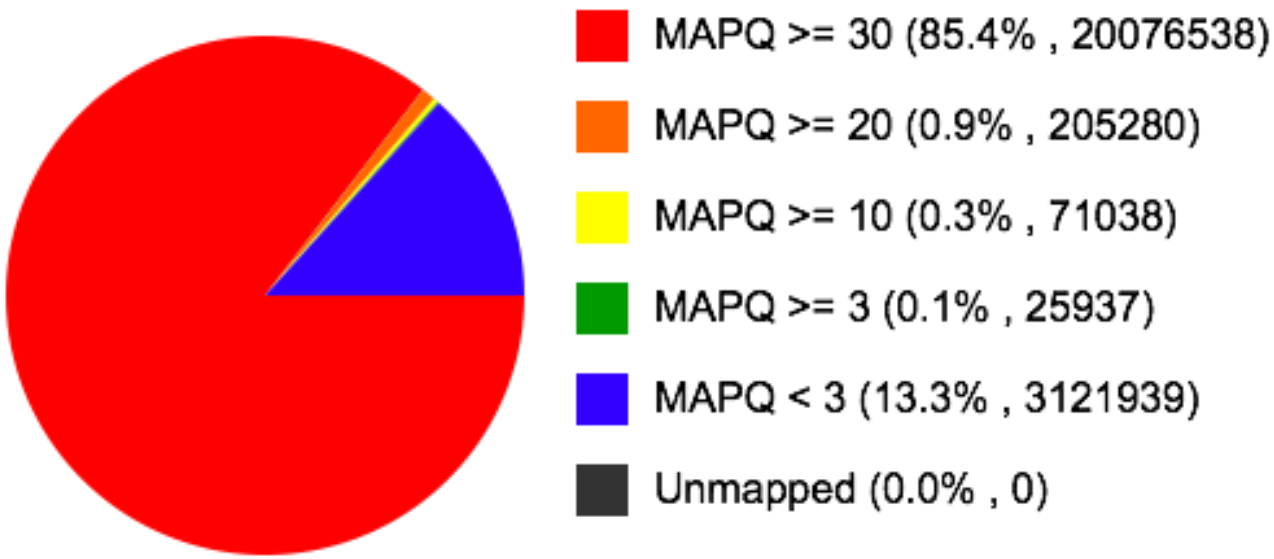


Figure 6.7 Chart of mapped read quality for strain Evo10_20_F10_CFPe_2.
100% of the reads had an alignment with a quality greater than 0. Duplicate reads had been removed prior to analysis.

6.3.2. DNA Sequence Analysis

Table 6.1 shows a summary of the results for this chapter, listing the genetic changes found in each sequenced strain.

Strain	Environment (% ethanol)	Gene/Region	Type	Zygosity
Both haploid starting strains	NA	*MEH1* (YKR007W)	Deletion	NA
Both haploid starting strains	NA	*APC1* (YNL172W)	Deletion	NA
Both haploid starting strains	NA	*CTR9/ CDP1* (YOL145C)	SNP (non-synonymous)	NA
CFPa haploid starting strain	NA	*EGT2* (YNL327W)	Deletion	NA
Evo6_11_B06_CFPe_4	4.5% (fixed)	*SSL2*	SNP (synonymous)	Heterozygous
Evo6_11_D06_YFPe_4.	4.5% (fixed)	-	-	-
Evo8_14_E07_CFPe_4	0% (control)	-	-	-
Evo8_14_F03_CFPe_2	4.5% (fixed)	*GPI18*	SNP (non-synonymous)	Homozygous
Evo8_14_F03_CFPe_2	4.5% (fixed)	*GCN1*	SNP (non-synonymous)	Heterozygous
Evo8_14_D03_YFPe_1	6.5% (fixed)	*COB1*	SNP (non-synonymous)	NA (mitochondrial)
Evo8_14_G04_YFPe_1	6.5% (fixed)	*COB1*	SNP (non-synonymous)	NA (mitochondrial)
Evo8_14_G04_YFPe_1	6.5% (fixed)	*PHO23*	SNP (non-synonymous)	Heterozygous
Evo10_20_F10_CFPe_2	9.0% (ramped)	-	-	-
Evo10_20_G10_YFPe_2	9.0% (ramped)	Chromosome V (450,00-500,000)	Deletion	Heterozygous
Evo10_20_G10_YFPe_2	9.0% (ramped)	*GDB1*	SNP (non-synonymous)	Homozygous

Table 6.1 Summary of genomic changes in sequenced strains.
Strains with no detected genetic changes are indicated with a dash.

6.3.3. Ancestral strains.

The documented multiple auxotrophies of the starting strains were clearly evident; there was no coverage of the documented deleted regions and BreakDancer was able to identify the published deletions and translocations (e.g. the *TDH3* promoter for the fluorophore at the *HO* locus). In addition to the engineered deletions in the starting strains (see Introduction, section 2.6), both starting strains have additional undocumented deletions in the genes *MEH1* (YKR007W; chrXI:451569-574), involved in vacuolar acidification and *APC1*(YNL172W; chrXIV:310636-315882), the largest subunit of the anaphase-promoting complex (an essential component of cell cycle progression). There are also two substitutions that are 9 bases apart in *CTR9* (YOL145C; chrXV:52788-49555), a component of the *Paf1p* complex involved in transcription elongation.

The *MEH1* deletion is in frame, removing a glutamine-histidine repeat in both ancestral fluorophores. The deleted sequence '`GAACAC`' removes amino acids 46-47 which is repeated in the reference gene (see highlighted region in Figure 6.8):

```
  1  MGAVLSCCRN HSGEENEALL REQQAGYGSQ GNANDEYDAE QMRLKEHEHE
 51  QKLLAREQEL RDIVANTNDK LIDISMINNS GIVIQGTDLQ EALDKRQQEE
101  GGDSREDERS AGDDNLSGHS VPSSGSAQAT THQTAPRTNT FTLLTSPDSA
151  KISKEQLKKL HSNILNEIFS QSQVNKPGPL TVPF*
```

Figure 6.8 Annotated amino acid sequence of MEH1 gene (YKR007W) showing (highlighted in yellow) the glutamine and histidine residues present in the S288c genome reference strain but absent from both CFP and YFP experimental evolution starting strains .
The two residue sequence is repeated in the reference strain but not in the starting strains.

The *APC1*(YNL172W) deletion is in frame and removes a GAC codon (aspartic acid) ChrXIV: 312329-312331 in all sequenced strains.

There are two non-synonymous SNPs in the ancestral *CTR9*/*CDP1* (YOL145C, ChrXV:52788-49555) locus in the ancestors. SNP clusters might indicate a small indel that local realignment has not been able to resolve however, this does not appear to be the case. Unresolved deletions usually create a variety of weakly supported SNPs, particularly in reads close to their ends at that position, whereas only two G→T SNPs are seen (at chrXV:50,002 and 50,012) and the bases around each SNP are not the same as is common with unresolved deletions. The *CTR9* protein contains 16 tetratricopeptide (TPR) repeats but the SNPs seen at amino acids 926 and 929 lie outside the repeated region. The two SNPs cause two substitutions in the 1173 amino acid protein, a glycine to valine substitution at position 926 and lysine to aspargine at 929 (Figure 6.9).

S288c 50,014*	CAA	GGA	TGG	ATG	AAG	CAG	49,999
Amino Acids	Q	G	W	M	K	Q	

CFPa/YFPa	CAA	GTA	TGG	ATG	AAT	CAG
Amino Acids	Q	V	W	M	N	Q

Figure 6.9 DNA and amino acid sequences of *S. cerevisiae* CTR9/ CDP1 (YOL145C, ChrXV:50,014-49,999) for the S288c and the parental CFP and YFP strains.
Yellow highlights in the DNA sequence show strain differences, green highlights show changes in the amino acid sequences. This coding sequence is on the Crick strand.

The sensu strictu strains (Byrne & K H Wolfe 2005) and the Kyokai No.7 strain (Akao et al. 2011) share lysine at position 929 apart from *S. castelli* which has glutamine. However, while *S. cerevisiae* has glycine at position 926, the rest of the sensu strictu has glutamic acid in this position as does Kyokai No.7 (which also has a substitution at position 927, tryptophan to arginine), see Figure 6.10.

S288c	901	EFEKEQSAKI DEARKILEEN ELKEQGWMKQ EEEARRLKLE
CFPa/YFPa	901	EFEKEQSAKI DEARKILEEN ELKEQVWMNQ EEEARRLKLE
Kyokai No.7	901	EFEKEQSAKI DEARKILEEN ELKEQERMKQ EEEARRLKLE

Figure 6.10 Amino acid sequence comparison between the *S. cerevisiae* strains S288c, the experimental evolution starting strains CFPa and YFPa, and the sake strain Kyokai No.7 for CTR9/ CDP1 (YOL145C), showing amino acids 901 to 1021.

In addition to these variations from the S288c reference that are common to both ancestral strains in these experimental evolutions, there is a deletion in the CFP ancestor that is not found in the YFP ancestor in the non-essential gene *EGT2* (YNL327W) on chromosome XIV. This codes for a glycosylphosphatidylinositol (GPI)-anchored cell wall endoglucanase required for proper cell separation after cytokinesis. The boundaries of the deletion are not clearly resolved but it is approximately 30bp in length at, or close to, 26733 and 26762 (amino acids 896-905 in the 1041 amino acid gene product). This deletion removes the first ten amino acids (amino acids 896 to 905, STTVGAAQYA) in the third of four 38 amino acid tandem repeats; 820-857, 858-895, 896-933, 934-972. The second and the third partially deleted repeats are identical (see Figure 6.11). The repetition may be causing the poor deletion boundary resolution and may make the true nature of the genetic change uncertain. Since the deletion boundaries are poorly resolved it is not certain that the deletion is in frame. However, the clear change in the CFP ancestor

and all strains derived from it and the absence of such change from all YFP strains strongly suggests that a genetic change has occurred.

```
820
GTCTESGQATEYGSLIPISTLDGSVIYTFTGESVVVGY
858
STTVGAAQYAQHTSLVPVSTIKGSKTSLSTEESVVAGY
896
STTVGAAQYAQHTSLVPVSTIKGSKTSLSTEESVVAGY
934
STTVDSAQYAEHTNLVAIDTLKTSTFQKATATEVCVTC
```

Figure 6.11 Amino acid sequence alignment of repeated motifs in EGT2p.
The section highlighted in teal shows the ten amino acid deletion seen solely in the ancestral CFP strain. The second and third repeats (starting at amino acid 858 and 896 respectively) are otherwise identical.

6.3.4. Evolution 6 (4.5% static ethanol stress)

In the evolution 6 strain Evo6_11_B06_CFPe_4, there is a synonymous heterozygous SNP in chromosome IX with strong support (map quality 175, depth 168), a transition from C to T at position 83021 SSL2 codon TAC to TAT, which both encode tyrosine. However, there are no non-synonymous SNPs, indels, copy number changes or other sequence changes with good support seen in either Evo6_11_B06_CFPe_4 or Evo6_11_D06_YFPe_4.

6.3.5. Evolution 8 (0%, 4.5%, 6.5% static ethanol stresses)

No fixed ethanol stress.

Subsequent work by Bharat Rash (pers. comm.) showed that Evo8_14_E07_CFPe_4 is heterozygous at the MAT locus, which confers histidine synthetic capacity (See chapter 5, section 5.3.5 for discussion). No other sequence changes with good support were seen in this strain.

4.5% Ethanol Stress

The sequenced strain Evo8_14_F03_CFPe_2 has two strong non-synonymous SNPs with good coverage. The first SNP is homozygous on chromosome II at position 244370 a C to A transversion from GCC (alanine) to GAC (valine), with mapping quality of 222 and coverage of 71. It is in the last codon prior to stop in the essential gene *GPI18* (YBR004C), a mannosyltransferase that transfers the second mannose in glycosylphosphatidylinositol (GPI) biosynthesis (Kang et al. 2005; A.-L. Fabre et al. 2005).

The second SNP is in a large regulatory protein GCN1 (YGL195W chrVII:131,525-139,543), a G to C transversion at position 134,717 with strong map support (225) and good coverage (102) supported by both forward and reverse reads. The SNP is

heterozygous (56% C to 44% G) and changes GAA (glutamic acid) to CAA (glutamine) at position 1065 in the protein. Apart from *Kluveromyces lactis* which has a histidine, the rest of the *Saccharomyces* sensu-strictu shares glutamic acid at this locus. *GCN1* contains 20 HEAT (Huntington, Elongation Factor 3, PR65/A, *TOR*) domains, which consist of 47 residues coding for two anti-parallel α-helices (Andrade & Bork 1995). The SNP is two residues from the end of the second HEAT repeat (1030-1067). Despite the large number of repeats the sequence around this SNP is unique and it is unlikely to be an error due to misalignment of repeated sequences although Sanger confirmation of changes in genes with repeats is highly desirable.

6.5%. Fixed Ethanol Stress

Both strains in this treatment Evo8_14_D03_YFPe_1 and Evo8_14_G04_YFPe_1 have the same mitochondrial SNP (M.41413, a transversion from T to G). For Evo8_14_D03_YFPe_1 the map quality is 222 (strong) with coverage x192, for Evo8_14_G04_YFPe_1 the map quality is 179 (medium) with coverage x32. No other base was called at this locus in either strain. The SNP is in the large (1.4kb) fourth intron of COB1, one of seven mitochondrial proteins essential for respiration. Bharat Rash (personal communication) confirmed that these strains can grow on glycerol and are hence able to respire. Mitochondrial introns are ribozymes that catalyse their own splicing in vitro (Pel & Grivell 1993). Part of the previous *COB* exon and the intron itself encode the mitochondrial mRNA maturase BI4 (Q0120) which mediates the splicing of intron 4 as well as the AI4 intron in *COX1* the gene coding for subunit I of cytochrome oxidase (Goguel et al. 1989). To do this the maturase forms a ternary complex with the leucine tRNA synthetase (*NAM2*) (Rho & Martinis 2000).

In Evo8_14_G04_YFPe_1 there is an additional strong (map quality 124, depth 250) non-synonymous heterozygous SNP in chromosome XIV.441743 a transversion from A to C in the non-essential gene *PHO23* (ChrXIV:442358 to 441366), AGA (arginine) to GGA (glycine) at position 206 (see Figure 6.12). This SNP is not present in Evo8_D03_YFPe_1.

```
                    201

S288/YFPa           NNTNSRKRAN AANTNNADPE

Evo8_14_G04_YFPe_1  NNTNSGKRAN AANTNNADPE

Evo8_14_D03_YFPe_1  NNTNSRKRAN AANTNNADPE
```

Figure 6.12 Amino acid sequence for part of PHO23 gene showing the amino acid change from arginine (R) to valine (G) in Evo8_14_G04_YFPe_1 at position 206.
The changed residue is highlighted in yellow**.**

6.3.6. Evolution 10 (ethanol stress ramped to 9%)

No genetic changes from the ancestral CFP strain were detected in Evo10_20_F10_CFPe_2. However, copy number variation shows an approximately 50kb heterozygous deletion in a gene-rich subtelomeric region of chromosome V (from ~450,000 to 500,000) of the Evo10_20_G10_YFPe_2 strain (see Figure 6.13). In this region genes are present as a single copy against a diploid background. There are copies of the *TYA Gag* and *TYB Pol* genes starting at 449024 (YER138C and YER137C-A) and 498124 (YER 160C and YER159C-A) which are likely sites for homologous recombination. No reports of intra or inter-chromosomal rearrangement were found using BreakDancer in this strain.

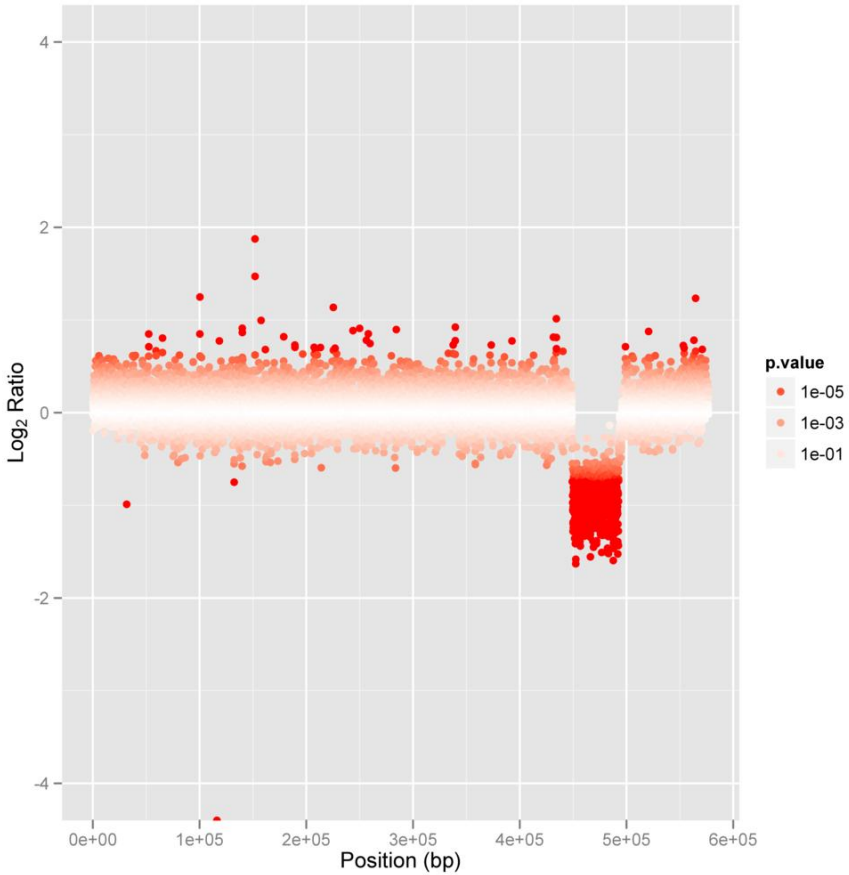


Figure 6.13 Copy number variation in chromosome V for strain Evo10_20_G10_YFPe_2 reads compared with ancestral YFP showing a deletion between ~450,000-500,000.
The reads are approximately halved, a significant reduction ($p<=1x10^{-5}$).

Evo10_20_G10_YFPe_2 also has a strong (mapping quality 222, coverage x180) homozygous, non-synonymous, SNP, a T to A transversion in chromosome XVI at 903652. The codon change is TGC (cysteine) to AGC (serine) in amino acid 537 of the non-essential gene for the glycogen debranching enzyme *GDB1* (YPR184W). *GDB1p* is a 1,536 amino acid enzyme containing glucanotransferase and alpha-1,6-amyloglucosidase activities required for glycogen degradation.

6.3.7. Copy number and structural variation

The CNV-seq comparison of the evolved strains with their haploid ancestor suggests that all sequenced strains are euploid. CNV-seq identified a 50kb deletion (450-500 in chromosome V in Evo10_20_G10_YFPe_2). BreakDancer looks for anomalously mapped read pairs (ARPs), reads where both of a pair have MAQ mapping quality >10 with a separation distance greater than 3 standard deviations (Chen et al. 2009). BreakDancer was able to identify the published deletions and translocations (e.g. the *TDH3* promoter for the fluorophore at the *HO* locus) of the starting strains but did not identify any strong candidates for structural change.

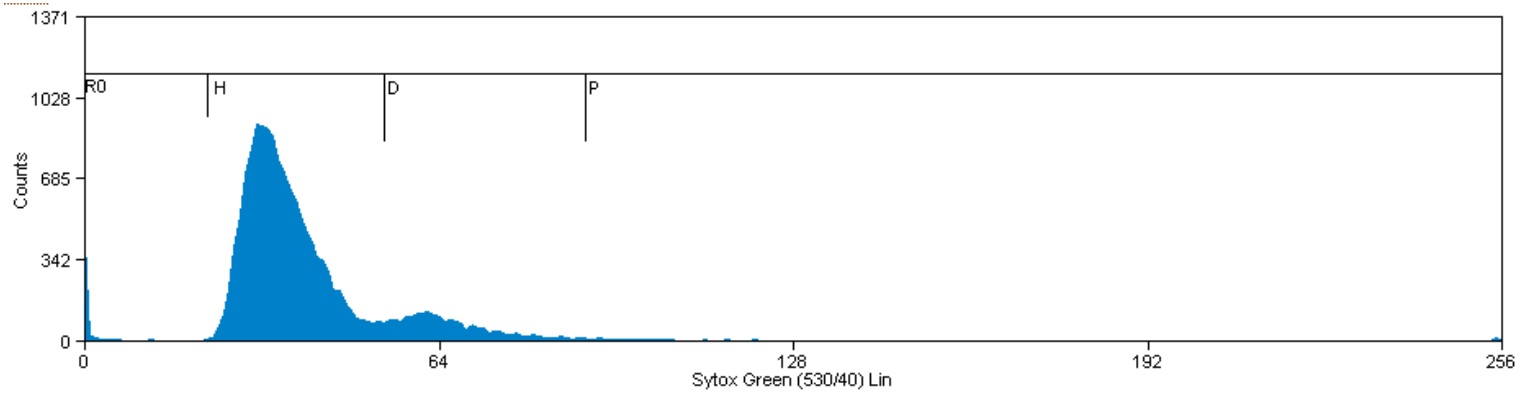
6.3.8. Fluorescence Cytometry - Analysis of Ploidy

Fluorescence cytometry with Sytox Green DNA staining showed a clear difference between the ancestral and evolved strains (see Figure 6.14 overleaf for histograms of DNA intensity for each strain). The mode of the dominant peak in for the ancestral strains has a fluorescence intensity of 31-34 arbitrary units, with a diffuse minor peak of 62-63 units. The small secondary peak, with slightly less than double the DNA intensity, is consistent with a signal from a minority of the cell population that has replicated but not completed division and separated from its bud. This is consistent with the Nexcelom micrographs of the diluted cultures (see Figure 3.1) which showed pairs of large and small cells consistent with budding cells being a visible fraction of the population even in the slow growth phase cultures used for cell size and fluorescence cytometric analysis.

All the evolved strains sequenced (from the final plates in each evolution) lack significant counts above background in the region of peak fluorescence intensity shown by the ancestral strains. However, they do have very similar major peaks with modes of 65-80 units, slightly more than twice the ancestral major peaks and close to the diffuse minor peak for post replicative cells. All sequenced strains have diffuse minor peaks with a mode of 90, consistent with a minority population of post-replicative cells. This is consistent with the evolved strains having twice the DNA content of the starting strains. Since the starting strains are described as

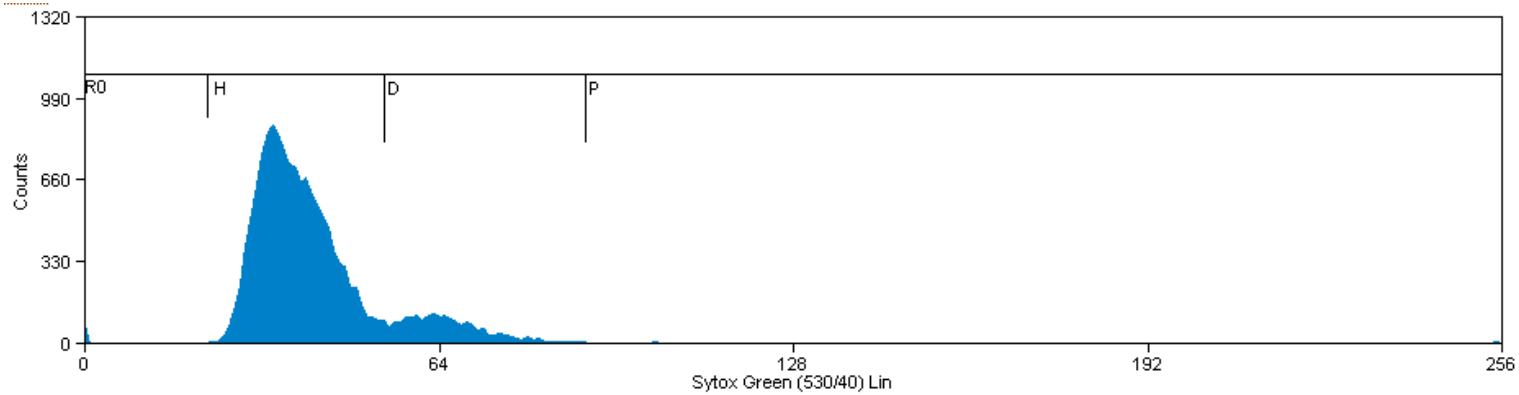
heterothallic haploids (DeLuna et al. 2008) the evolved strains appear broadly diploid in DNA content, although fluorescence cytometry is not sufficiently accurate to determine that they are euploid.

Despite adherence to standardised methods and machine calibration, the intensity of stain uptake in each batch varies as does the intensity estimates by the cytometer, accordingly the results shown in Figure 6.14 and Figure 6.20 come from single batches of fixed and stained cells processed in the same run of the cytometer.



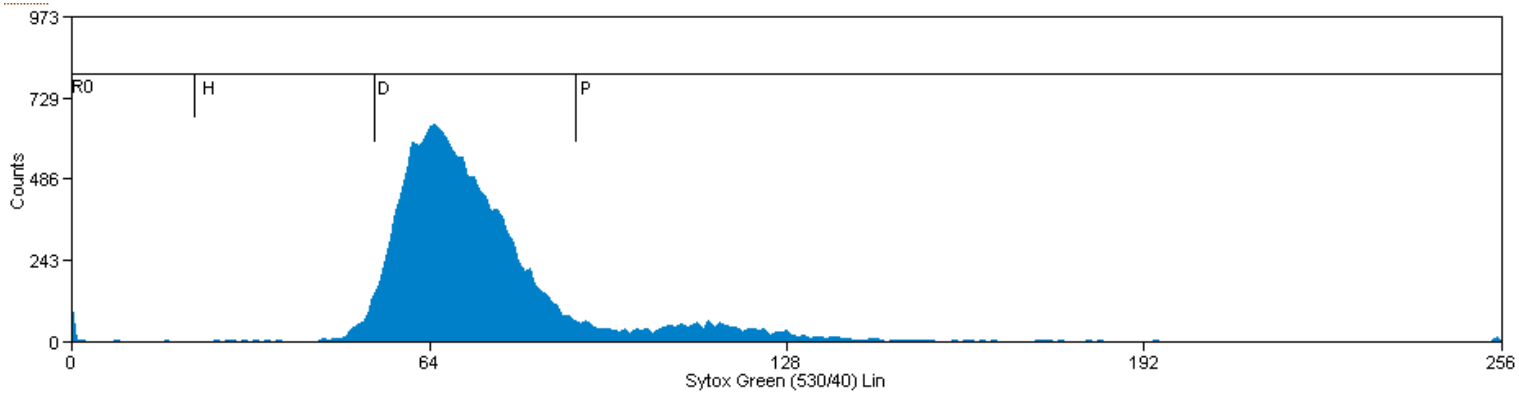
1. CFPa

Mode of DNA intensity=31 (minor peak =62)



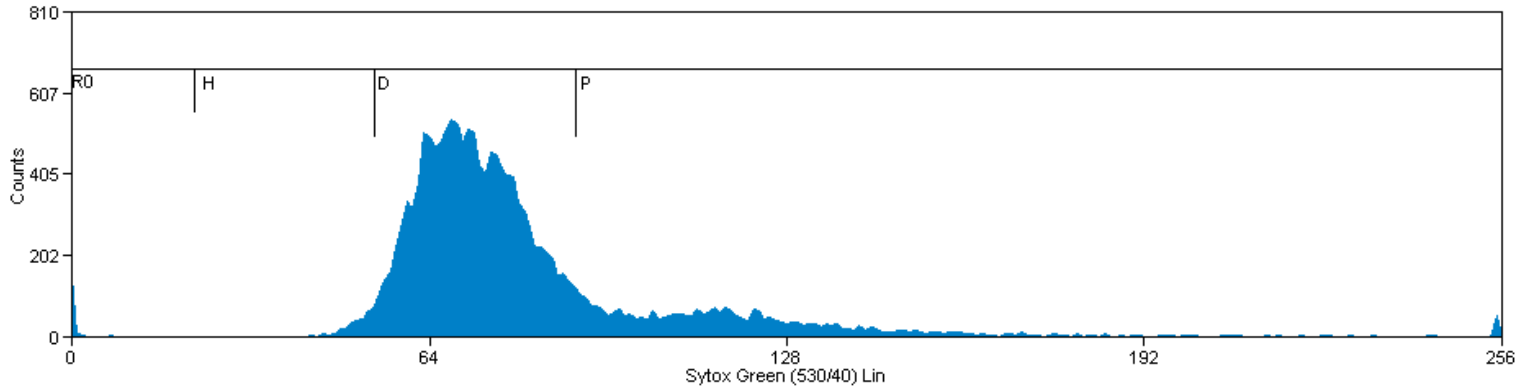
2. YFPa

Mode of DNA intensity=34 (mode of minor peak = 63)



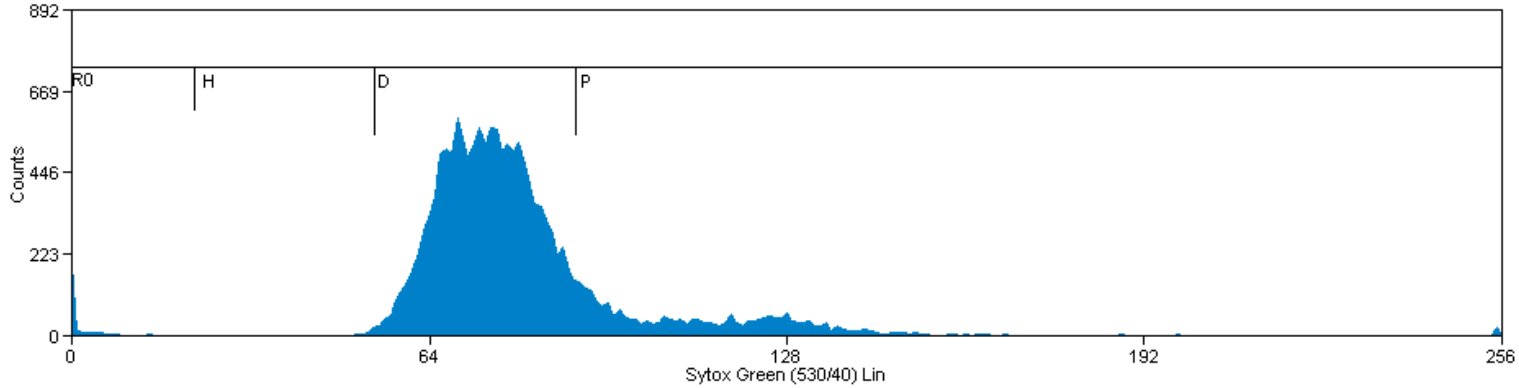
3. Evo6_11_B06_CFPe

Mode of DNA intensity=65 (mode minor peak=90)



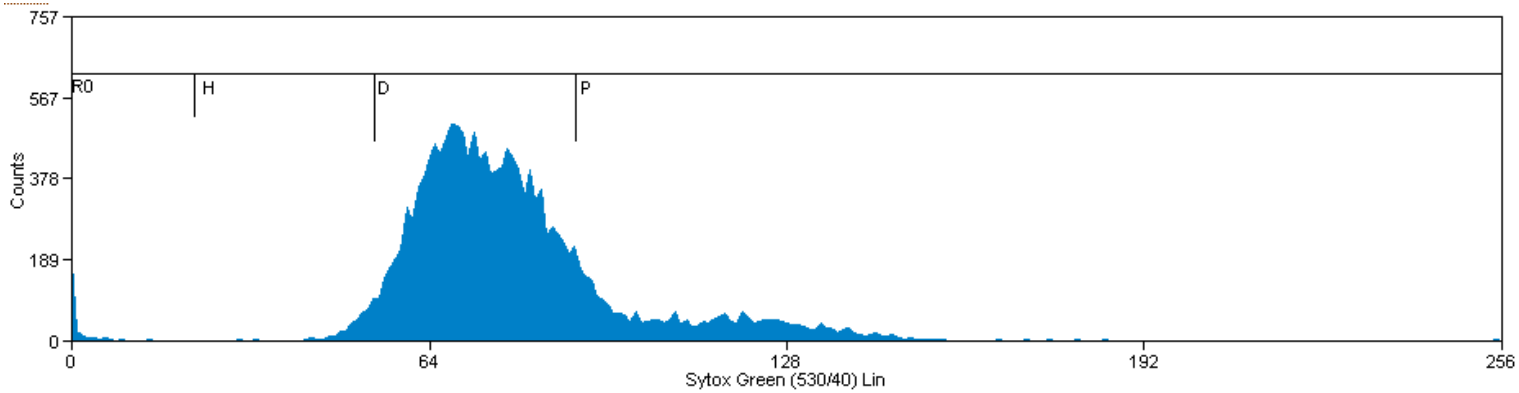
4. Evo6_11_D06_YFPe

Mode of DNA intensity=68 (mode of minor peak=90)



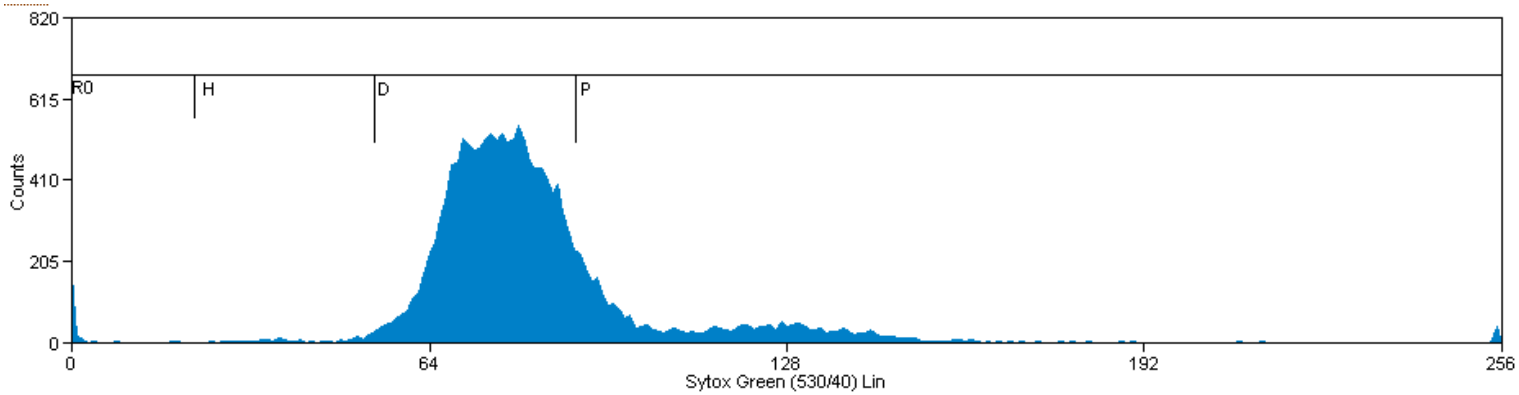
5. Evo8_14_D03_YFPe

Mode of DNA intensity=69 (mode of minor peak=90)



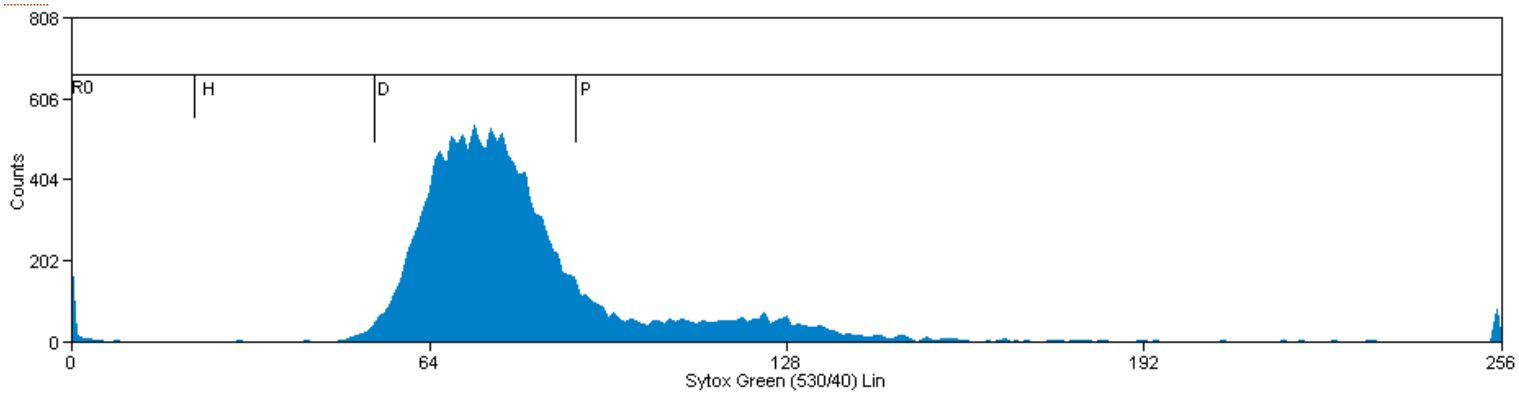
6. Evo8_14_E07_CFPe

Mode of DNA intensity=68 (mode of minor peak=90)



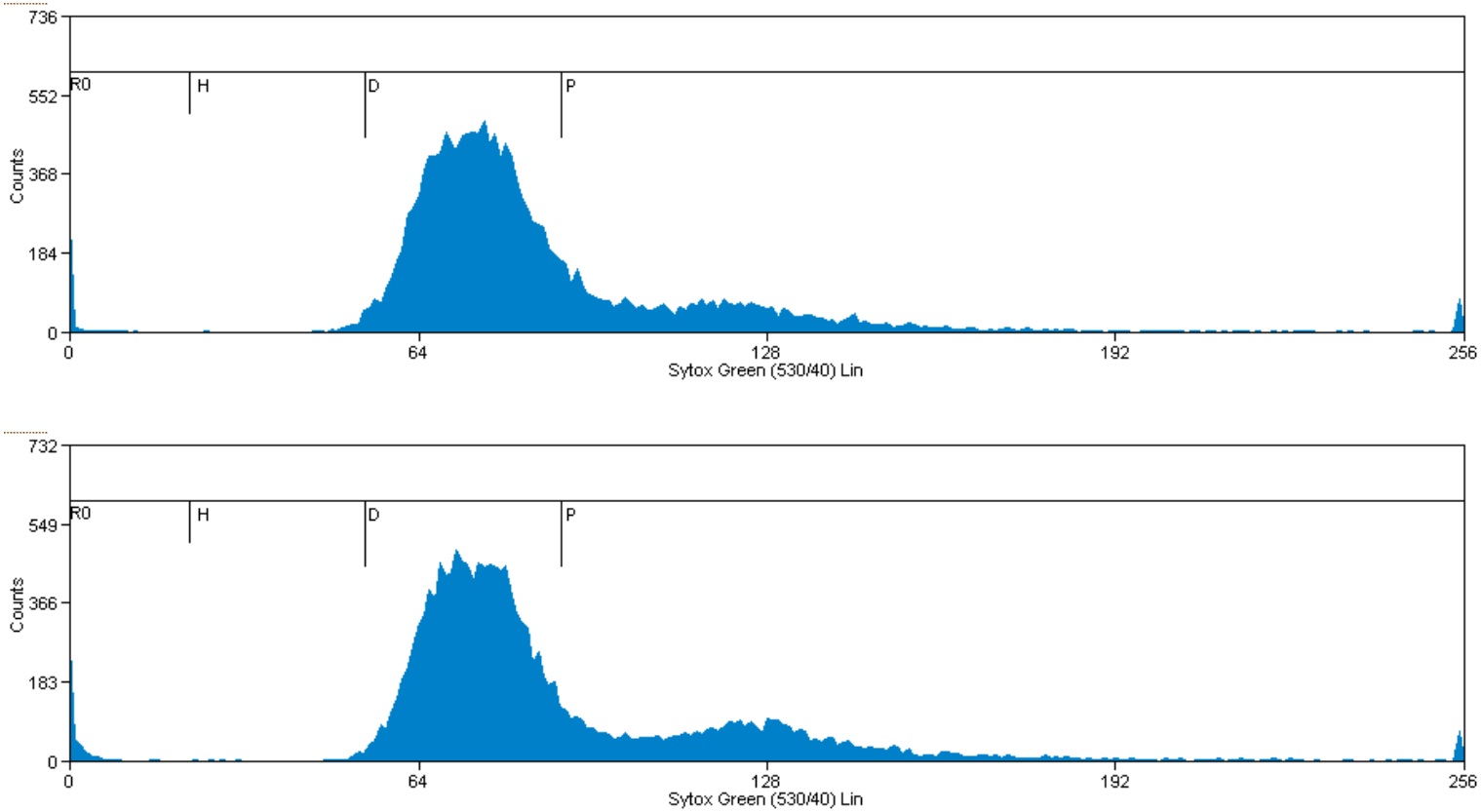
7. Evo8_14_F03_CFPe

Mode of DNA intensity=80 (mode of minor peak=90)



8. Evo8_14_G04_YFPe

Mode of DNA intensity=72 (mode of minor peak=90)



9. Evo10_20_F10_CFPe

Mode of DNA intensity=76 (mode of minor peak=90)

10. Evo10_20_G10_YFPe

Mode of DNA intensity=71 (mode of minor peak=90)

Figure 6.14 Histogram of DNA intensity FACS Run 1

Correlated with the change in ploidy in these evolved strains there was a clear and consistent increase in cell size. The haploid starting strains have a mean cell diameter of approximately 3 microns whereas the mean of evolved diploids is close to 4.4 microns (see Table 6.2). The cell size data is a mean of the longest and shortest dimensions in each cell image, and hence cannot capture any potential changes in cell shape.

Strain	**Mean Diameter (microns)**
CFPa	3.07
YFPa	3.06
Evo6-B06-CFPe-4	4.26
Evo6-D06-YFPe-4	4.56
Evo8-D03-YFPe-1	4.32
Evo8-E07-CFPe-4	4.44
Evo8-F03-CFPe-2	4.25
Evo8-G04-YFPe-1	4.33
Evo10-F10-CFPe-2	4.44
Evo10-G10-YFPe-2	4.48

Table 6.2 Mean of cell diameters (microns) for strains in fluorescence cytometry run 2.

Density plots of cell size for haploid and diploid strains in Figure 6.15, Figure 6.16, and Figure 6.17 show the cell size distributions for ancestral and evolved cells in evolutions 6, 8 and 10 respectively. The ancestral haploid and evolved diploid cell size distributions have significant overlap but well-separated peaks and the modes and profiles of these peaks are very similar in every evolved strain. The diploid density distributions are more negatively skewed than their haploid counterparts.

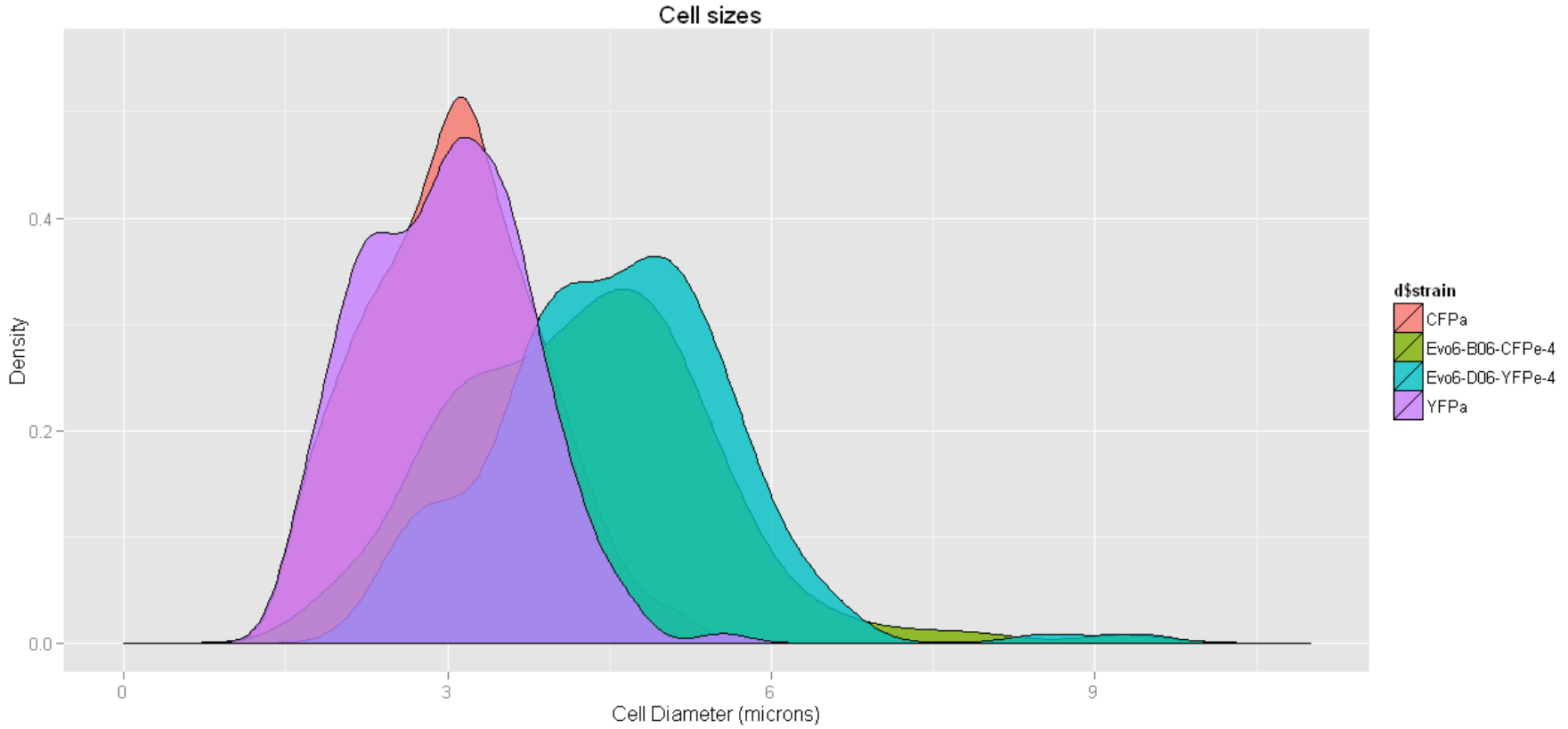


Figure 6.15 A density plot of cell diameter (microns) for the ancestral strains and the evolution 6 strains Evo6_11_B06_CFPe_4, Evo6_11_D06_YFPe_4.

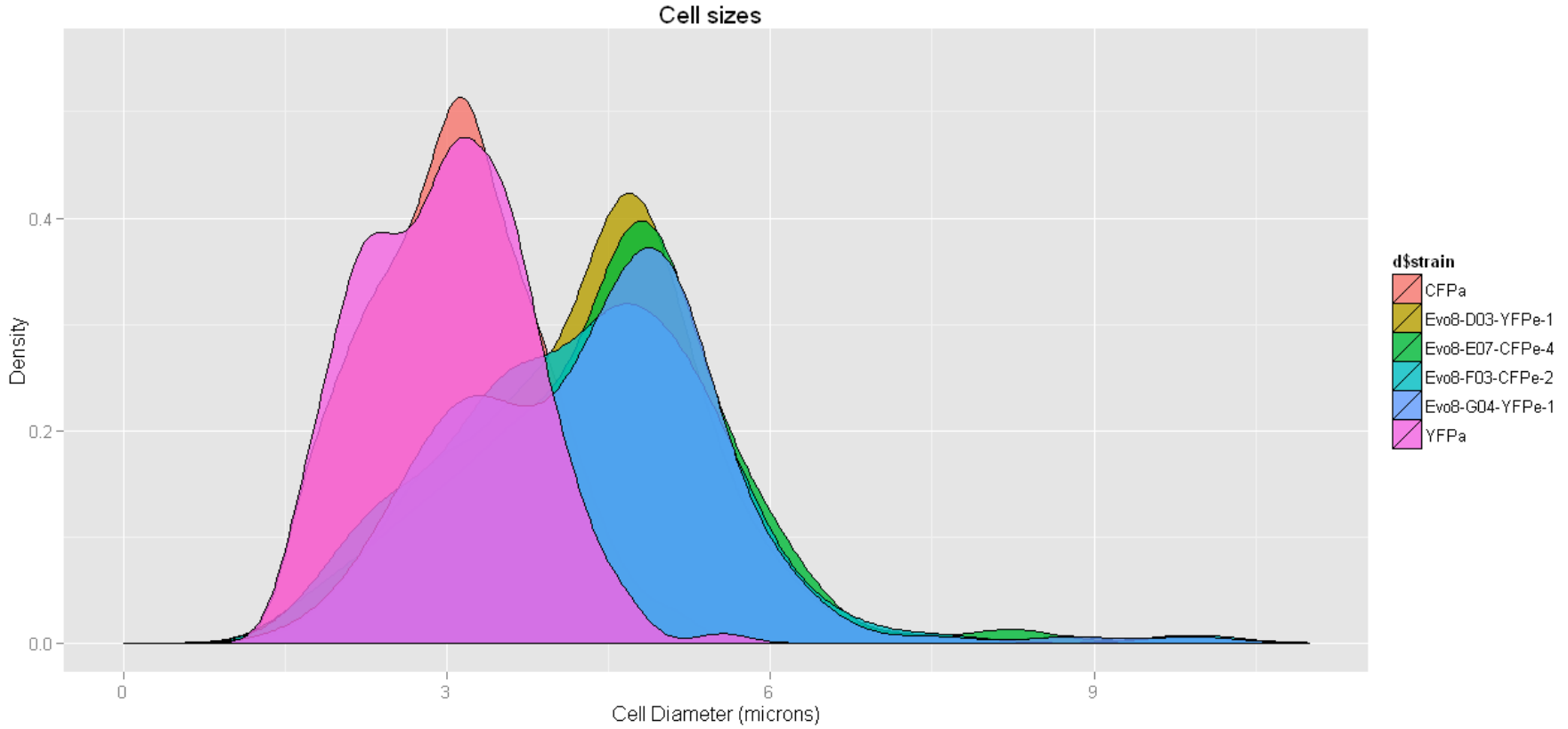


Figure 6.16 Density plot of cell diameter (microns) for ancestral strains and evolution8 strains Evo8_14_D03_YFPe_1, Evo8_14_E07_CFPe_4, Evo8_14_F03_CFPe_2, Evo8_14_G04_YFPe_1.

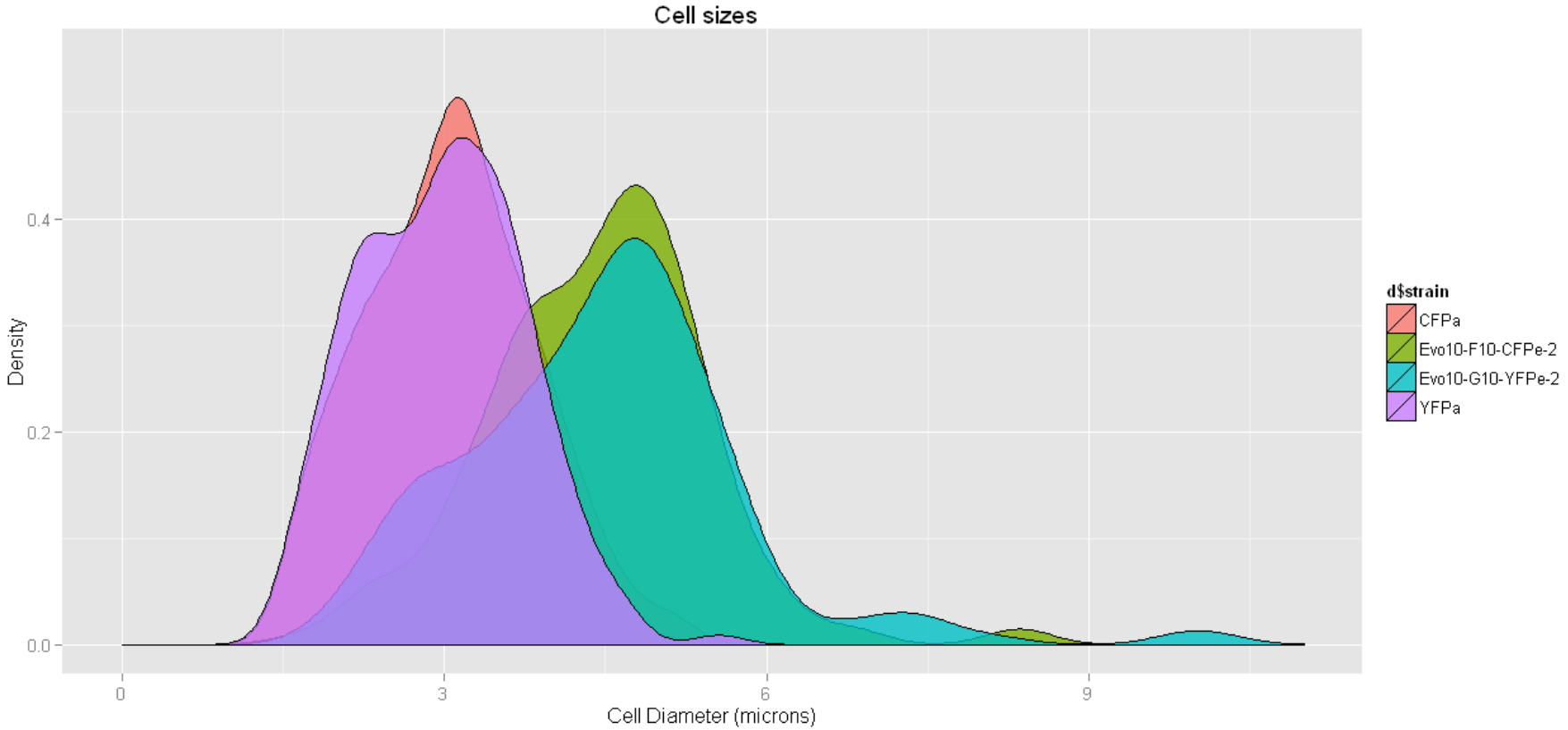


Figure 6.17 Density plot of cell diameter (microns) for ancestral strains and evolution10 strains Evo10_20_F10_CFPe_2, Evo10_20_G10_YFPe_2 in fluorescence cytometry run 1.

The second fluorescence cytometry run included three strains derived from the final plates of evolution 6 and 10. Evo6_11_B06_YFPe_1 had no demonstrably increased fitness compared to the ancestors and has a mean cell diameter of 3.36 microns close to that of the haploid ancestor (see Table 6.3), while Evo10_20_B03_YFPe_1 has a mean diameter of 4.51microns, the larger size characteristic of diploids and Evo10_20_B03_CFPe_5 was larger still with a mean cell size of 5.75microns.

Strain	**Mean Size (microns)**
YFPa	3.67
Evo6_11_B06_YFPe_1	3.36
Evo10_20_B03_YFPe_1	4.51
Evo10_20_B03_CFPe_5	5.75

Table 6.3 Mean cell diameter in microns for strains run in fluorescence cytometry run 2 (part I).

Density plots of these strains (see Figure 6.18 below) showed a good correspondence between Evo6_11_B06_YFPe_1 and its haploid ancestor with well separated and equidistant peaks for the two larger strains.

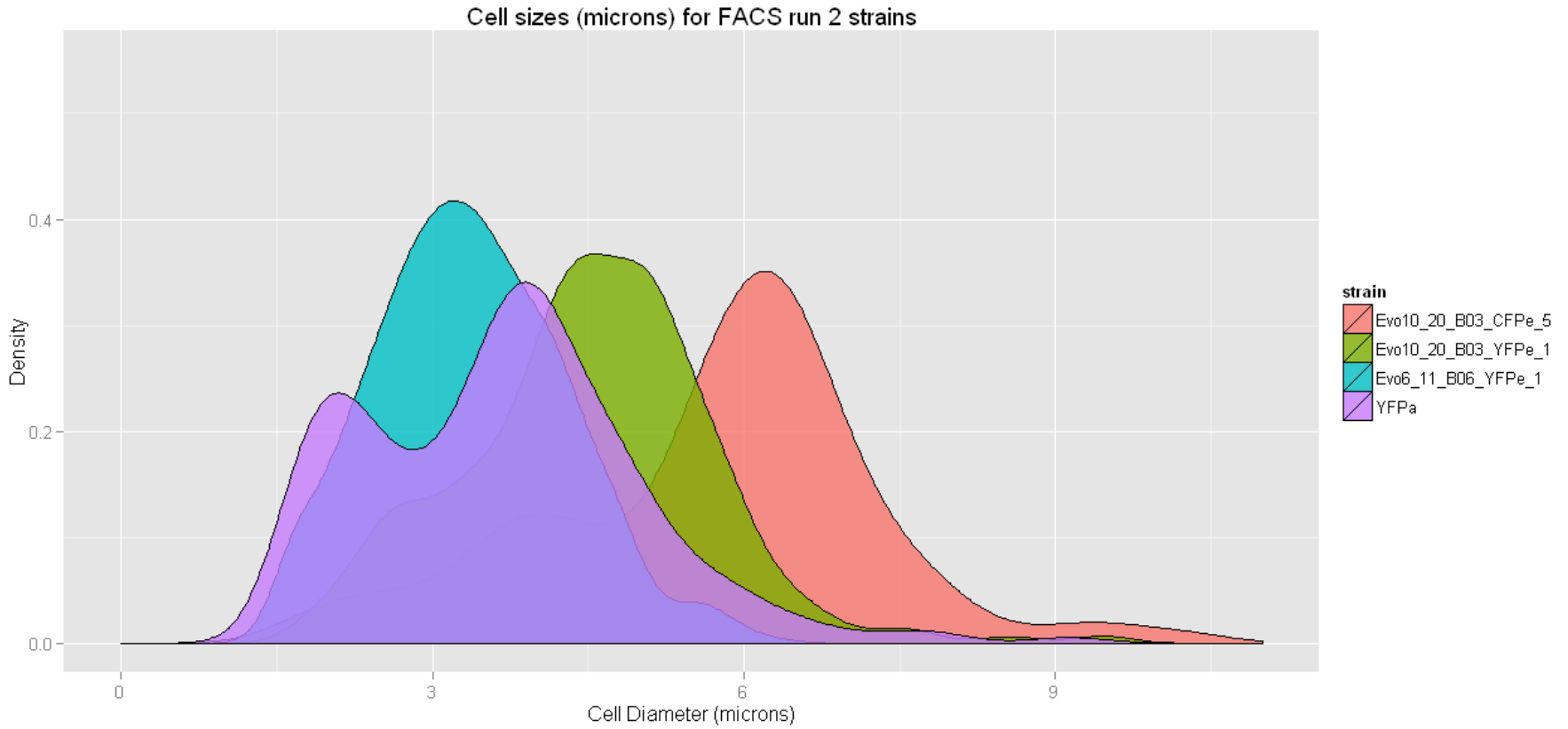


Figure 6.18 Density plot of mean cell diameter (microns) for the ancestral YFP strain (YFPa) and strains from evolution 6 (Evo6_11_B06_YFPe_1) and evolution 10 (Evo10_20_B03_CFPe_5, Evo10_20_B03_CFPe_5) in fluorescence cytometry run 2.

Sytox Green DNA Fluorescence cytometry of these strains confirmed that Evo6_11_B06_YFPe_1 had a DNA content profile not discernibly different from the haploid ancestors (see Figure 6.18) while Evo10_20_B03_YFPe_1 had a little more than double the DNA content with its major peak at an intensity of 68 compared with 24 for the YFP starting strain. The mode of the major peak of Evo10_20_B03_CFPe_5 is 104 units, very close to a 50% increase of the 68 unit diploid value and consistent with triploidy, but substantially less than the doubling of DNA required for a tetraploid. This in turn suggests that for this strain each increase in ploidy adds ~1.2 microns to the mean cell diameter.

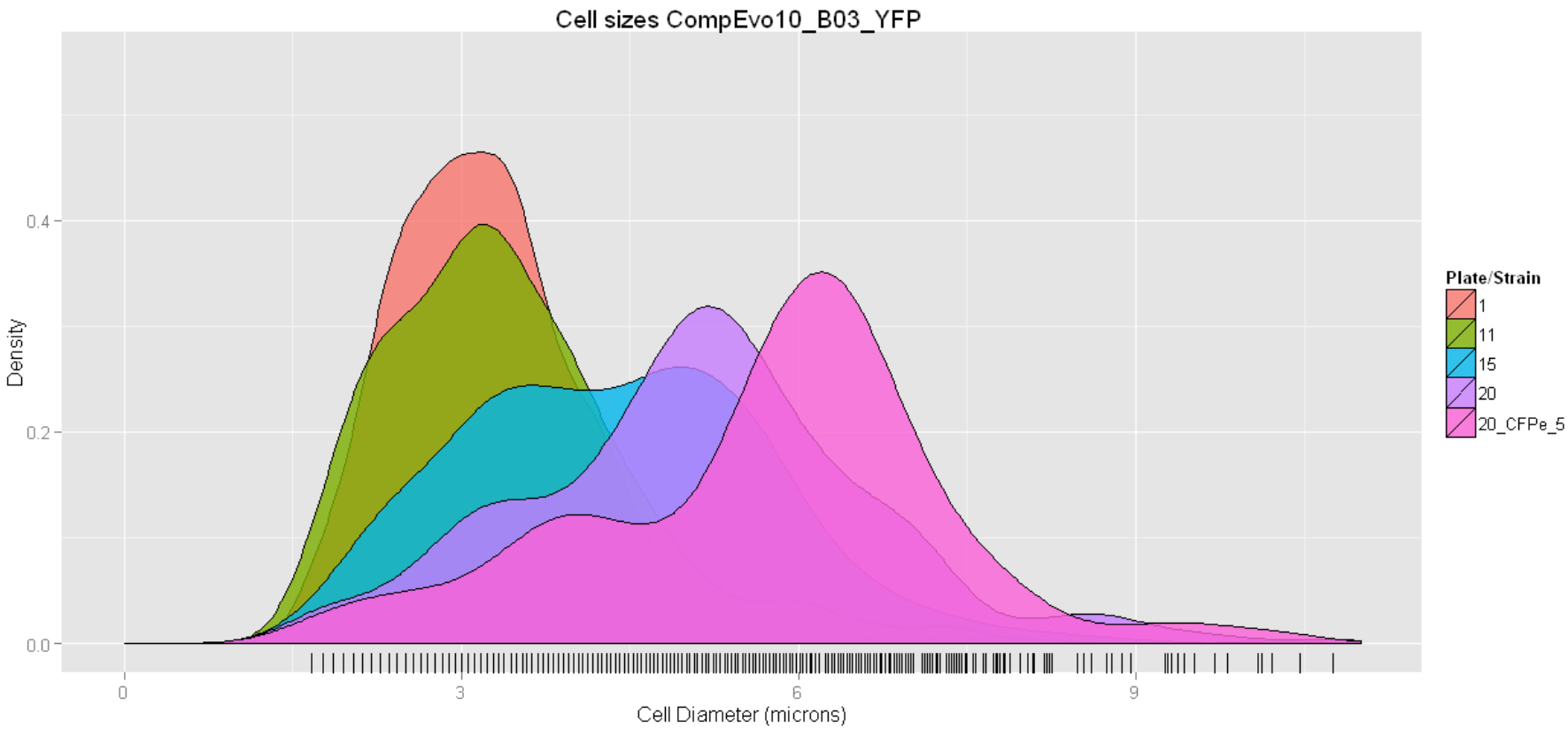
The evolution 10 populations of the same well (B03) as the putative triploid strain Evo10_20_B03_CFPe_5 were grown up for cell size measurement and Sytox Green DNA fluorescence cytometry. Table 6.4 shows plates 1 and 11 having mean cell diameters very similar to the starting haploid cultures. However, plates 15 and 20 show a significant increase in cell size and the Evo10_20_B03_CFPe_5 strain isolated from the plate 20 population is larger still.

Strain	**Mean Size (microns)**
YFPa	3.67
Evo10_B03_Plate_1	3.26
Evo10_B03_Plate_11	3.45
Evo10_B03_Plate_15	4.38
Evo10_B03_Plate_20	5.13
Evo10_20_B03_CFPe_5	5.75

Table 6.4 Mean cell diameter in microns for strains run in fluorescence cytometry run 2 (part II).

Density plots of the populations in two evolution 10 wells B03 and G10 (see Figure 6.19 a) and b) respectively) show the mode of the cell size profile spreading rightwards to larger sizes by plate 15 with the population becoming progressively more negatively skewed as it does so. The population in well B03 clearly converges on a larger modal cell size than G10. The population mode of well G10 is close to that of the sequenced diploids whilst that of well B03 approaches the probable triploid strain Evo10_20_B03_CFPe_5. This shows clear diversity in the well populations of evolution 10.

a) well B03


Cell sizes CompEvo10_B03_YFP
Density
0.0
0.2
0.4
0
3
6
9
Cell Diameter (microns)
Plate/Strain
1
11
15
20
20_CFPe_5

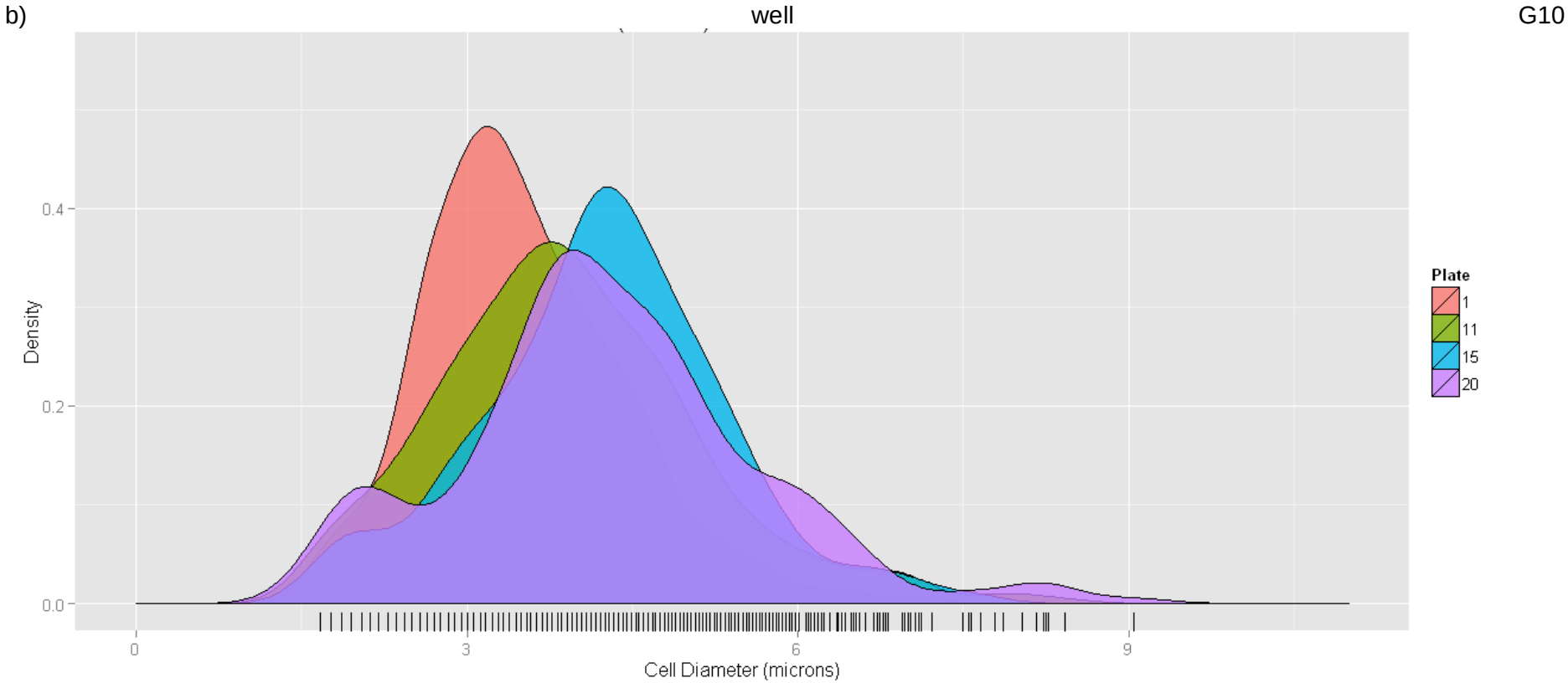
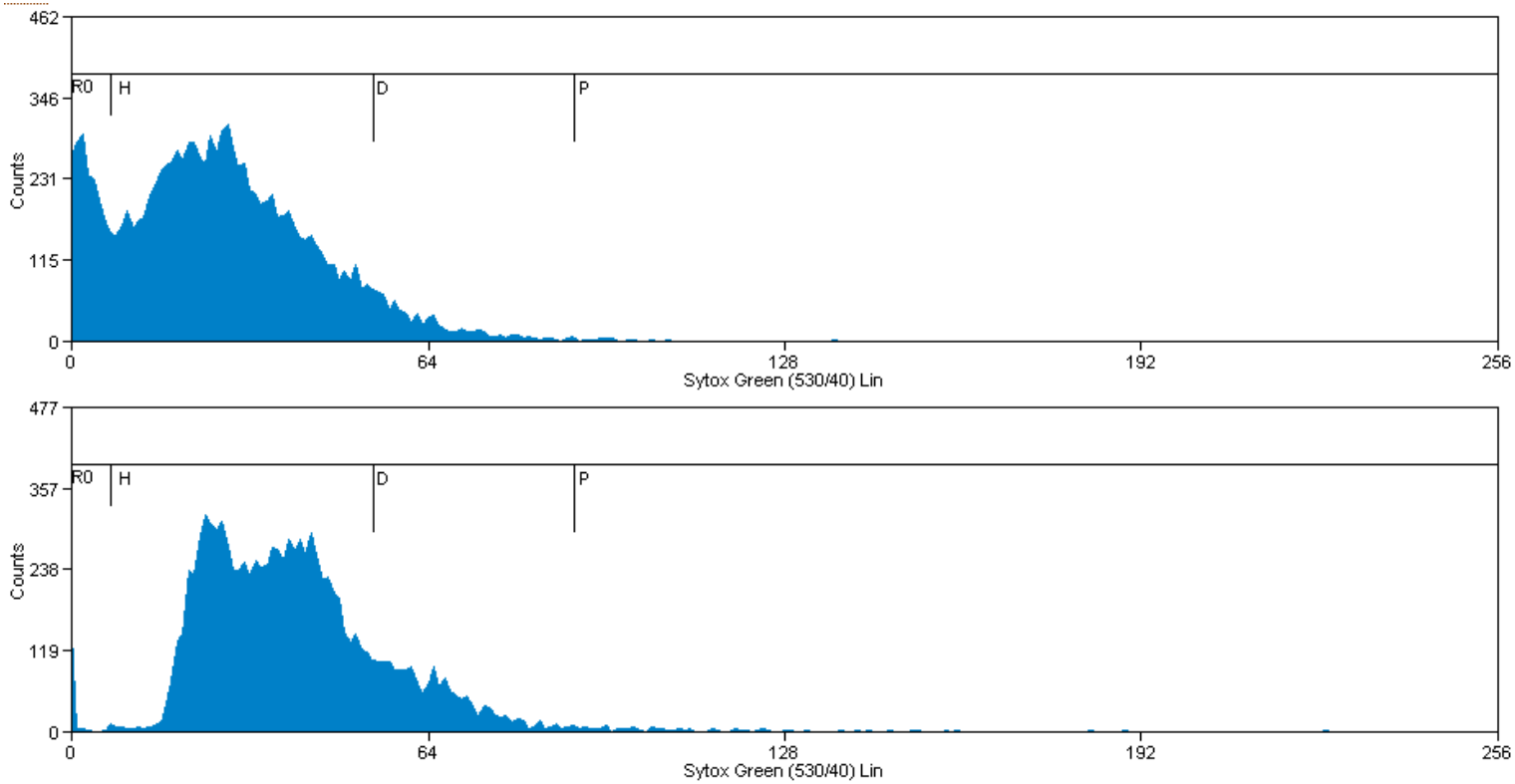


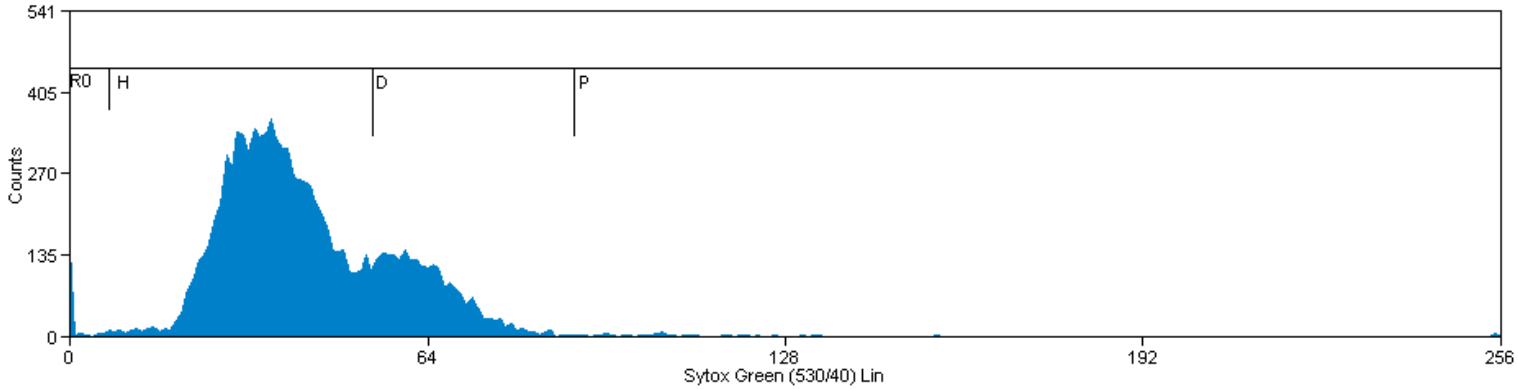
Figure 6.19 Cell size density plots for two wells in evolution 10: a) well B03 b) well G10.
Each plot shows the data for plates 1,11,15 and 20 and the well B03 plot also has the cell size distribution for the isolated strain Evo10_20_B03_CFPe_5.

DNA fluorescence cytometry of the evolutions 10 B03 population (see Figure 6.19.a.) is consistent with the strong and well-established correlation between size and ploidy; the populations in plates 1 and 11 look very similar to the haploid ancestors, with a major DNA peak at 30 units (close to the haploid values of 24 and 28 units) and (diploid post-replicative) minor peaks at 61 and 55 units respectively. However, the plate 15 population (Figure 6.20, histogram 8) is substantially more complex with major DNA intensity peaks at 32 and 96 units with a significant peak at approximately 63 units. This suggests that the plate 15 population has haploid, diploid and triploid subpopulations. However in plate 20 (Figure 6.20, histogram 9), the haploid population signal has almost disappeared and the diploid peak is barely discernible in the negative skew of a peak with a mode at 97 units. This is close to the 104 unit peak of the putative triploid Evo10_20_B03_CFPe_5. This strongly suggests that the triploid subpopulation has out-competed the haploid and diploid subpopulations and the final population is largely triploid.

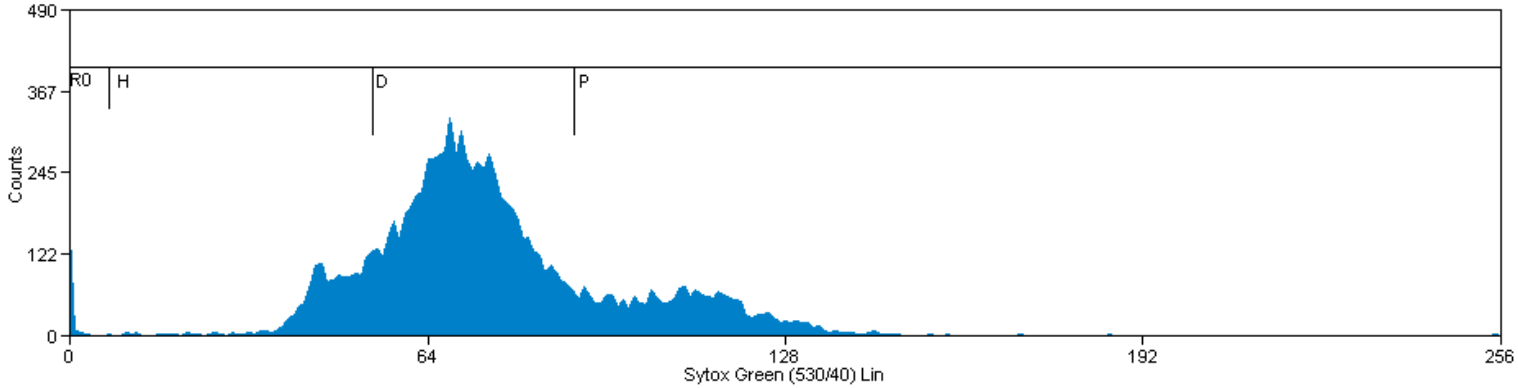


1. CFPa
Mode of DNA intensity= 28 (mode of minor peak=54)

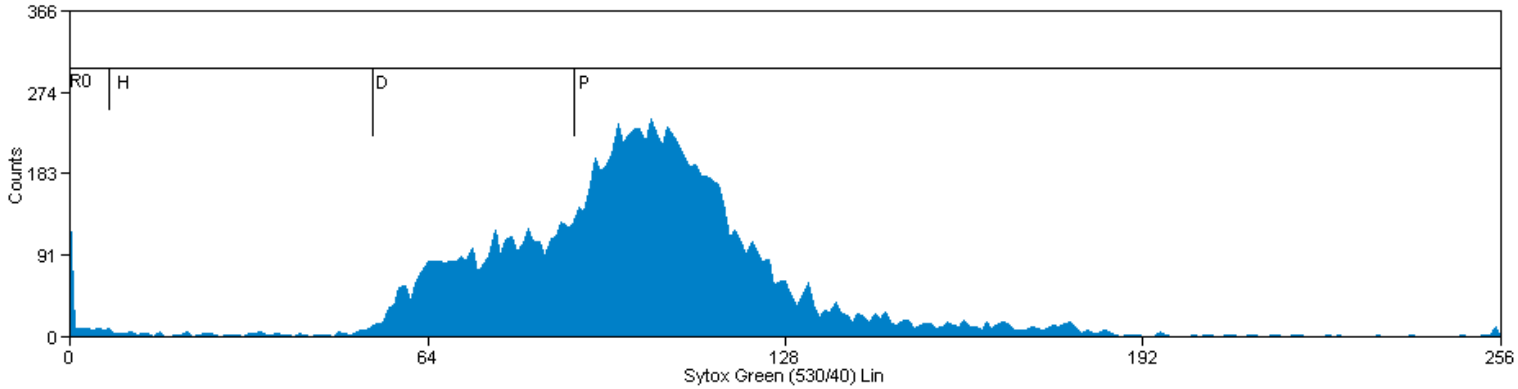
2. YFPa
Mode of DNA intensity= 24 (mode of minor peak=54)



3. Evo6_11_B06_YFPe_1
Mode of DNA intensity= 36 (mode of minor peak=60)

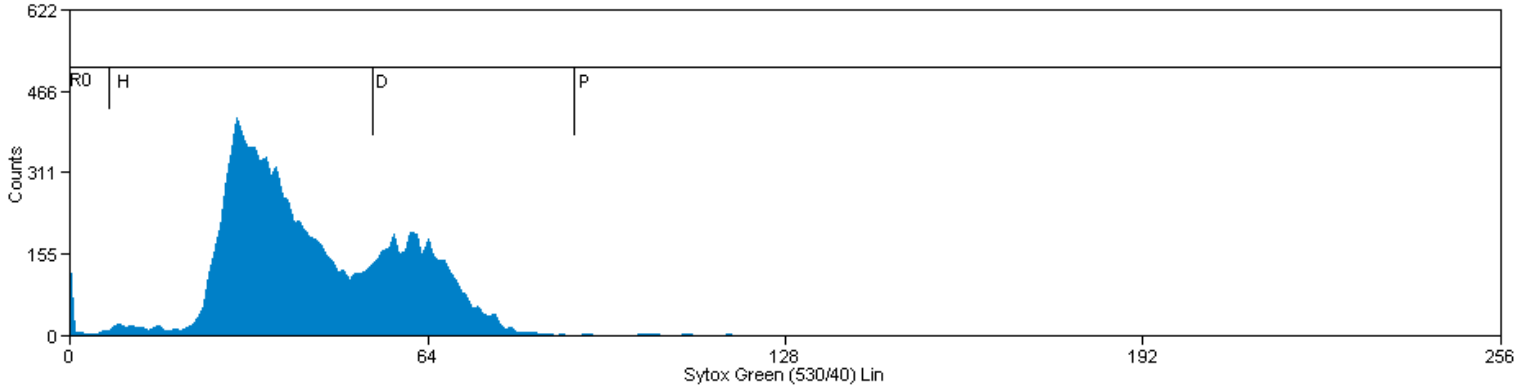


4. Evo10_20_B03_YFPe_1
Mode of DNA intensity= 68 (mode of minor peak=92)



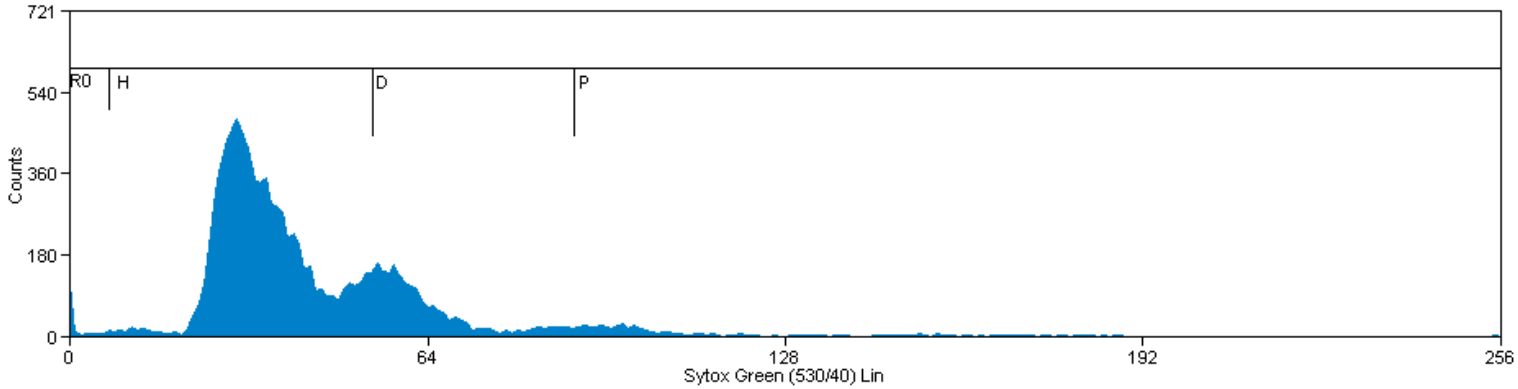
5. Evo10_20_B03_CFPe_5

Mode of DNA intensity= 104 (mode of minor diploid peak=88)



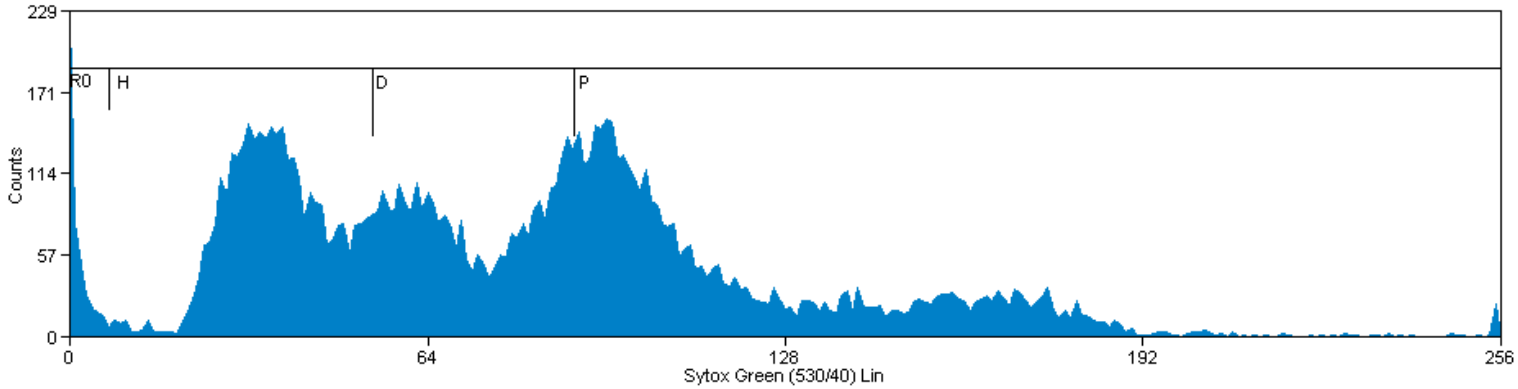
6. Evo10_B03 Plate1 Population

Mode of DNA intensity= 30 (mode of minor peak=61)



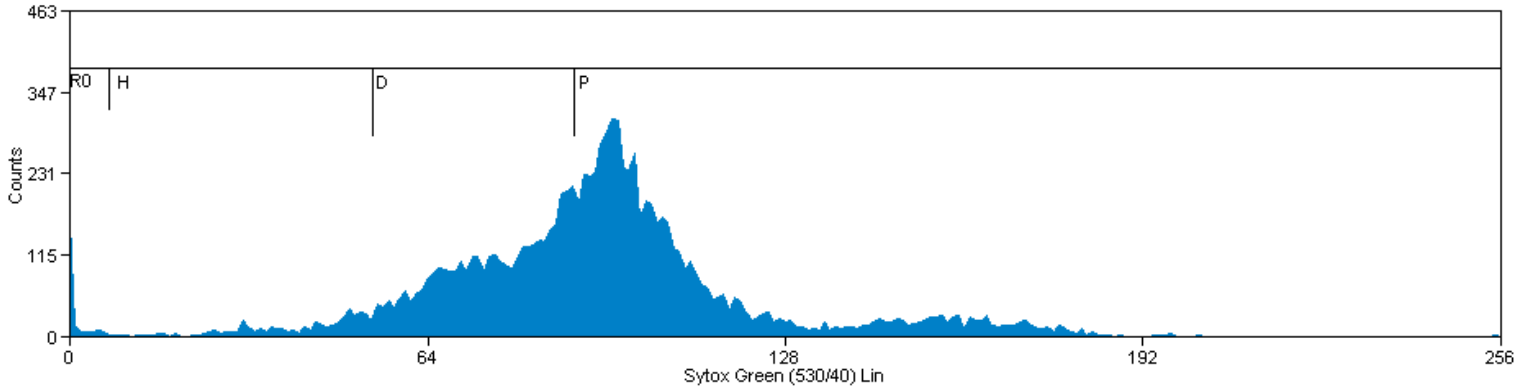
7. Evo10_B03 Plate11 Population

Mode of DNA intensity= 30 (mode of minor peak=55)



8. Evo10_B03 Plate15 Population

Peaks have modes at DNA intensities of 32,63, and 96



9. Evo10_B03 Plate20 Population

Mode of DNA intensity= 97

Figure 6.20 Histogram of DNA intensity for fluorescence cytometry Run 2.
In addition to ancestral controls and Evo6_11_B06_YFPe_1 well B03 samples were taken from four plates during evolution 10, plates 1, 5, 11, 15, 20.

Figure 6.21a is a ternary plot of the partitioned histogram of the evolution 10 well B03 populations seen in Figure 6.20 (histograms 6-9). This plot clearly shows the pattern seen in the DNA fluorescence intensity histograms (Figure 6.20) and cell size density plots (Figure 6.19); little change in ploidy for the first eleven plates, a mixed population of all three ploidies for plate 15 (which has point close to the centre of the ternary plot) and a predominantly polyploid population in plate 20 and in the isolated triploid strain Evo10_20_B03_CFPe_5. Figure 6.21b is a ternary plot of well G10 in the same evolution. The ploidy trajectory on the ternary phase plane is substantially less steep and the ploidies of the final populations are largely diploid. These ternary plots correlate well with the population cell sizes for these wells (see Figure 6.19) which show the well G10 population converging on diploidy while the B03 population appears substantially triploid.

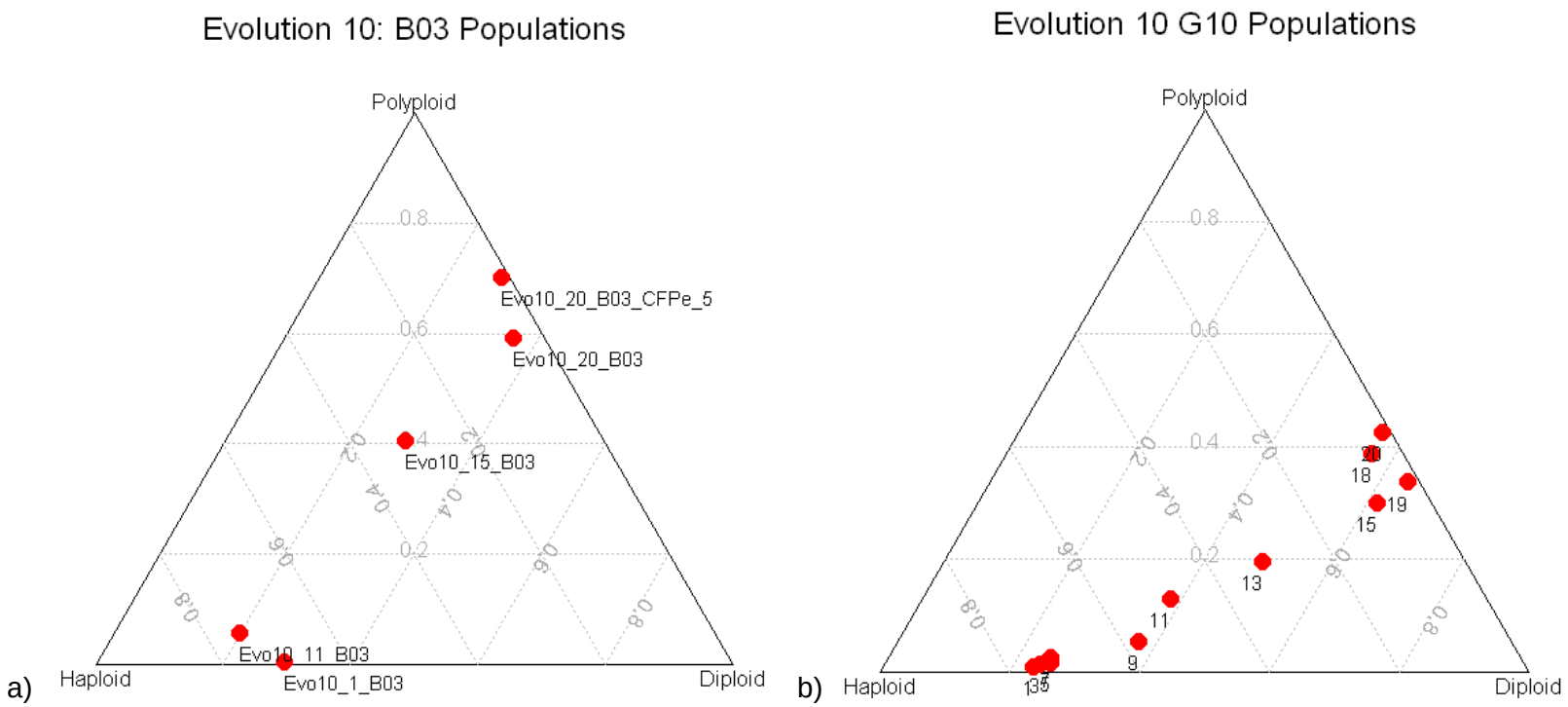
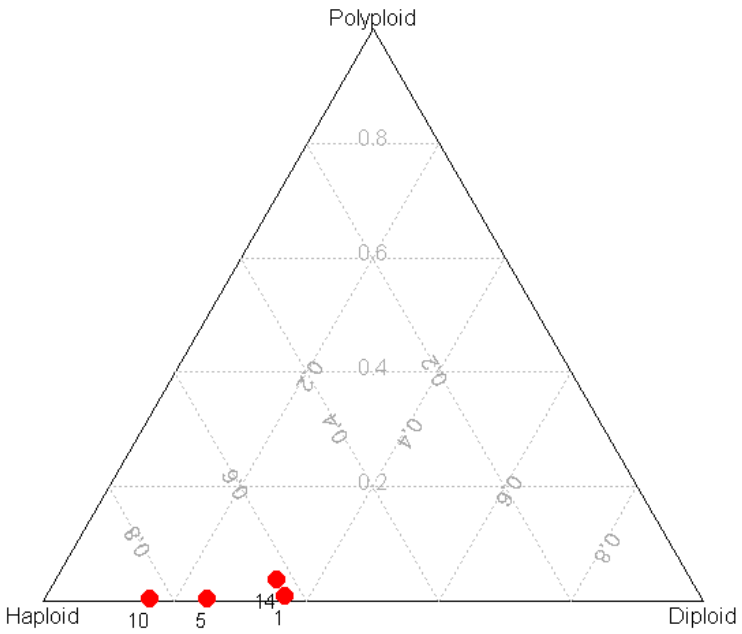


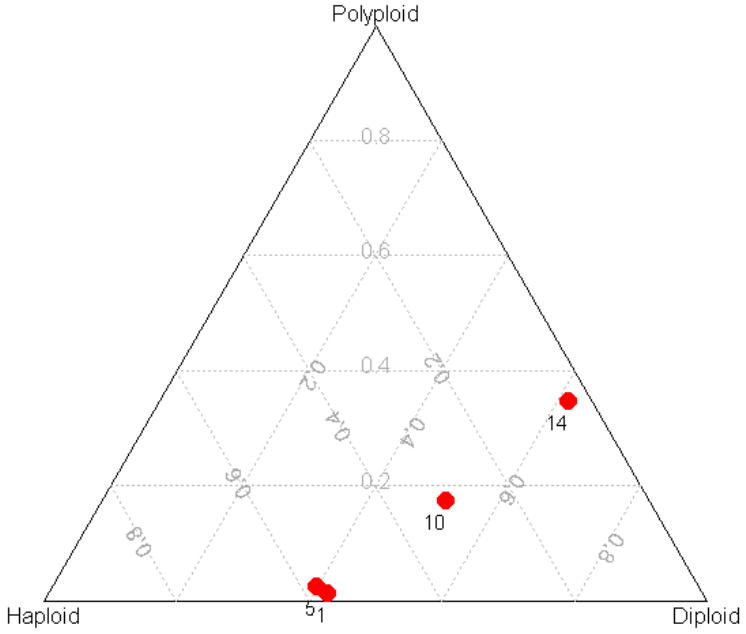
Figure 6.21 A ternary plot of Sytox Green DNA fluorescence cytometry of populations from evolution 10 wells B03 (a) and G10 (b).
In this evolution the ethanol stress ramped by 0.5% per plate to 9%. In a) B03 populations from plates 1, 11, 15 and 20 (the final plate) are shown together with the plate 20 isolate Evo10_20_B03_CFPe_5. b) G10 populations from plates 1, 3, 5, 9, 11, 13, 15, 18, 19 and 20

Ternary plots of the change in DNA fluorescence cytometry for evolution 8 populations also show diversity between wells. Figure 6.22a shows that the population of well B10 has remained haploid over the fourteen plates of evolution 8, with the point for plate 14 very close to that of plate 1. However, after the fifth plate the well E07 population (see Figure 6.22b) makes a linear transition across the ternary plot phase plane culminating in a population that appears substantially diploid. Both of these wells received no added ethanol treatment.

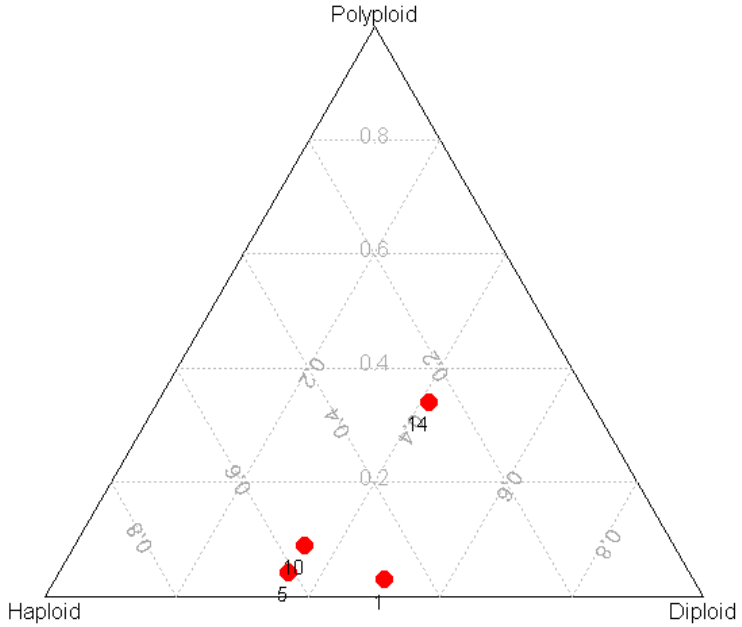
The wells G04 and D03 received a 6.5% added ethanol treatment (see Figure 6.22d and e) and both show a change to a population with a substantial diploid component. The well F03 population received 4.5% added ethanol treatment and Figure 6.22c shows its change in ploidy is ambiguous, with little change for the first ten plates and then a plate 14 population with a somewhat higher polyploid component.



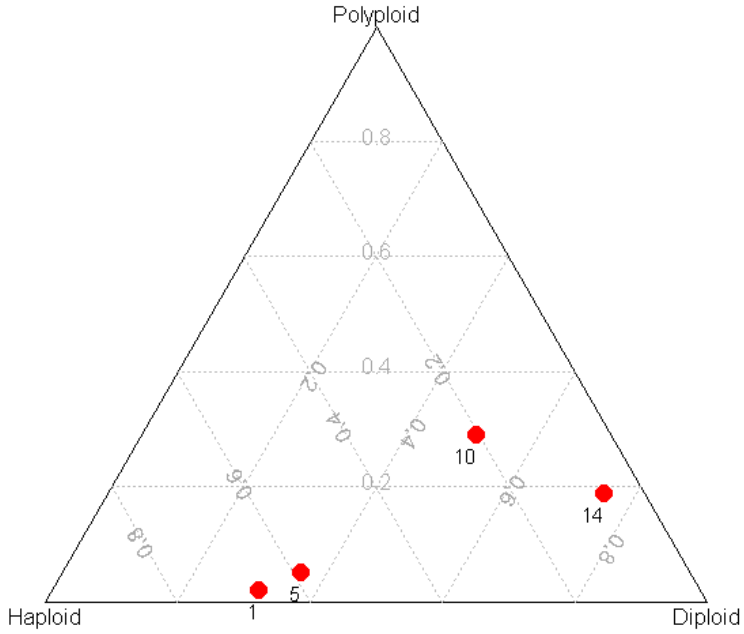
a) Well B10, 0% added ethanol



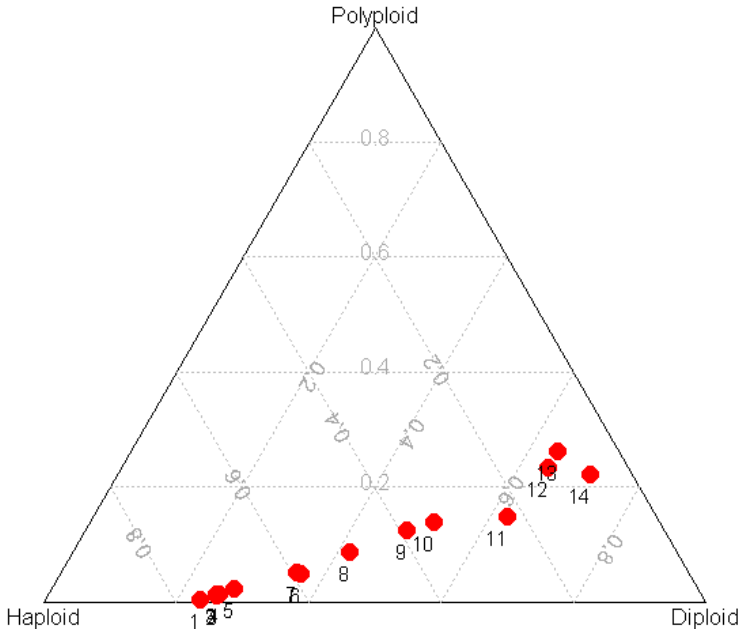
b) Well E07, 0% added ethanol



b) Well F03, 4.5% added ethanol



d) Well G04, 6.5% added ethanol



e) Well D03, 6.5% added ethanol

Figure 6.22 Ternary plots of DNA fluorescence cytometry counts binned by ploidy for populations in different wells of evolution throughout the course of evolution 8.

Numbers next to points identify the plate the population was taken from. Figs.a) well B10, and b) well E07 had no added ethanol treatment, c) well F03 had 4.5% added ethanol, and d) well G04 and e) well D03 had 6.5% added ethanol.

Figure 6.23 shows the mean cell sizes for the ternary plotted populations in Figure 6.22. The evolution 8 populations in Figure 6.23.a) all show an increase in cell size with plate number. Linear best fits of the E07 (0% added ethanol), 4.5% and 6.5% populations have very close intercepts and show similar significant gradients.

Ethanol	Well	Slope	Significance
0%	B10	0.059596	0.021*
0%	E07	0.09212	0.018*
4.5%	F03	0.06984	0.019*
6.5%	D03	0.09324	4.80e-06***
6.5%	G04	0.06803	0.021*

Table 6.5 Slope of linear best fit for the mean cell sizes for the ternary plotted populations in Fig. 5.22. showing the ethanol treatment of each well and the probability that there is a non-zero slope.

The well B10 population, which displayed little or no change in ploidy in Figure 6.22.a) has a very similar slope to the 4.5% treatment well F03 although B10 has an intercept approximately 0.3 of a micron lower. A strain was sequenced from the E07 population because it showed a clear change in fluorophore ratio however, the fluorophore ratio of the B10 population, in common with most populations without added ethanol, did not. The apparent increase in cell size in B10, a control well with no added ethanol and no increase in ploidy (see Figure 6.22a) suggests that there was an improvement in strain growth, possibly due to sequence change.

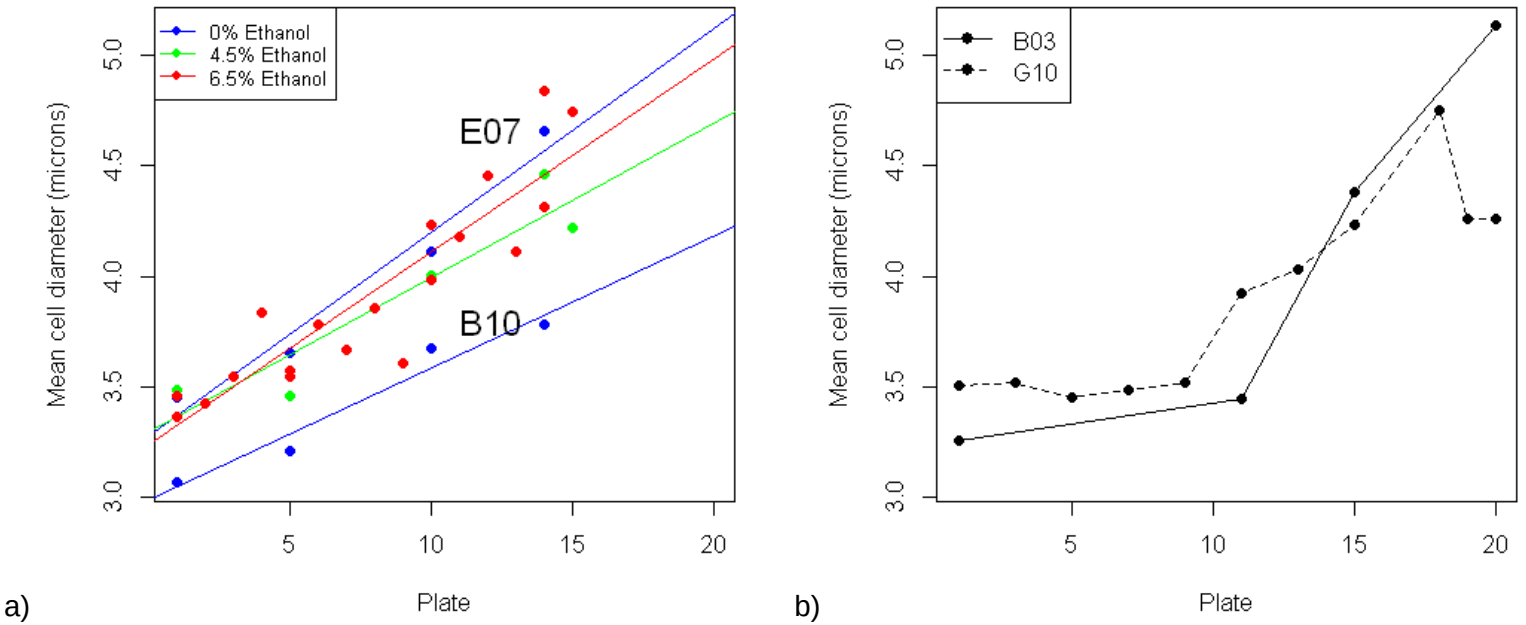
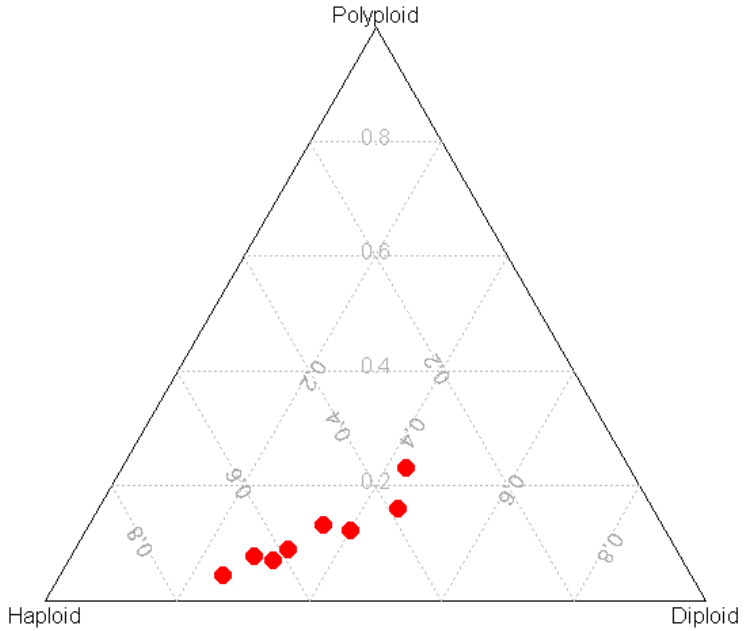


Figure 6.23. Mean cell diameter against plate for populations from evolution 8 and 10.

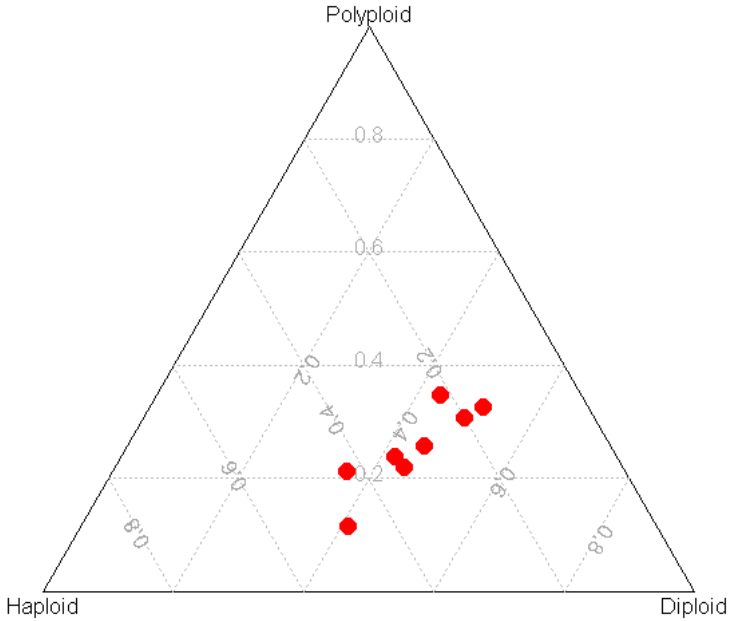
a) Mean cell diameter of evolution 8 populations from wells: B10, E07 (0% added ethanol treatment - blue points and best fit lines); F03 (4.5% added ethanol treatment - green points and best fit line); G04 and D03 (6.5% added ethanol treatment - red points and best fit line). Figure 6.23.b) Mean cell diameter of two evolution 10 populations from wells B03 and G10.

Figure 6.23.b) shows the evolution 10 cell size change in wells B03 and G10. In contrast to the linear increase see in evolution 8, there is little or no increase in cell size for the first 9-11 plates of evolution 10, until the increasing ethanol concentration exceeds 3.5%-4.5%. This is consistent with the little change in population ploidy (see Figure 6.21). Thereafter both B03 and G10 show a steep rise in mean cell size. The largely triploid B03 population has a final mean cell diameter of 5.13 microns while the largely diploid G10 population is 4.25 microns, although this difference in cell size is not statistically significant.

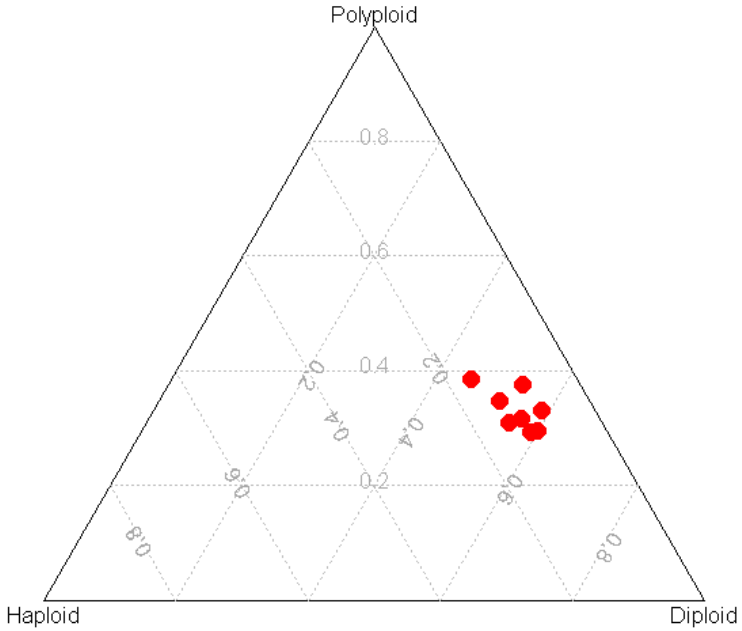
Figure 6.24.a)-c) show ternary plots of several populations from plate 14 of evolution 8 subjected to different ethanol treatments, while Figure 6.24d) shows populations from plate 15 of evolution 10. The three evolution 8 figures are subject to increasing added ethanol treatment. Figure 6.24.a) with no added ethanol, has some wells with the haploid ploidy characteristic of the starting strains but others with a significant diploid component which are spread out and appear to describe a linear transition across the ternary phase plane. Figure 6.24.b) with 4.5% added ethanol, has no sampled populations with solely haploid characteristics and has a larger diploid component spread out along a similar line in the phase plane to Figure 6.24.a). The populations subjected to 6.5% ethanol in Figure 6.24.c) are substantially diploid and found in a tight cluster. The evolution 10 populations in Figure 6.24.d) are from plate 15 and have been subjected to an ethanol stress that has increased 0.5% per plate to 6.5%. This figure has a cluster of largely diploid populations very similar to Figure 6.24.c) however it also has a populations with significant triploid content.



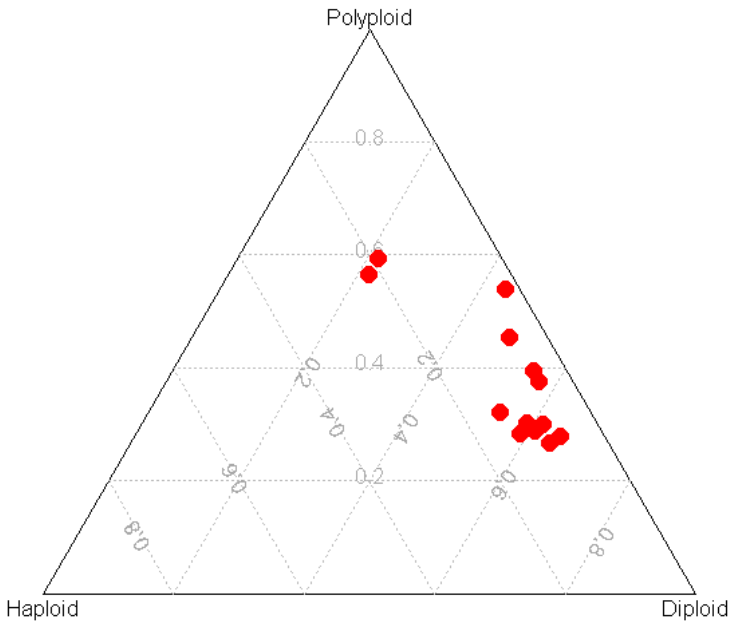
a) Evolution 8, 0% added ethanol



b) Evolution 8, 4.5% added ethanol



c) Evolution 8, 6.5% added ethanol



d) Evolution 10, ethanol ramped to 9%

Figure 6.24 Ternary plots of DNA fluorescence cytometry counts binned by ploidy for populations in different wells of evolutions 8 and 10.
Fig.a) shows evolution 8 populations with no added ethanol treatment, b) and c) show evolution 8 populations with 4.5% and 6.5% added ethanol respectively. All evolution 8 populations were derived from plate 14. Fig. d) shows the ploidy ratios of populations from evolution 10 plate 15, with 6.5% added ethanol concentration.

6.4 Discussion

6.4.1. Summary

A key research aim (see Section 2.7 “Research aims and goals”) was the identification of genotypic changes in the experimentally evolved strains that have demonstrably increased fitness under ethanol stress (see Chapters 4 & 5). No repeated genetic sequence changes were found in independently evolving strains. However, all of the isolated strains shown to out-compete their ancestors when growing with an ethanol stress had increased their ploidy. Such ploidy increase has been reported with salt stress in a similar multiply auxotrophic haploid background to these experiments, in a paper entitled “Genomic convergence toward diploidy in *Saccharomyces cerevisiae*” (Gerstein et al. 2006) seed discussion in chapter 2, section 2.6. Over 1800 generations Gerstein et al. showed convergence to diploidy from haploid and tetraploid ancestors. However, when a salt stress applied haploids became diploid more rapidly and the stressed tetraploids largely became triploid. In the experimental evolutions described in this chapter, an increase in ploidy from haploid to diploid was found within 200 generations. This more rapid transition to increased ploidy in these experimental evolutions may be due to differences in the stress applied or the starting strains which may carry particular auxotrophies or mutations which might further predispose the ancestral strains to adapt by increasing ploidy. One diploid strain was isolated from a control well with no added ethanol in evolution 8 (well E07) within this number of generations so transition to diploidy is not exclusively a feature of stressed cells. However (see Chapter 4, Fig. 4.26) changes in fluorophore ratio indicative of adaptive change (and likely ploidy increase) occur later in wells without added ethanol but in the overnight culture prior to preparation of the first plate with 6.5% ethanol suggesting that ethanol greatly increases the rate of ploidy increase. There is also support for this in Fig 6.24 in which ternary plots of population ploidy show a more rapid transition to diploidy (and for triploidy for evolution 10) as ethanol concentration increases. This increased rate of transition to higher ploidy with ethanol stress does not appear to have been previously documented. Either the rate of production of (largely endoreduplicated) higher ploidy strains increases with ethanol stress or there may be a larger fitness advantage for higher ploidy with ethanol stress with makes such strains sweep faster.

Strains of yeast may adapt in different ways due to small differences in their genotype. An anomalous or incomplete genotype may radically alter the accessible genetic landscape for evolution and prevent generalisation from the specific results

of these experiments to *S. cerevisiae* generally. Accordingly it is important to scrutinise the genotype of the starting strains and consider whether divergence from standard strains may predispose them to particular adaptive outcomes. Genotypic differences between each ancestral strain are also potentially problematic.

Phenotypic differences were observed between the ancestral strains and the strains that evolved from them in these experiments. The CFP ancestor is approximately 0.25 microns larger than the YFP. Furthermore, three out of the four evolved CFP strains evaluated also gain the capacity to grow on organic acids which do not support the growth of either ancestor, or of any YFP strain tested. This suggests that adaptive evolutions that start with the CFP ancestor may be more likely to gain an organic acid growth phenotype than those starting with YFP. However, the numbers of samples tested is small (see Table 6.6), too few for a chi-squared test, and a Fisher's exact test is not significant with a p-value = 0.1429.

	CFP	YFP
Acid tolerant	3	0
Non-acid tolerant	1	4

Table 6.6 Contingency table for fluorophore and development of organic acid tolerance in experimentally evolved strains.

6.4.2. Starting strain sequences

Aside from their documented multiple auxotrophies, which were verified using BreakDancer and inspection of the read depth at these loci, the starting strains in these experiments share several genetic features that differ from the S288c reference sequence. Both fluorescently tagged ancestors have minor deletions in *MEH1* and *APC1*, and two non-synonymous SNPs in *CTR9*/*CDP1*. These differences may be significant but could also be common strain variation that has no impact on the fitness of the strains or the trajectories of adaptation under ethanol stress.

The deletion in *MEH1* seen in all sequenced strains in this experiment removes a duplicated pair of amino acids, glutamine and histidine. There has been prior speculation that changes to *MEH1* could affect ethanol tolerance. Ethanol stress increases the permeability of the plasma membrane to H+ ions which leak into the cytoplasm and lower pH (see Introduction, section 2.1.3). Part of this cytoplasmic H+ is sequestered in vacuoles and the *MEH1* null mutant has a defect in vacuolar

acidification (Gao et al. 2005), which accordingly reduces tolerance of ethanol (Yoshikawa et al. 2009) and acetic acid (Kawahata et al. 2006) but also has several pleiotropic effects; reducing starvation tolerance (Davey et al. 2012) and increasing glycogen stores (Wilson et al. 2002), probably due to impaired autophagy and mobilisation of glycogen. Because of the ethanol sensitivity of its deletant, *MEH1* has been suggested (Yoshikawa et al. 2009) as a candidate gene for adaptive increases in ethanol tolerance.

However, the deletion of the glutamine-histidine repeat in *MEH1* is also found in the highly ethanol tolerant sake strain Kyokai No.7 (Akao et al. 2011), which has two additional substitutions relative to S288c not seen in the experimental evolution starting strains. The deletion is also seen in the reference sequences for *S. paradoxus*, *S. bayanus*, *S. mikatae, S. kudriavzevii* (Scannell et al. 2011) which are otherwise well conserved; *S. paradoxus* has 88% identity and *S. kudriavzevii* 76%. This suggests that this deletion is not likely to grossly impair the function of the *MEH1* protein.

Similarly, *APC1* is a component of the anaphase promoting complex/cyclosome (APC/C) required for separation of sister chromatids in anaphase (Peters 2006), which suggests that mutations in *APC1* could increase the rate at which cells with increased ploidy arise. In the *APC1* gene in the S288c reference strain there are two adjacent aspartic acid residues (residues 566 and 567) but every sequenced strain in these experiments has only one (see Figure 6.25):

```
S288C       (551) FLWQSAYSIL LSRANDDVVG GLKMEHDAFS LVLSLLILPI

Ancestors   (551) FLWQSAYSML LSRAND-VVG GLKMEHDAFS LVLSLLILPI
```

Figure 6.25 Amino acid sequence of APC1 (YNL172W) showing an alignment of the *S. cerevisiae* S288c and the ancestral CFP and YFP sequences.
The location of the aspartic acid residue where these strains differ is highlighted in yellow. All sequenced strains, both starting strains and those experimentally evolved, had a single aspartate at this locus.

However, the sake strain Kyokai No.7 (Akao et al. 2011) also has a single aspartate at this locus. This strain has a high ethanol tolerance. The well-conserved alignment of the *Saccharomyces* sensu-strictu for *APC1* shows an insertion of the repeated GAC codon at this position only in *S. cerevisiae* (see Figure 6.26), sequence data downloaded (http://wolfe.gen.tcd.ie/ygob) from the Yeast Gene Order Browser (Byrne & Wolfe 2005). This suggests that *APC1* is functional in the sequenced strains which lack the repeat.

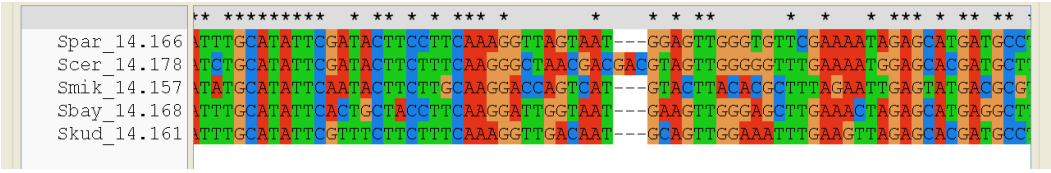


Figure 6.26 DNA sequence alignment of APC1 (YNL172W) *S. cerevisiae* (S288c) sequences and the homologous sequences in other members of the Saccharomyces sensu-strictu.
Alignment in Clustal X2.1. The aligned species are *S. paradoxus* (Spar_14.166), *S. mikatae* (Smik_14.157), *S. bayanus* (Sbay_14.168), and *S. kudriavzevii* (Skud_14.161).The GAC triplet is a repeated insertion in S288c relative to other members of the sensu-strictu. The sequence data is downloaded from the Yeast Gene Order Browser (http://wolfe.gen.tcd.ie/ygob).

CTR9 is involved in cyclin regulation (Koch et al. 1999) and transcription (Marton & Desiderio 2008). Growing populations of *CTR9* mutant strains accumulate large cells and the mutant is not viable at 37°C (Koch et al. 1999). Teixeira (Teixeira et al. 2009) but not Yoshikawa (Yoshikawa et al. 2009) suggested that *CTR9* might be a candidate ethanol tolerance gene due to poor growth of the null mutant under ethanol stress.

The ancestral CFP specific deletion is at the start of a repeat in *EGT2*. Confirmation of mutations seen in next generation sequencing by Sanger sequencing is always desirable but it is especially so in genes with repeats due to possible difficulties with uniquely mapping reads. *EGT2p* is one of three glucanases for separation of mother and budded daughter cells, the others being *DSE2p* and *DSE4p* (Colman-Lerner et al. 2001) and *EGT2* deletions show delayed bud separation although cytokinesis and is unaffected (Kovacech et al. 1996). As measured in the Nexcelom Cellometer, the ancestral CFP is 0.25 microns larger than the YFP ancestor from 0-8% ethanol (see chapter 5, “Evolved strain phenotypic analysis”, Figure 5.51). This cell enlargement could be linked to this *EGT2* deletion, although the null mutant phenotype shows typical cell growth in chains of three or four adherent cells (Kovacech et al. 1996; Yeong 2005), a pattern rarely seen in either CFP or YFP strains.

The *EGT2* null mutation has reduced starvation tolerance (Davey et al. 2012) and decreased competitive fitness (Breslow et al. 2008). However, *EGT2Δ* also has enhanced tolerance of lactic, acetic and hydrochloric acids (Kawahata et al. 2006). The Biolog Gen III has several wells testing growth on the organic acids: acetic acid, propionic acid, acetoacetic acid, and α-keto-butyric acid. Neither ancestral strain grew in these wells, suggesting that the CFP-specific *EGT2* deletion did not confer any substantial benefit in the haploid background. However, the only evolved strains to show substantial growth in organic acids were derived from the CFP ancestor exposed to a 4.5% ethanol stress: from evolution 6 there was Evo6_11_B06_CFPe_4, and from evolution 8 Evo8_14_F03_CFPe_2, and

Evo10_20_F10_CFPe_2. None of the four evolved YFP strains grew but neither did Evo8_14_E07_CFPe_4; CFP diploidy appears insufficient for good growth on organic acids. However, no other genetic changes likely to change phenotype, no non-synonymous SNPs, INDELS, copy number or structural changes were found in either Evo6_11_B06_CFPe_4 or Evo10_20_F10_CFPe_2 to explain their enhanced growth on organic acids. Evo10_20_F10_CFPe_2, is also larger than the other diploid strains isolated (see Chapter 5, Figure 5.54) with the exception of the other evolution 10 strain, Evo10_20_G10_YFPe_2, further implying that genetic change(s) remain undetected in this strain.

6.4.1. Genetic changes in evolved strains

In Evo8_14_F03_CFPe_2 sequence changes were found in two genes: *GPI18* and *GCN1*. Neither has been cited as involved in improving growth on organic acids. However, each of these genes has some limited literature support as part of a possible adaptive response to ethanol stress. A genome-wide analysis of expression changes during a fifteen day wine fermentation identified the essential gene *GPI18* as one of a group of up-regulated genes, although the maximum fold change was relatively modest at x5.38 and the group was a large one containing 223 genes (Marks et al. 2008).

The *GPI18* SNP replaces alanine with valine as the final amino acid residue. GPI18 is a mannosyltransferase that transfers the second mannose in glycosylphosphatidylinositol (GPI) biosynthesis (Kang et al. 2005; Fabre et al. 2005). GPI forms a lipid anchor for lipoproteins and is localized to the endoplasmic reticulum and mitochondria. *S. cerevisiae GPI18p* is a highly conserved protein, it is a functional homolog of the human *PIG-V* gene which is able to rescue the null *GPI18* mutant from lethality (Kang et al. 2005). A putative trans-membrane domain precedes the last two amino acids (Kang et al. 2005) which are relatively well conserved; in the human, rat and *D.melanogaster* the last two amino acids are tryptophan-threonine but the *Saccharomyces* sensu-strictu all terminate *GPI18* with proline-alanine. The substitution of alanine with valine is conservative and, as an essential gene, this substitution in *GPI18* must leave at least part of its function intact.

GCN1 is a non-essential gene and its deletion makes growth under ethanol stress poorer which has suggested that it might play a role in ethanol tolerance (Yoshikawa et al. 2009). However, *Gcn1p* is both large and multifunctional and the structure/function relationships of much of this protein remain unclear. What is

known, in addition to its putative role in ethanol tolerance, is that *GCN1* is a required part of the "general control response" to amino acid starvation and its appearance in an experimentally evolved multiple auxotroph is intriguing. A mutation in *GCN1* raises the prospect that the largest accessible fitness improvements with these starting strains in experimental evolution may be not in ethanol tolerance but in adapting to impaired amino acid synthesis and absorption.

The general control response is triggered by uncharged tRNA and a complex of *Gcn1p* and *Gcn20p* which activate the *Gcn2p* protein kinase (Garcia-Barrio et al. 2000), a process in which *Gcn1p* is required (Marton et al. 1993), possibly as a transporter of uncharged tRNAs from ribosomes (Wout et al. 2009). *Gcn2p* phosphorylates the alpha-subunit of translation initiation factor *eIF2* (*Sui2p*) which inhibits translational initiation (Kubota et al. 2001) but stimulates the translation of *GCN4*, a master gene regulator with 500-1000 targets (Hinnebusch & Natarajan 2002) including increased transcription of amino acid biosynthetic genes (Rolfes & Hinnebusch 1993). *GCN4* activation by starvation increases the replicative lifespan (Steffen et al. 2008). The minimum region required to bind *Gcn2p* is from amino acid residue 2064 to 2382 (Kubota et al. 2001) and the Evo8_14_F03_CFPe_2 heterozygous SNP is at position 1065 in the protein suggesting that any non-neutral effect of this mutation is unlikely to be in affecting binding to *Gcn2p* directly but possibly by a regulatory effect via the binding of other proteins.

This *GCN1* mutation is found in a single strain and has an unquantified phenotypic effect. However, it raises an intriguing possibility. In the methods used in these experimental evolutions there is a relatively short, consistent interval between the exhaustion of the medium in one plate and the provision of new media in the next plate. Accordingly it is tempting to speculate that one route to enhanced growth in this experimental system might be to modify the metabolic changes preparatory for withstanding prolonged starvation of uncertain duration. Accordingly, changing or delaying the starvation response, and possibly other changes that commit cells to enter quiescence, might be adaptive in these experiments and in those industrial processes that 'repitch' yeast cultures in batch brewing or bioethanol production.

For wild-type strains it is claimed that basal rates of endogenous amino acid synthesis prevent the general control response on minimal media; elicitation of the response on such media requires amino-acid analogues that inhibit biosynthesis (Albrecht et al. 1998). However, the ancestral strains in these evolutions are multiple auxotrophs unable to synthesise methionine, cysteine, leucine or histidine (in the

MATα/*MATα* background of Evo8_14_F03_CFPe_2) and also lack the lysine and arginine permeases. In such auxotrophs depleted intracellular amino acids may trigger the general control response more readily but with a poorer outcome since synthesis of some amino acids cannot be increased because the necessary genes are missing and this could result in persistently impaired translational initiation.

Shortage of purine (Rolfes & Hinnebusch 1993) or glucose (Yang et al. 2000) or salt stress (Goossens et al. 2001) also induce the *GCN2* mediated activation of *GCN4* by phosphorylation of *eIF2α*. Goosens postulated that over-activation of *GCN4* activation by salt stress might be toxic and found that loss of *GCN1* function reduced the deleterious effects of sodium stress (Goossens et al. 2001).

The presence of both a homozygous and a heterozygous SNP in Evo8_14_F03_CFPe_2 has a bearing on the likelihood of events in the divergence of the strain from the ancestor. The homozygosity of the *GPI18* SNP makes it improbable that diploidisation is the first event in the divergence of the strain from the ancestor. If the diploid was produced by endoreduplication, as its *MATα*/*MATα* genotype suggests (personal communication Bharat Rash). The likeliest sequence of genetic changes is for the homozygous *GPI18* SNP to be present before endoreduplication with the heterozygous *GCN1* SNP arising after it in one copy of chromosome VII.

The fluorophore ratio increased in favour of YFP in all the 13 evolution 8 populations with 6.5% added ethanol that showed non-trivial fluorophore ratio change. This was the only evolution and treatment where the ratio of prevailing fluorophores departed from chance ($p=2.4x10^{-4}$). Dirac curve fitting had similar low alpha fitness parameter values (mean 0.176 +/- 0.018 s.e.) and tau onset values with a negative mean value (-0.245 +/- 0.898 s.e., see Table 4.5), see Chapter 4, section 4.3.8. The two strains that were sequenced (Evo8_14_D03_YFPe_1 and Evo8_14_G04_YFPe_1) both had the same mitochondrial intron SNP in *COB1*. The fluorophore curves and sequencing data together suggest that the *COB1* mutation had swept from rare in the YFP population prior to the wells of the first plate in the 6.5% ethanol treatment of evolution 8 being filled. It is likely that the mutation arose during overnight culture of YFP for this evolution 8 treatment or in the YFP colony selected for culture. If the stock YFP population had a beneficial mutation sweeping, it is intriguing that three wells showed only trivial fluorophore ratio change; either the beneficial YFP mutation was stochastically lost from these three populations or a beneficial mutation of matching fitness size arose in the CFP population causing clonal interference.

Failure to splice introns at either *COB1* or *COX1* would cause respiratory deficiency characterised by failure to grow on respiratory carbon sources such as glycerol. However, the maturase SNP is a conservative substitution at position 360 from GTA (valine) to GGA (glycine) and it is claimed (http://www.uniprot.org/uniprot/P03879) that residues 385 to 638 are sufficient for maturase activity. Both Evo8_14_D03_YFPe_1 and Evo8_14_G04_YFPe_1 produce colonies of normal size, rather than the petite colonies produced by strains with defective mitochondria. Both also grow well on the non-fermentable carbon source glycerol (personal communication, Bharat Rash).

The copy number of mitochondrial genes is unclear and more copies of the mitochondrial genome are physically transferred than inheritance would suggest. Although there are estimated to be ~50-100 copies of the genome per yeast cell and ~40 copies are passed into the bud (Sena et al. 1976; Piskur 1994) the typical inherited number of mitochondrial genome copies is 3-4 and some mutations can come to be exclusively represented - i.e. mutations can be inherited as though the mitochondrion were haploid. This may be due to selection or intracellular genetic drift.

In addition to the homozygous mitochondrial SNP in *COB1* shared with Evo8_14_D03_YFPe_1, the Evo8_14_G04_YFPe_1 strain also has a non-synonymous, heterozygous SNP in the non-essential gene *PHO23*. The most likely sequence of events for the evolution 8, 6.5% ethanol treatment is that *COB1* arose in the haploid YFP culture before becoming diploid, either in a single event prior to being split between wells or in multiple events after. Subsequently the *PHO23* mutation occurred on one chromosome of the diploid after the overnight culture was split into wells on the first plate.

The carboxyl terminal region of *PHO23* has a conserved region very similar to the yeast genes *YNG1* (YOR064C), *YNG2* (YHR090C) and the human tumour suppressor *ING1* (R Loewith et al. 2000), containing a plant homeodomain (PHD) region involved in chromatin-mediated transcriptional regulation (Cell & Pfeffer 1995). However, the mutation at position 206 lies outside this region and the rest of the protein is not highly conserved. Null mutants in *PHO23* stimulate constitutive expression of *PHO5* encoding a cell surface acid phosphatase glycoprotein that scavenges phosphate from the environment and which is normally repressed in a high phosphate medium (Lau et al. 1998).

PHO23 is also required for the normal functioning of the *RPD3* histone deacetylase complex (HDAC) as a chromatin-mediated regulator of transcription (Loewith et al. 2001). There is a strong positive correlation between gene expression and histone acetylation (Hebbes & Thorne 1988) with deacetylation as a repressor of transcription (Grunstein 1997). *RPD3* is associated with the transcriptional silencing of telomeres and the cryptic mating-type cassettes (Sun & Hampsey 1999), null mutations of *PHO23* enhance silencing at these loci (Loewith et al. 2001). Acetylation is also involved in the general control response to amino acid starvation. Such starvation causes the targeted hyperacetylation of a small region surrounding the *HIS3* promoter (Kuo et al. 2000). A homozygous null mutation of *PHO23* therefore would stimulate phosphate uptake by acid phosphatase and synthesis of histidine in wild type strains. By extension a gain of function mutation might be expected to reduce these responses to starvation.

Glycogen degradation in yeast requires both glycogen phosphorylase (*GPH1p*) which can break α1-4 glucose bonds and a glycogen debranching enzyme like *GDB1p* to resolve α1-6 branches; deletion of *GDB1p* reduces the mobilisation of glycogen (Teste et al. 2000). In humans, deletion of the glycogen debranching enzyme causes type III glycogen storage disease (Cori's or Forbes disease) in which enlarged cells accumulate a highly branched form of glycogen that is largely inaccessible to metabolism (Chen 2011).

Genes involved in both the storage and mobilisation of glycogen (including *GDB1*) are induced late in exponential growth when the bulk of glycogen storage occurs (Teste et al. 2000), glycogen then falls during the slow growth phase (Lillie & Pringle 1980). Glycogen is also broken down prior to the resumption of growth when cells are transferred to fresh medium, a period in which the inhibitor *IGD1* is downregulated (Kresnowati et al. 2006). This lag phase breakdown of glycogen is disrupted by the deletion of *GDB1* (Teste et al. 2000).

The *GDB1* promoter has a stress responsive 'STRE' element -191 bp before the start of the open reading frame. This element is responsible for a 2 or 3 fold increase in expression with, respectively, a heat or osmotic shock (Teste et al. 2000). Similarly, after 30 minutes a 7% ethanol shock increases the mRNA expression of enzymes involved in both glycogen synthesis and degradation; *GPH1* and *GDB1* expression is increased by 4 and 4.2 fold respectively, increases thought likely to result in a net mobilisation of glucose from stored glycogen (Alexandre et al. 2001). *GDB1p* is bound by an inhibitor *IGD1p* (YFR017C) which is not thought to have a homologue

outside *Saccharomyces* (Walkey et al. 2011). Deletion of *GDB1* increases glycogen stores whilst deletion of the inhibitory *IGD1* reduces them (Walkey et al. 2011).

The *GDB1* protein is highly conserved in the *Saccharomyces* sensu strictu and beyond; the mammalian and yeast homologues have 39% identity including the cysteine residue changed by this SNP (see Figure 6.27).

```
487 AYLRREVIVWGDCVKLRYGKSPEDSPYLWERMSKYIEMNAKIFDGFRIDNCHSTPIHVGE 546 Yeast

479 VYLRRELICWGDSVKLRYGNKPEDCPYLWAHMKKYTEITATYFQGVRLDNCHSTPLHVAE 538 Dog

478 VYLRRELICWGDSVKLRYGNKPEDCPYLWAHMKKYTEITATYFQGVRLDNCHSTPLHVAE 537 Human

501 VYLRRELICWGDSVKLRYGTKPEDCPYLWAHMRKYTEIIATYFQGVRLDNCHSTPLHVAE 560 Rabbit

479 VYLRRELICWGDSVKLRYGNKPEDCPYLWAHMKKYTEITATHFQGVRLDNCHSTPIHVAE 538 Horse
```

Figure 6.27 An alignment of the highly conserved *S. cerevisiae* gene GDB1 (YPR184W) with mammalian homologues.
The Evo10_20_G10_YFPe_2 SNP changes the cysteine residue highlighted in yellow (amino acid 537) to serine. The aligned data is derived from Uniprot (Uniprot accession codes in brackets): yeast (Q06625), dog (Q2PQH8), human (P35573), rabbit (P35574), horse (A8BQB4).

Uniprot suggests that residues 535 and 538 are part of an active site. If *GDB1* activity is impaired, it is likely to enlarge cells and this might have a physiological benefit for ethanol tolerance. However, it is difficult to see that such enlargement would outweigh the problematic inaccessibility of glycogen stores. If the SNP reduces binding of the *GDB1* inhibitor then this could reduce stored glycogen, this may have growth benefits in a system where the waiting time for fresh media is relatively short, a matter of hours in these experiments, rather than the weeks and months that may occur naturally.

Sequence coverage was in excess of x100 across the genome (see this chapter, section 6.3.1) but locally poor coverage may prevent detecting variation against the background noise of Next Generation sequencing error. SAMTools and Pindel variant calling did not always identify the ancestral genetic variations from the reference in all strains, although they were apparent when the mapped read pile-ups were viewed in IGViewer or SAMTools view. The discrepancy between phenotypic strain differences and detectable genetic difference change may indicate genetic changes in genes, or parts of genes, that do not uniquely map (such as genes already present in several copies or with large numbers of tandem repeats, areas of the genome present in low coverage). Changes in genes present in the evolved strains but absent from the S288c gene reference would be invisible. However, Yamaguchi and Roth claim (Engel & Cherry 2013) that there are only 'miniscule' genetic differences between the S288c strain used to construct the genome

reference, and BY4741 used to generate the starting strains in these experiments. Genetic variation may also not be detectable in genes present in multiple copies or containing many tandem duplications that are difficult to uniquely map, like the *FLO* flocculation genes.

6.4.2. Ploidy

All of the strains isolated which out-competed their ancestors were diploid or triploid and fluorescence cytometry clearly shows populations increasing in ploidy (see this Chapter, section 6.3.8). While an increase of ploidy from haploid to diploid has been reported with salt stress in a multiply auxotrophic haploid background (Gerstein et al. 2006), it does not appear to have been previously documented as a response to ethanol stress. Katou et al noted the occurrence of diploids when trying to isolate haploids by sporulating the Kyokai No. 7 sake strain (Katou et al. 2008), but did not observe the appearance of such strains from haploid starting cultures. The ancestral strains for these experimental evolution experiments have similar multiple auxotrophies to the Gerstein strains, being unable to synthesise or import several amino acids, or make pyrimidines (see chapter 2, section 2.6).

Added ethanol stress is not required for ploidy increase in these evolutions; the diploid Evo8_14_CFPe_E07_4 arose in a control population without added ethanol (see Figure 6.22). However, fluorescence cytometry suggests that increasing the added ethanol stress also increases the rate of transition to diploidy (see Figure 6.24). The ternary plots in Figure 6.24 appear to show the populations making a faster progression in the three dimensional phase plane described by the ratio of haploid, diploid and triploid DNA content counts from positions characteristic of haploidy to those typical of diploids as the added ethanol stress increases. Statistical analysis of such phase plane trajectories is not straightforward and quantitative analysis is further complicated by DNA content change throughout the cell cycle; the diploid counts have a contribution from haploids (after replication but prior to division) and the polyploids not only have a contribution from post-replication diploids but combine their pre- and post-replication counts unlike the lower ploidy groupings.

The changes to the fluorophore ratios discussed in chapter 4 also clearly show earlier onset (lower values for τ) with 4.5% ethanol rather than 0% ethanol controls. However, such a change in the onset of fluorophore ratio change could also be due to genetic sequence change (see chapter 4, section 4.3.8). Ethanol stress may increase the rate of endoreduplication, perhaps by disrupting the mitotic spindle, or

cause diploids to sweep faster through haploid populations by increasing their relative fitness.

The failure to detect any triploid populations in the experimental evolutions with fixed added ethanol stresses may be significant. Following the method of Gerstein et al. (Gerstein et al. 2006), PCR amplification of the *MAT* locus produces bands consistent with the triploid strain Evo10_20_B03_CFPe_5 being *MATα*/*MATα*/*MATα* (Bharat Rash, personal communication). This strongly suggests that the triploid strain has arisen from a tetraploid intermediate, produced by two rounds of endoreduplication, which subsequently lost (at least close to) a euploid number of chromosomes. Such ploidy reduction of tetraploids to apparently euploid triploids has been observed before (Mable & Otto 2001; Gerstein et al. 2008). The competitive fitness of the triploid strain Evo10_20_B03_CFPe_5 is substantially better than any other strain produced by these experimental evolutions over a range of ethanol concentrations, having a shorter lag phase and faster growth (see chapter 5, section 5.3.4). This contradicts Galitski who observed in an isogenic ploidy series both an increase in lag and a 10-15% decrease in growth rates in triploids and tetraploids when compared with haploids and diploids (Galitski 1999, see note 6). Evo10_20_B03_CFPe_5 was not sequenced and may have genetic sequence changes which enhance its fitness despite a deleterious effect of a triploid background. However, fluorescence cytometry indicates that strains with ploidys higher than diploid swept through other evolution 10 populations (see Figure 6.24 d.) and such a deleterious effect of triploidy seems unlikely; fitness in this strain background is likely to increase with ploidy up to triploid.

If triploids have high fitness then it may be significant that the fluorescence cytometry signature of ploidies greater than diploid was not seen in evolutions 6 and 8 with fixed ethanol stresses. This may be due to the longer duration of evolution 10, about 200 generations rather than the approximately 140-150 generations for evolutions 6 and 8; triploids might have emerged in the shorter experimental evolutions given enough generations. However, tetraploids generally have lower fitness and stress resistance than haploids and diploids (see Chapter 2, section 2.2.2) and there are suggestions that components of the mitotic spindle may scale incompatibly provoking genome instability (Storchová et al. 2006). Any tetraploid intermediates that might persist and lose a set of chromosomes to become triploids in an experimental evolution that increases by 0.5% per plate may be lost too rapidly to leave offspring with constant ethanol stresses of 4.5% or 6.5% throughout the evolution. Alternatively, it may solely be chance that a triploid has arisen and been

detected in evolution 10 but not in the other experiments. Despite the apparent fitness benefit of triploidy, not all the wells in evolution 10 form triploid populations; whilst there is a population increase in the ploidy of well B03 from haploid to triploid, the G10 population displays a constant trajectory from haploid to diploid (see Fig. 6.21). Producing a viable triploid from two rounds of endoreduplication and the loss of a set of chromosomes may be sufficiently rare that triploids do not arise in all wells or in all experiments; a surviving triploid arose in B03 but not in G10 in evolution 10.

The diploids that spontaneously emerged from haploid starting populations in the experimental evolution of Gerstein and Otto (Gerstein et al. 2006) required 200 generations with 0.6M NaCl stress (700 generations without stress) to sweep through population; broadly comparable to these experimental evolutions given that the x100 dilution rate for Gerstein et al was an order of magnitude lower than in these experiments. The faster emergence and sweep of diploidy with increasing ethanol stress in these experimental evolutions (see section 6.3.8, Figure 6.24) is comparable to that seen with salt stress (Gerstein et al. 2006). Accordingly increasing the ploidy of haploids may be a general response to stress rather than a specific feature of ethanol.

The Gerstein and Otto starting strains (obtained from B. Mable (B K Mable & S P Otto 2001)) shared many of the same auxotrophies as the starting strains in this experiment, being unable to synthesise histidine, leucine and pyrimidine (*his4Δ leu2Δ* and *ura3Δ*, respectively) or absorb arginine (*can1Δ*). The Gerstein and Otto strains also have a deletion in their respective pheromone receptor locus (*ste6Δ8*-694), plus a *MAT* locus mutation (*mata*-a1) and a tryptophan auxotrophy (*trp1Δ*). It is possible that, even in rich YPD medium, multiple auxotrophies could cause haploinsufficiencies that are exacerbated by stress, particularly if ethanol impairs plasma membrane transport (Leao & Van Uden 1982) or triggers amino acid efflux from the cell (Dombek & Ingram 1986), see Chapter 2, section 2.1.3. Gene amplification is induced by stress in *E. coli* (Slack et al. 2006) and stress can promote aneuploidy in *S. cerevisiae* and gene copy number increases associated with drug resistance (Chen et al. 2012). Increased ploidy may be a response to the stresses imposed by multiple auxotrophy and additional ethanol stress may increase the fitness benefit of increasing copy number by raising ploidy.

Whilst increased ploidy might confer a general benefit, fitness can increase in some environments if genes are present as single copies in a diploid background, a phenomenon dubbed haploproficiency to contrast with the haploinsufficient fitness

deficit of some hemizygous deletions. Delneri et al competed barcoded hemizygous strains from the yeast deletion collection in continuous culture in white grape juice and three minimal medium environments, limited for carbon, nitrogen or phosphate (Delneri et al. 2008). Enrichment or depletion of the bar codes in the populations was used to infer haploproficiency and haploinsufficiency for a large number of genes (see Table 6.7). The deletion collection is based on the BY4743 strain which has a similar complement of engineered deletions to this study.

	Environment			
	Limited glucose	Limited nitrogen	Limited phosphate	Grape juice
Haploproficient genes	815	1,194	733	654
Haploinsufficient genes	685	765	1,277	217

Table 6.7 Counts of haploproficient and haploinsufficient genes determined by diploid competition in continuous cultures (data from Delneri et al 2008).
Each competition was between a hemizygous deletant strain and a strain homozygous wild-type at that locus. Strains were cultured in media with limited glucose, nitrogen or phosphate or in white grape juice.

Copy number variation detected by CNVseq suggests that there is an approximately 50kb hemizygous deletion from ~450Mb to 500Mb on chromosome V of Evo10_20_G10_YFPe_2 (see Chapter 6, section 6.3.6). Gresham et al noted that diploid starting strains were more likely than haploid to have large amplifications or deletions after ~200 generations of chemostat growth limited for glucose, sulphate or phosphate (Gresham et al. 2008). The location of two copies of the *TY1* retrotransposon at the likely boundaries of this deletion (YER138C/YER137C-A and YER160C/YER159C-A) suggests that it may have arisen as a recombination event between these homologous genes or their LTR repeats. A similar 13kb deletion of three genes in chromosome X by homologous recombination between TY elements has been well characterised in the *DEL1* mutator strains (Liebman et al. 1981). *TY* boundaries were also noted by Gresham et al in the large structural changes they detected (Gresham et al. 2008). If the *TY* transposons do mark the boundaries of the chromosome V deletion in strain Evo10_20_G10_YFPe_2, then there are 27 genes present as a single copy in the diploid background in this strain. A priori, such a large hemizygous deletion, encompassing four essential genes (*LSM5*, *SCC4*, *SPT15* and *BUR6*), would be expected to be deleterious or at best neutral. However, according to Delneri et al., just under half of these genes (13) were haploproficient in at least one environment; 5 of these were haploproficient in two or more

environments and 2 of these in three environments (see Table 6.8). *SPT15* (YER148W) in the middle of the deletion encodes the TATA-binding protein and has been successfully mutated to produce strains with increased ethanol tolerance (Alper et al. 2006; Gao et al. 2010). Delneri et al. found *SPT15* to be haploproficient during carbon starvation and haploinsufficient in nitrogen starvation.

However, the Evo10_20_G10_YFPe_2 deletion cannot be construed from the work of Delneri et al. to be wholly beneficial. *FTR1* (YER145C) and YER158C were haploinsufficient in 2 and 3 environments respectively and the strain hemizygous for the *RTR1* (YER139C) gene was rapidly eliminated in all environments tested, suggesting strong haploinsufficiency. Homozygous deletion of *RTR1* causes a temperature sensitive growth phenotype (Gibney et al. 2008). It is uncertain what the phenotypic effects are of copy number reduction of all 27 genes.

		Environment			
Gene	Alias	Limited Carbon	Limited Nitrogen	Limited Phosphate	Grape Juice
YER138C	*TYA* Gag and *TYB Pol*				
YER137C-A					
YER138W-A					
YER139C	*RTR1*	-	-	-	-
YER140W		0	0	+	0
YER141W	*COX15*	0	0	0	0
YER142C	*MAG1*	0	0	+	0
YER143W	*DDI1*	+	+	+	0
YER144C	*UBP5*	0	0	0	0
YER145C	*FTR1*	-	-	0	0
YER145C-A		0	0	0	0
YER146W	*LSM5*	+	+	0	0
YER147C	*SCC4*	0	0	0	0
YER147C-A		0	0	0	0
YER148W	*SPT15*	+	-	0	0
YER148W-A		0	0	0	0
YER149C	*PEA2*	0	0	0	+
YER150W	*SPI1*	+	+	+	0
YER151C	*UBP3*	+	+	0	0
YER152C		0	+	0	0
YER152W-A		0	0	0	0
YER153C	*PET122*	0	0	0	0
YER154W	*OXA1*	0	0	+	0
YER155C	*BEM2*	0	+	+	0
YER156C		0	0	0	+
YER157W	*COG3*	0	0	+	0
YER158C		-	-	-	0
YER158W-A		0	0	0	0
YER159C	*BUR6*	0	0	0	0
YER160C	*TYA Gag and TYB Pol*	0	0	0	0
YER159C-A		0	0	0	0

Table 6.8 Haploproficiency and haploinsufficiency of genes possibly made hemizygous by the deletion on chromosome V from the TYA Gag and TYB Pol genes at 449,024 to 498,119 in the strain Evo10_20_G10_YFPe_2 (data from Delneri et al, 2008).
A '+' indicates haploproficiency in an environment, a '-' indicates haploinsufficiency and '0' indicates no observable change when a gene is present as a hemizygous deletant.

The strength of continuous culture for assessing haploproficiency/insufficiency is that there is competition in a defined culture medium over many generations (50

generations or fewer to avert adaptive evolution) which can amplify the selective effects of small growth rate differences. However, this strength could also be a potential weakness when attempting to generalise such findings for batch culture. The conditions in minimal media continuous culture may be experienced only briefly in rich medium batch culture and may not be the key drivers of adaptive evolution in that environment. It has also been previously noted that yeast may be adapted for episodic plenty rather than consistent scarcity (see Introduction, section 2.4.7). Grape juice is closest to the YPD rich medium and Delneri et al found that haploinsufficiency was much rarer in grape juice than in the minimal media; only 3.8% of genes haploinsufficient in grape juice (Delneri et al. 2008), (see Table 6.7). This accords with the work of Deutschbauer et al. in YPD which showed haploinsufficiency for 3% of genes in YPD rich media (Deutschbauer et al. 2005).

6.4.3. Have the genotypic causes of phenotype been found?

Whole genome sequencing and ploidy evaluation were done to explore the genotypic correlates of phenotypic change. Identifying adaptive genetic changes is non-trivial (see Introduction Section 2.4.7 “Drawing conclusions from experimental evolution”). Moving from identifying genetic correlates of change to positing a direct causal link requires a causative mutation or mutations to be present without other confounding factors (eg. sequence variation without ploidy change or vice versa) and would ideally arise in multiple parallel evolutions and be transferrable into the ancestral background to effect the same phenotypic changes.

None of the genetic changes in ploidy or sequence change meet these strict criteria. It is also an open question whether all the genetic correlates of fitness change have been found by the methods used. Certainly all the evolved strains that out-competed their ancestors had increased their ploidy; both fluorescence cytometry and cellular enlargement strongly support this. Ploidy increase may explain the common features of increased competitive fitness e.g. evolved strains that show lower maximum growth rates than their ancestors but which depress the maximum growth rate of their ancestral competitors by a much greater amount (see Chapter 5). However, there are some phenotypic features of these evolved strains that are not wholly explained by ploidy alone. For example, only some of the experimentally evolved diploid CFP strains showed enhanced tolerance of organic acids while none of the YFP diploids or the haploid CFP ancestor did so. Similarly, the two sequenced strains from evolution 6 showed longer duration of active growth, something not

observed in evolutions 8 or 10 despite no sequence changes being found in their genomes.

Whilst a bulk property like the ploidy of a homogenous population of cells is reasonably straightforward to determine, the complete genome sequence of a eukaryotic cell, albeit one with a relatively modest genome size (~12.1Mb), is both complex and impossible to completely determine with current technology. There are sources of positive errors, including sequence specific errors, particularly inverted repeats and 'GGC' sequences (Nakamura et al. 2011) but there are also sources of false negative outcomes in which sequence changes can be missed. Next generation sequencing has short read lengths (100 bases in these experiments) and the reads are error prone with each read likely to have one or more random errors. The shortness of the reads causes problems with both assembly and mapping. Short reads are more difficult to establish unambiguous mapping positions, particularly for genes with paralogs, genes present in more than one copy (particularly determining potential variation in copy number in experimentally evolved strains), or in genes containing repeated domains. If reads cannot be unambiguously mapped then sequence variation is likely to be distributed between multiple loci and genuine sequence changes will be difficult to distinguish from errors. Some likely genetic targets for improved ethanol tolerance are refractory to sequence determination by next generation sequencing in this manner. For example, *S. cerevisiae* strains engineered to over-express the flocculation gene *FLO1* late in fermentation have improved growth and ethanol fermentation (Li et al., 2012). Both *FLO1* and its paralog *FLO5*, are mannose binding, lectin-like proteins that contain a number of repeated flocculin domains and reads mapped to both of these genes have very low mapping quality due to the occurrence of these repeats.

Aside from the difficulties of finding variants in a complex genome with short reads, there are also imperfections in both the current sample preparation technology and the bioinformatic detection of variants. These experimental evolution experiments had very good sequenced read coverage (approximately x100). However, coverage variation may arise due to uneven amplification during the polymerase chain reaction (PCR) step in sample preparations causing redundant over-amplification of some sequences (which was removed in these experiments prior to variant calling) and correspondingly poor representation of other sequences (Kozarewa et al. 2009). Park et al. consider that variant detection in next generation sequences is also demonstrably "still prone to errors" particularly false negative errors (Park et al. 2014).

7. General discussion

The first research objective was to identify experimental evolutions in which adaptive evolution had occurred amongst dozens run in parallel in YPD batch culture subjected to both fixed and increasing ethanol stresses. Analysis of fluorophore ratio curves proved to be a practical method for the identification of such cultures as well as the fluorescent tag of the prevailing population. The Dirac equation was used to fit the fluorophore ratio trajectory for experimentally evolving populations. This equation has parameters which are indicators of the fitness and the generation of onset of adaptive change. The fitting of the Dirac curve frequently required truncation of the data and heuristics were devised to ensure that data was not over fitted by excessive truncation. Null modelling of different bottleneck sizes showed that the $2x10^3$ bottleneck size of one evolution (number 6) is close to the lower limit below which large fluorophore ratio changes can arise due to neutral drift that are indistinguishable from adaptive mutation. However, the $1x10^4$ and $1.5x10^4$ bottleneck sizes of the later evolutions make them refractory to substantial drift. In evolution 8, the tau (onset) parameter for the fluorophore ratios in the 4.5% treatment diverged substantially earlier than the 0% ethanol control suggesting that ethanol stress decreases the waiting time for adaptation or increases the speed of sweep.

All of the strains isolated which out-competed their haploid ancestors were diploid or triploid. Such an emergence of diploid strains would be unexceptional in wild-type cells since *S. cerevisiae* appears to favour diploidy over haploidy. In common with many other known yeast species *S. cerevisiae* will mate whenever haploid cells of opposite mating type are in proximity but atypically, it requires starvation to enter meiosis and sporulate (see Section 2.2.2). However, although wild-type *S. cerevisiae* are homothallic, these starting strains are generally unable to switch mating type due to a deletion of the *HO* gene locus (see section 2.23). Further work (Bharat Rash, pers comm) has identified that the only diploid strain isolated from an experimental evolution without added ethanol (well E07 in experimental evolution 8) was heterozygous at the *MAT* locus, i.e. it had formed following mating after a mating-type switch had occurred. He was also able to isolate *MAT* heterozygotes from some, but not all, experimentally evolving populations although their fluorophore was not always that prevailing in competition despite the mating type switching conferring a histidine biosynthetic competence lacking in the starting strains (see chapter 4, section 4.4.1 "*MAT* switching and histidine auxotrophy"). This is not

surprising since there may be little or no fitness advantage in histidine biosynthetic ability in the rich media used in these experiments and the stochastic nature of experimental evolution means that adaptive mutations may need to arise multiple times before any sweep, and may not ultimately prevail due to clonal interference (Elena & Lenski, 2003). Sequencing confirmed the deletion of the *HO* gene in all strains and therefore it is likely that mating type switch followed rare recombination events, likely following a DNA break in or near the *MAT* locus. The other isolated strains were all homozygous MATα and hence had arisen from endoreduplication. Two endoreduplicated strains, those sequenced from from well D06 in evolution 6, and well F10 in evolution 10, lacked detectable genome sequence variation from their ancestral strain and, all were histidine auxotrophs like the ancestors. If the sole change likely to affect fitness from the ancestors in these strains is endoreduplication then the preponderance of endoreduplicants compared to the diploids derived by mating suggests that endoreduplication is commoner than mating type switching. It is possible that the rate of endoreduplication may increase with ethanol stress. Fluorescence cytometry clearly shows populations increasing in ploidy at a rate proportional to ethanol stress. Such an increase in rate of diploid formation could account for the faster progression to diploidy with increasing ethanol (see Figure 6.22 in Chapter 6, section 6.3.8 “Analysis of ploidy”) and the smaller values for tau, the parameter related to onset time of adaptive mutation, when the Dirac equation was fitted to the evolution 8 experiments with 4.5% ethanol rather than to the controls lacking ethanol (see Figure 4.25 in Chapter 4, section 4.3.8). The rate of endoreduplication in these starting strains may be higher than wild type due to a single amino acid deletion in *APC1*, the largest subunit of the anaphase-promoting complex, which is an essential component of cell cycle progression, and may render mutants more prone to spindle failure/non-disjunction that might increase ploidy. This might explain the frequent and relatively rapid emergence of higher ploidy in these experimental evolutions (100-200 generations) compared with the larger number of generations seen by Gerstein et al. in a similar strain backgrounds (Gerstein et al. 2006). However, this might also be due to the smaller population sizes in these experiments and the stress applied; Gerstein et al. used a salt stressor.

Any fitness advantage for diploids over haploids is a likely confounding factor in a search for genetic sequence changes in experimental evolutions with haploid starting strains. Ploidy increase was universally detected in the experimentally evolved strains able to out-compete their ancestors and an endoreduplicated triploid strain

(not sequenced) showed the largest increase in fitness across the range of ethanol stresses whereas no sequence changes were detected in two strains. Accordingly, it is highly unlikely that ploidy increase confers no fitness advantage in these experimental evolutions. Further analysis of fitness competitions with isogenic diploids derived from mating type switching (Bharat Rash, pers comm.) may confirm this, although such diploids are not exactly comparable since they gain histidine biosynthetic capability from mating type switching.

The converse argument that at least some of the emerging sequence changes in the experimental evolutions are neutral or deleterious rather than adaptive is much more tenable. Only 8 strains were sequenced and few sequence variants were found in each strain. The paucity of variants was partly to be expected; the evolutions ran for fewer than 200 generations consistent with the research goal of identifying initial adaptive sequence changes. Only one mutation was seen in more than one strain (an identical *COB1* intron mutation in wells D03 and G04 in the evolution 8) and that was extremely likely to have arisen in an overnight culture. The absence of common sequence variants may be because there are many genetic targets for adaptive evolution in these experiments (a rugged fitness landscape). That would be consistent with the evolution tolerance literature which identified hundreds of potential ethanol tolerance genes by evaluating deletion library strains and analysing gene expression. Alternatively at least some mutations may be neutral hitchhikers in strains selected for higher ploidy.

There is a great deal of work suggesting that damage to both cytoplasmic and intracellular membranes is a critical aspect of ethanol toxicity (See Introduction, section 2.1.3). Accordingly, the beneficial effects of a reduced cell surface area to volume ratio may be a driver for the widespread increase in ploidy seen. However, a global increase in gene copy number under stress may also be adaptive irrespective of cell size, particularly if Kron is right to suggest that laboratory strains are “recessive mutants at multiple loci” (Kron 1997) since such genetic lesions may promote widespread haploinsufficiency and a corresponding fitness increment with higher ploidy. Higher euploidy may however, be a key attribute of ethanol tolerance; it is not uncommon in industrial yeast strains such as the bioethanol strain ZK2, an aneuploid strain with a ploidy between 3n and 4n (Zheng et al. 2013).

The cell sizes in an *S.cerevisiae* population depend both on the sizes of mother cells at bud initiation of budding and the cell growth during budding. Further work to evaluate the ethanol tolerance of isogenic populations with different cell size profiles

could be carried out by arresting division using temperature sensitive mutations (Johnston et al. 1979) or adding mating pheromone (Dolan 1996) to enlarge cells. It is difficult to see how it would be possible to evaluate the ethanol tolerance of cells throughout a batch culture cycle using these means, but a measure of ethanol tolerance such as percentage viability following ethanol shock could be informative of the relationship between tolerance and cell size. Only one experimentally evolved diploid strain was characterised and sequenced that arose in a control without an added ethanol stress (in well E07 in evolution 8) and this proved to be a *MATα*/*MATa* diploid that had acquired the capacity to grow on histidine deficient minimal media.

However, there does appear to be phenotypic variation between different strains and between the evolution series which suggests that phenotype cannot wholly be explained by ploidy increase alone. All of the experimentally evolved strains were able substantially to depress the maximum growth rate of their ancestral competitors, whilst also showing rates of maximum growth lower than either strain in an ancestral competition (suggesting a trade-off between ethanol tolerance and maximum growth rate). However, the strains from evolution 6 also showed a longer duration of growth due to earlier emergence from lag dormancy. Those strains isolated from media with constant amounts of added ethanol throughout the evolution showed changes in the rates, duration and amount of growth that enabled them to outgrow their competitors across the range of ethanol concentrations tested (0%, 4.5%, 6.5%). However, the strains isolated from the ramped ethanol stress of evolution 10 (from wells F10 and G10) appeared to grow worse than their ancestors with no added ethanol while their performance improved as ethanol concentrations increased suggesting antagonistically pleiotropic adaptation. This suggests that there are mutations adaptive at higher ethanol concentrations that are detrimental with low concentrations. The strain isolated from well F10 in evolution 10 had no identified sequence variations from the ancestor, definitively lacking the glycogen debranching enzyme mutation found in the other evolution 10 strain from well G10. These phenotypic characteristics suggest that at least some genetic variation from the ancestors in the evolved strains has been missed.

All the evolved strains that out-competed the ancestors were larger than the haploid starting strains, with the evolution 10 strains largest of all. All the evolved diploids had a lower maximum growth rate than their ancestor but they depressed the maximum rate of growth of a haploid competitor much further. Larger cells may be better able to maintain an active transport gradient under ethanol stress (see Section 2.1.3) allowing them to absorb nutrients more efficiently than the smaller haploids

although their large size may reduce their maximum growth rate. The larger diploids from evolution 10 (Evo10_20_F10_CFPe_2, Evo10_20_G10_YFPe_2) may then be sufficiently large that their growth is actually poorer if their competing haploid strains are not inhibited by a sufficiently large ethanol stress. However, the evolution 10 triploid strain Evo10_20_B03_CFPe_5 is larger than the diploids and shows both a faster rate of growth and greater overall growth in all ethanol concentrations. It is possible that this (unsequenced) strain has genetic changes in addition to ploidy which confer a fitness benefit. Two strains showed a shorter lag than their haploid competitors Evo6_11_B06_CFPe_4 and Evo10_20_G10_YFPe_2 (although the gap between the lag of Evo10_20_F10_CFPe_2 and its haploid competitor narrows with increased addition of ethanol). Cell enlargement could affect the duration of growth lag in several ways; large cells may have greater stored reserves to withstand the poor nutrient levels post-diauxy without entering a prolonged quiescence which would lengthen growth lag. There may be a trade-off between such benefits and a possible requirement for a longer duration for the preparatory biosynthesis required before cell growth and divison can properly commence, a corrolary of the lower maximal growth rates of large cells.

The addition of relatively low levels of ethanol 4.5% and 6.5% in evolution 6 and 8 clearly favour the cell enlargment of diploidy. However, the 9.5% ethanol stress may provide a fitness benefit for further cell enlargement, producing diploids larger than those seen in other evolutions as well as at least one triploid strain. Equally, ethanol stress may induce haploinsufficiency which may be ameliorated by diploidy at lower ethanol levels but may only favour higher ploidies over 6.5% ethanol.

A key research goal was the isolation of genetic sequence changes that confer ethanol tolerance. Completely satisfying this goal would have required the verification of mutations identified by next generation sequencing using Sanger sequencing and a transfer of the mutation to the starting genotypic background that confers the same fitness benefit. Accordingly, whilst some sequence candidates have been isolated that have not been associated with ethanol tolerance previously this research goal is far from being fully met.

Three of the four evolved CFP strains tested grew well on organic acids in Biolog GenIII plates whilst neither the remaining evolved CFP strain nor any of the YFP strains failed to do so. This is not statistically significant ($p=0.1429$), and would require the testing a larger number of samples to confirm whether there is a strain bias in the emergence of this phenotype. There is a partial deletion in *EGT2* in the

ancestral CFP strain, a gene which has been associated with organic acid tolerance. However, an *egt2* partial deletion, or an undetected additional genetic difference in the strains, could not be sufficient for a strain bias in the development of this phenotype since the ancestral CFP did not show organic acid growth. Neither could any bias in phenotypic emergence be solely an interaction between *egt2Δ* and ploidy, since one CFP strain (also *egt2Δ* and diploid) does not show organic acid tolerance. It is possible that a genetic difference between the strains produces a difference in the 'adaptive landscape' of each strain, allowing a fitness benefit to be readily accessed in one strain by a single mutation that might require more than one genetic change in the other. No firm conclusions about this phenotypic trait can be made without testing more experimentally evolved strains and the linkage between this phenotypic trait and *egt2* or any other strain-specific mutation would require that restoration of the wild-type gene in the CFP strain removed the bias while the transfer of the mutation into the YFP strain conferred it.

In addition to the amino acid deletion in *APC1* (YNL172W), both ancestors also have a small deletion in *MEH1* (YKR007W) and share a SNP in *CTR9*/*CDP1* (YOL145C). The *meh1* mutation suggests that the starting strains may have poorer ethanol tolerance than wild type S. cerevisiae. In the evolved sequenced strains non-synonymous heterozygous SNPs were found in *GCN1* and *PHO23*, homozygous SNPs in *GPI18* and *GDB1*, and one mitochondrial SNP was found in *COB1*. None of these genes are well established in the literature as loci associated with enhanced ethanol tolerance. Two strains had two genetic changes; one strain having a 50Kb heterozygous deletion and the *GDB1* SNP, the other strain had the *COB1* mitochondrial SNP and the *PHO23* SNP. This suggests that batch culture can produce more than one mutation/structural change within 150-200 generations. The 50Kb heterozygous deletion is likely to have been generated by recombination between two *TY1* loci.

A strength of experimental evolution is that it places few constraints on the outcome of the experiment; adaptive evolution may change unexpected genetic targets. However, this does mean that selection pressures other than those intended may influence the outcome. The toxic effects of added ethanol were the intended selection pressures. However, there may have been other pressures, related to growth on plates shaken and analysed in these plate readers. For example, sampling every five minutes may have caused reactive oxygen species to be liberated as a result of photo bleaching. To some extent adaptations to these selection pressures are covered by controls in which no ethanol was added. However, with further time it

would have been desirable to have run a series of experimental evolutions without an added ethanol stress to look at the phenotypic and genotypic changes associated with adaptation to this experimental system. However, it is highly desirable that the starting strains for subsequent experimental evolutions with ethanol stresses would be isogenic and having any adaptations to the experimental system. Accordingly it would have been necessary to take a pre-adapted strain and construct an isogenic strain with another fluorophore from it (and ideally verify by sequencing that both strains were isogenic). This would have required a minimum of two rounds of sequencing with the cost and time that would have entailed.

8. Bibliography

Abe, H. et al., 2009. Ethanol-tolerant Saccharomyces cerevisiae strains isolated under selective conditions by over-expression of a proofreading-deficient DNA polymerase delta. *J Biosci Bioeng*, 108, pp.199–204.

Adams, J. et al., 1985. Physiological characterization of adaptive clones in evolving populations of the yeast, Saccharomyces Cerevisiae. *Genetics*, 110(2), pp.173–185.

Adams, J. & Hansche, P.E., 1974. Population studies in microorganisms I. Evolution of diploidy in Saccharomyces cerevisiae. *Genetics*, 76(2), pp.327–335.

Akao, T. et al., 2011. Whole-genome sequencing of sake yeast Saccharomyces cerevisiae Kyokai no. 7. *DNA Res*, 18(6), pp.423–34.

Albrecht, G. et al., 1998. Monitoring the Gcn4 protein-mediated response in the yeast Saccharomyces cerevisiae. *J Biol Chem*, 273(21), pp.12696–702.

Alexandre, H. et al., 2001. Global gene expression during short-term ethanol stress in Saccharomyces cerevisiae. *FEBS Lett*, 498(1), pp.98–103.

Alexandre, H., Rousseaux, I. & Charpentier, C., 1994. Relationship between ethanol tolerance, lipid composition and plasma membrane fluidity in Saccharomyces cerevisiae and Kloeckera apiculata. *FEMS Microbiol Lett*, 124(1), pp.17–22.

Allen, C. et al., 2006. Isolation of quiescent and nonquiescent cells from yeast stationary-phase cultures. *J Cell Biol*, 174(1), pp.89–100.

Alper, H. et al., 2006. Engineering yeast transcription machinery for improved ethanol tolerance and production. *Science*, 314(5805), pp.1565–1568.

Andalis, A.A. et al., 2004. Defects arising from whole-genome duplications in Saccharomyces cerevisiae. *Genetics*, 167(3), pp.1109–21.

Anderson, J.B., Sirjusingh, C. & Ricker, N., 2004. Haploidy, diploidy and evolution of antifungal drug resistance in Saccharomyces cerevisiae. *Genetics*, 168(4), pp.1915–23.

Andrade, M. & Bork, P., 1995. HEAT repeats in the Huntington's disease protein. *Nat Genet*, 11, pp.115–116.

Andreasen, A. & Stier, T., 1953. Anaerobic nutrition of Saccharomyces cerevisiae. I. Ergosterol requirement for growth in a defined medium. *J Cell Physiol*, 41(1), pp.23–36.

Anon, Product description: Greiner 96 “F” well Polystyrene microplate. Available at: http://www.greinerbioone.com/en/england/articles/catalogue/article/68_4/16664/ [Accessed June 5, 2013].

Aragon, A. & Rodriguez, A., 2008. Characterization of differentiated quiescent and nonquiescent cells in yeast stationary-phase cultures. *Mol Biol Cell*, 19, pp.1271–1280.

Ashburner, M., 1998. Speculations on the subject of alcohol dehydrogenase and its properties in Drosophila and other flies. *BioEssays*, 20(11), pp.949–954.

Atwood, K.C., Schneider, L.K. & Ryan, F.J., 1951. Periodic selection in Escherichia coli. *Proc Natl Acad Sci U S A*, 37(3), pp.146–155.

Auesukaree, C. et al., 2009. Genome-wide identification of genes involved in tolerance to various environmental stresses in Saccharomyces cerevisiae. *J Appl Genet*, 50, pp.301–310.

Auger, E. a et al., 1989. Construction of lac fusions to the inducible arginine- and lysine decarboxylase genes of Escherichia coli K12. *Mol Microbiol*, 3(5), pp.609–20.

Austin, K.T. & Butzke, C.E., 2000. Spectrophotometric Assay for Arginine in Grape Juice and Must. *Am. J. Enol. Vitic*, 51(3), pp.227–232.

Baganz, F. et al., 1997. Suitability of replacement markers for functional analysis studies in Saccharomyces cerevisiae. *Yeast*, 13(16), pp.1563–73.

Bandas, E.L. & Zakharov, I.A., 1980. Induction of rho- mutations in yeast Saccbaromyces cerevisiae by ethanol. *Annu Rev Microbiol.*, 71, pp.193–199.

Baroni, M.D. et al., 1992. cAMP-mediated increase in the critical cell size required for the G1 to S transition in Saccharomyces cerevisiae. *Exp Cell Res.*, 201(2), pp.299–306.

Barrett, P.M. & Willis, K.J., 2001. Did dinosaurs invent flowers? Dinosaur – angiosperm coevolution revisited. *Biol Rev Camb Philos Soc*, 76, pp.411–447.

Barrick, J.E. et al., 2009. Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature*, 461(7268), pp.1243–7.

Bates, DM & Watts, D., 1988. *Nonlinear Regression Analysis and Its Applications*, Wiley.

Bauer, B.E. et al., 2003. Weak organic acid stress inhibits aromatic amino acid uptake by yeast , causing a strong influence of amino acid auxotrophies on the phenotypes of membrane transporter mutants. *Eur J Biochem*, 3195, pp.3189–3195.

Benbadis, L. et al., 2009. Isolation of two cell populations from yeast during high-level alcoholic fermentation that resemble quiescent and nonquiescent cells from the stationary phase on glucose. *FEMS Yeast Res*, 9(8), pp.1172–86.

Bereiter-Hahn, J. & Voth, M., 1994. Dynamics of Mitochondria in Living Cells: Shape Changes, Dislocations, Fusion, and Fission of Mitochondria. *Microsc Res Tech.*, 27, pp.198–219.

Betz, C., Schlenstedt, G. & Bailer, S.M., 2004. Asr1p, a novel yeast ring/PHD finger protein, signals alcohol stress to the nucleus. *J Biol Chem.*, 279(27), pp.28174–81.

Blank, D. et al., 2014. The predictability of molecular evolution during functional innovation. *Proc Natl Acad Sci USA*, [epub], pp.1–6.

Blateyron, L. & Sablayrolles, J., 2001. Stuck and slow fermentations in enology: statistical study of causes and effectiveness of combined additions of oxygen and diammonium phosphate. *J Biosci Bioeng.*, 91(2), pp.184–189.

Bokman, S. & Ward, W., 1981. Renaturation of Aequorea green-fluorescent protein. *Biochem Biophys Res Commun.*, 101(4), pp.1372–1380.

Boulton, B. et al., 1996. Yeast and biochemistry of ethanol fermentation. In *Principles and Practices of Winemaking*. New York: Chapman and Hall, pp. 139–172.

Boulton, C.A. & Quain, D.E., 2001. *Brewing Yeast and Fermentation.*, Oxford: Blackwell Science.

Brachmann, C.B. et al., 1998. Designer Deletion Strains derived from Saccharomyces cerevisiae S288C: a Useful set of Strains and Plasmids for PCR-mediated Gene Disruption and Other Applications. *Yeast*, 132, pp.115–132.

Breslow, D. et al., 2008. A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nat Methods.*, 5(8), pp.711–718.

Brown, S. & Oliver, S., 1982. Isolation of ethanol-tolerant mutants of yeast by continuous selection. *Eur J Appl Microbiol Biotechnol*, 16(2-3), pp.119–122.

Bryson, V. & Szybalski, W., 1952. Microbial Selection. *Science*, 116(3003), pp.45–51.

Butow, R. a & Avadhani, N.G., 2004. Mitochondrial signaling: the retrograde response. *Mol Cell*, 14(1), pp.1–15.

Byrne, K.P. & Wolfe, K H, 2005. The Yeast Gene Order Browser: combining curated homology and syntenic context reveals gene fate in polyploid species. *Genome Res*, 15, pp.1456–1461.

Cabeqa-Silva, C. et al., 1982. Temperature relations of ethanol-enhanced petite mutation in Saccharomyces cerevisiae: Mitochondria as targets of thermal death. *FEMS Microbiol. Lett.*, 11, pp.149–151.

Cakar, Z.P. et al., 2005. Evolutionary engineering of multiple-stress resistant Saccharomyces cerevisiae. *FEMS Yeast Res.*, 5, pp.569–78.

Campos, P.R. a et al., 2008. Environmental heterogeneity enhances clonal interference. *Evolution*, 62(6), pp.1390–9.

Carter, B.L.A. & Jagadish, M.N., 1978. The relationship between cell size and cell division in the yeast Saccharomyces cerevisiae. *Exp Cell Res.*, 112(1), pp.15–24.

Castrejón, F. et al., 2002. Acetaldehyde and ethanol are responsible for mitochondrial DNA (mtDNA) restriction fragment length polymorphism (RFLP) in flor yeasts. *System. Appl. Microbiol.*, 25(3), pp.462–7.

Cell, I. & Pfeffer, L.M., 1995. The PHD finger : implications for chromatin-mediated transcriptional mediation. *Trends Biochem Sci*, 20(2), pp.56–59.

Chandler, M. et al., 2004. A genomic approach to defining the ethanol stress response in the yeast. *Annal Microbiol*, 54(4), pp.427–454.

Chen, G. et al., 2012. Hsp90 stress potentiates rapid cellular adaptation through induction of aneuploidy. *Nature*, pp.1–5.

Chen, K. et al., 2009. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods*, 6(9), pp.677–681.

Chen, M., 2011. Glycogen Storage Diseases. In *Molecular Pathology Library, volume 5*. pp. 677–681.

Chen, X. & Clark-Walker, G., 2000. The petite mutation in yeasts: 50 years on K. W. Jeon, ed. *Int. Rev Cytol.*, 194, pp.197–238.

Chopra, R., Sharma, V.M. & Ganesan, K., 1999. Elevated growth of Saccharomyces cerevisiae ATH1 null mutants on glucose is an artifact of nonmatching auxotrophies of mutant and reference strains. *Appl Environ Microbiol*, 65(5), pp.2267–8.

Chowdhury, S., Smith, K.W. & Gustin, M.C., 1992. Osmotic stress and the yeast cytoskeleton: phenotype-specific suppression of an actin mutation. *J Cell Biol*, 118(3), pp.561–71.

Cipollina, C. et al., 2005. SFP1 is involved in cell size modulation in respiro-fermentative growth conditions. *Yeast*, 22(5), pp.385–99.

Colegrave, N. & Buckling, A., 2005. Microbial experiments on adaptive landscapes. *BioEssays*, 27(11), pp.1167–73.

Colman-Lerner, A., Chin, T.E. & Brent, R., 2001. Yeast Cbk1 and Mob2 activate daughter-specific genetic programs to induce asymmetric cell fates. *Cell*, 107(6), pp.739–50.

Conant, G.C. & Wolfe, Kenneth H, 2007. Increased glycolytic flux as an outcome of whole-genome duplication in yeast. *Mol Syst Biol*, 3(129), pp.1–12.

Coote, H. & Kirsop, B., 1976. Factors responsible for the decrease in pH during beer fermentations. *J Inst Brew*, 82, pp.149–153.

Costa, V. et al., 1997. Mitochondrial superoxide dismutase is essential for ethanol tolerance of Saccharomyces cerevisiae in the post-diauxic phase. *Microbiology*, 143(Pt 5), pp.1649–1656.

Daniel, J. a et al., 2006. Eliminating gene conversion improves high-throughput genetics in Saccharomyces cerevisiae. *Genetics*, 172(1), pp.709–11.

Davey, H.M. et al., 2012. Genome-wide analysis of longevity in nutrient-deprived Saccharomyces cerevisiae reveals importance of recycling in maintaining cell viability. *Environ Microbiol*, 14(5), pp.1249–60.

Davies, T.J. et al., 2004. Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proc Natl Acad Sci U S A*, 101(7), pp.1904–9.

De Deken, R.H., 1966. The Crabtree effect: a regulatory system in yeast. *J. gen. Microbiol.*, 44(2), pp.149–56.

Delneri, D. et al., 2008. Identification and characterization of high-flux-control genes of yeast through competition analyses in continuous cultures. *Nat Genet*, 40(1), pp.113–7.

DeLuna, A. et al., 2008. Exposing the fitness contribution of duplicated genes. *Nat genet.*, 40(5), pp.676–81.

Deutschbauer, A.. et al., 2005. Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. *Genetics*, 169, pp.1915–1925.

Devineni, A. V. & Heberlein, U., 2009. Preferential Ethanol Consumption in Drosophila Models Features of Addiction. *Curr Biol.*, 19(24), pp.2126–2132.

Dhar, R. et al., 2011. Adaptation of Saccharomyces cerevisiae to saline stress through laboratory evolution. *J.Evol.Biol.*, 24(5), pp.1135–53.

Diez-Gonzalez, F. & Karaibrahimoglu, Y., 2004. Comparison of the glutamate-, arginine- and lysine-dependent acid resistance systems in Escherichia coli O157:H7. *J Appl Microbiol*, 96(6), pp.1237–44.

Dinh, T.N. et al., 2008. Adaptation of Saccharomyces cerevisiae cells to high ethanol concentration and changes in fatty acid composition of membrane and cell size. *PLoS ONE*, 3(7), p.e2623.

Dixit, R. et al., 2006. Using intrinsically fluorescent proteins for plant cell imaging. *Plant J*, 45, pp.599–615.

Dolan, J.W., 1996. Novel aspects of pheromone-induced cell-cycle arrest in yeast. *Curr Genet*, 30, pp.469–475.

Dombek, K. & Ingram, L., 1986. Magnesium limitation and its role in apparent toxicity of ethanol during yeast fermentation. *Appl Environ Microbiol*, 52(5), p.975.

Du, X & Takagi, H, 2007. N-Acetyltransferase Mpr1 confers ethanol tolerance on Saccharomyces cerevisiae by reducing reactive oxygen species. *Appl Microbiol Biotechnol*, 75, pp.1343–1351.

Du, Xiaoyi & Takagi, Hiroshi, 2007. N-Acetyltransferase Mpr1 confers ethanol tolerance on Saccharomyces cerevisiae by reducing reactive oxygen species. *Appl Microbiol Biotechnol*, 75(6), pp.1343–51.

Dudley, R., 2004. Ethanol, Fruit Ripening, and the Historical Origins of Human Alcoholism in Primate Frugivory. *Integr. Comp. Biol.*, 44(4), pp.315–323.

Dujon, B., 2010. Yeast evolutionary genomics. *Nature*, 11(7), pp.512–524.

Dunham, Maitreya J et al., 2002. Characteristic genome rearrangements in experimental evolution of Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A*, 99(25), pp.16144–9.

Edgar, B.A. & Orr-Weaver, T.L., 2001. Endoreplication Cell Cycles : More for Less. *Cell*, 105, pp.297–306.

El-Behhari, M. et al., 2000. Evidence for a lectin in Kluyveromyces sp. that is involved in co-flocculation with Schizosaccharomyces pombe. *FEMS Microbiol Lett.*, 184(1), pp.41–6.

Elena, S.F. & Lenski, Richard E, 2003. Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat Rev Genet*, 4(6), pp.457–69.

Elsliger, M.A. et al., 1999. Structural and spectral response of green fluorescent protein variants to changes in pH. *Biochemistry*, 38(17), pp.5296–301.

Engel, S.R. & Cherry, J.M., 2013. The new modern era of yeast genomics : community sequencing and the resulting annotation of multiple Saccharomyces cerevisiae strains at the Saccharomyces Genome Database. *Database*, 2013(bat012).

Fabre, A.-L., Orlean, P. & Taron, C.H., 2005. Saccharomyces cerevisiae Ybr004c and its human homologue are required for addition of the second mannose during glycosylphosphatidylinositol precursor assembly. *FEBS J*, 272(5), pp.1160–8.

Feklistov, A. et al., 2008. Rifamycins do not function by allosteric modulation of binding of Mg2+ to the RNA polymerase active center. *Proc Natl Acad Sci USA*, 105(39), pp.14820–5.

Fernandes, A. et al., 2005. Saccharomyces cerevisiae adaptation to weak acids involves the transcription factor Haa1p and Haa1p-regulated genes. *Biochem Biophys Res Commun.*, 337(1), pp.95–103.

Fernandes, P., 2005. How does yeast respond to pressure? *Braz J Med Biol Res*, 38(2), pp.1239–1245.

Fiedurek, J., Skowronek, M. & Gromada, A., 2011. Selection and adaptation of Saccharomyces cerevisae to increased ethanol tolerance and production. *Pol J Microbiol.*, 60(1), pp.51–8.

Fleet, G.H., 2003. Yeast interactions and wine flavour. *Int J Food Microbiol*, 86, pp.11–22.

Flores, C.-L. et al., 2000. Carbohydrate and energy-yielding metabolism in non-conventional. *FEMS Microbiol Rev*, 24, pp.507–529.

Fuge, E.K., Braun, E.L. & Werner-Washburne, M, 1994. Protein synthesis in long-term stationary-phase cultures of Saccharomyces cerevisiae. *J Bacteriol*, 176(18), pp.5802–13.

Fujita, K et al., 2004. Comprehensive gene expression analysis of the response to straight-chain alcohols in Saccharomyces cerevisiae using cDNA microarray. *J Appl Microbiol.*, 97(1), pp.57–67.

Fujita, Katsuhide et al., 2006. The genome-wide screening of yeast deletion mutants to identify the genes required for tolerance to ethanol and other alcohols. *FEMS Yeast Res*, 6(5), pp.744–50.

Gale, E.F., 1946. The bacterial amino acid decarboxylases. *Adv Enzymol*, 6, pp.1–32.

Galhardo, R.S., Hastings, P.J. & Rosenberg, S.M., 2007. Mutation as a Stress Response and the Regulation of Evolvability. *Crit Rev Biochem Mol Biol*, 42(5), pp.399–435.

Galitski, T., 1999. Ploidy Regulation of Gene Expression. *Science*, 285(5425), pp.251–254.

Gao, C. et al., 2010. Global transcription engineering of brewer's yeast enhances the fermentation performance under high-gravity conditions. *Appl Microbiol Biotechnol*, 87(5), pp.1821–7.

Gao, X.D. et al., 2005. ERS1 encodes a functional homologue of the human lysosomal cystine transporter. *FEBS J*, 272, pp.2497–2511.

Garcia-Barrio, M. et al., 2000. Association of GCN1-GCN20 regulatory complex with the N-terminus of eIF2alpha kinase GCN2 is required for GCN2 activation. *EMBO J*, 19(8), pp.1887–99.

Gerrish, P J & Lenski, R E, 1998. The fate of competing beneficial mutations in an asexual population. *Genetica*, 102/103, pp.127–144.

Gerstein, A. et al., 2011. Haploids adapt faster than diploids across a range of environments. *J. Evol. Biol*, 24(3), pp.531–40.

Gerstein, A.C. et al., 2006. Genomic convergence toward diploidy in Saccharomyces cerevisiae. *PLoS Genet*, 2(9), p.e145.

Gerstein, A.C., McBride, R.M. & Otto, Sarah P, 2008. Ploidy reduction in Saccharomyces cerevisiae. *Biol Lett*, 4(1), pp.91–4.

Gerstein, A.C. & Otto, Sarah P, 2011. Cryptic fitness advantage: diploids invade haploid populations despite lacking any apparent advantage as measured by standard fitness assays. *PloS ONE*, 6(12), p.e26599.

Gerstein, A.C. & Otto, Sarah P, 2009. Ploidy and the causes of genomic evolution. *J Hered*, 100(5), pp.571–81.

Giaever, G. et al., 1999. Genomic profiling of drug sensitivities via induced haploinsufficiency. *Nat Genet*, 21(3), pp.278–83.

Gibney, P.A. et al., 2008. Rtr1 Is the Saccharomyces cerevisiae Homolog of a Novel Family of RNA Polymerase II-Binding Proteins. *Eukaryot Cell*, 7(6), pp.938–948.

Gilbert, A, Sangurdekar, D.P. & Srienc, F, 2009. Rapid strain improvement through optimized evolution in the cytostat. *Biotechnol Bioeng*, 103, pp.500–512.

Gilbert, Alan & Srienc, Friedrich, 2009. Optimized evolution in the cytostat: a Monte Carlo simulation. *Biotechnol Bioeng*, 102(1), pp.221–31.

Gilbert, D.G., 1980. Dispersal of yeasts and bacteria by drosophila in a temperate forest. *Oecologia*, 46(1), pp.135–137.

Goddard, M.R., 2008. Quantifying the complexities of saccharomyces cerevisiae's ecosystem engineering via fermentation. *Ecology*, 89(8), pp.2077–2082.

Goguel, V., Perea, J. & Jacq, C., 1989. Synthesis and function of the mitochondrial intron - encoded bI4 RNA maturase from Saccharomyces cerevisiae. *Curr Genet*, 16, pp.241–246.

Goldring, E., Grossman, L. & Marmur, J., 1971. Petite mutation in yeast. *J Bacteriol*, 107(1), pp.377–381.

Goossens, A. et al., 2001. The protein kinase Gcn2p mediates sodium toxicity in yeast. *J Biol Chem*, 276(33), pp.30753–60.

Görner, W. et al., 1998. Nuclear localization of the C2H2 zinc finger protein Msn2p is regulated by stress and protein kinase A activity. *Gene Dev*, 12(4), pp.586–97.

Gray, J. V et al., 2004. " Sleeping Beauty ": Quiescence in Saccharomyces cerevisiae. *Microbiol Mol Biol Rev*, 68(2), pp.187–206.

Greenbaum, L. et al., 2000. Green fluorescent protein photobleaching: a model for protein damage by endogenous and exogenous singlet oxygen. *Biol Chem*, 381(12), pp.1251–1258.

Greig, D, 2009. Reproductive isolation in Saccharomyces. *Heredity*, 102(1), pp.39–44.

Greig, Duncan, Borts, R.H., et al., 2002. Epistasis and hybrid sterility in Saccharomyces. *Proc. R. Soc. Lond. B*, 269(1496), pp.1167–71.

Greig, Duncan, Louis, E.J., et al., 2002. Hybrid speciation in experimental populations of yeast. *Science*, 298(5599), pp.1773–5.

Gresham, D. et al., 2008. The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet*, 4(12), p.e1000303.

Gross, V. & Smith, D., 1972. The effect of nalidixic acid on growth and petite formation in Saccharomyces cerevisiae. *Microbios*, 6(22), pp.139–46.

Grunstein, M., 1997. Histone acetylation in chromatin structure and transcription. *Nature*, 389, pp.349–352.

Guidi, F. et al., 2010. Effect of different glucose concentrations on proteome of Saccharomyces cerevisiae. *Biochim Biophys Acta.*, 1804(7), pp.1516–1525.

Gunge, N. & Nakatomi, Yasuo, 1972. Genetic mechanisms of rare matings of the yeast Saccharomyces cerevisiae heterozygous for mating type. *Genetics*, 70(January), pp.41–58.

Haber, J.E., 2012. Mating-Type Genes and MAT Switching in Saccharomyces cerevisiae. *Genetics*, 191(1), pp.33–64.

Haldane, J.B.S., 1933. The part played by recurrent mutations in evolution. *Am Nat*, 67, pp.5–19.

Hall, D.W. & Joseph, S.B., 2010. A high frequency of beneficial mutations across multiple fitness components in Saccharomyces cerevisiae. *Genetics*, 185(4), pp.1397–409.

Hamada, K., Nakatomi, Y & Shimada, S., 1992. Direct induction of tetraploids or homozygous diploids in the industrial yeast Saccharomyces cerevisiae by hydrostatic pressure. *Curr Genet*, 22(5), pp.371–6.

Harsch, M.J. et al., 2010. Optimized fermentation of grape juice by laboratory strains of Saccharomyces cerevisiae. *FEMS yeast research*, 10(1), pp.72–82.

Hazel, J.R. & Williams, E.E., 1990. The role of alterations in membrane lipid composition in enabling physiological adaptation of organisms to their physical environment. *Prog. Lipid Res.*, 29(3), pp.167–227.

Hebbes, T.R. & Thorne, A.W., 1988. A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J*, 7(5), pp.1395–1402.

Heckman, N.E. & Ramsay, J.O., 1998. *Penalised Regression with Model-based Penalties*,

Hegreness, M. et al., 2006. An Equivalence Principle for the Incorporation of Favorable Mutations in Asexual Populations. *Science*, 311, pp.1615–1617.

Hennaut, C., Hilger, F. & Grenson, M., 1970. Space limitation for permease insertion in the cytoplasmic membrane of Saccharomyces cerevisiae. *Biochem Biophys Res Commun*, 39(4), pp.666–671.

Heude, M. & Fabre, F., 1993. a/alpha-Control of DNA repair in the yeast Saccharomyces cerevisiae: genetic and physiological aspects. *Genetics*, 133, pp.489–498.

Hin, A., Tong, Y. & Boone, Charles, 2007. High-Throughput Strain Construction and Systematic Synthetic Lethal Screening in Saccharomyces cerevisiae. In M. Stansfield, I; Stark, ed. *Methods in Microbiology Volume: 36*. p. 369.

Hinnebusch, Alan G & Natarajan, K., 2002. Gcn4p , a Master Regulator of Gene Expression , Is Controlled at Multiple Levels by Diverse Signals of Starvation and Stress. *Eukaryot Cell*, 1(1), pp.22–32.

Hu, X. et al., 2007. Genetic dissection of ethanol tolerance in budding yeast S. cerevisiae. *Genetics*, 175, pp.1479–1487.

Hughes, T.R. et al., 2000. Widespread aneuploidy revealed by DNA microarray expression profiling. *Nat Genet*, 25(3), pp.333–337.

Hutner, S.H., Kaplan, H.M. & Enzmann, E. V, 1937. Chemicals attracting Drosophila. *Am Nat*, 71(734), p.193.

Hutter, a. & Oliver, S. G., 1998. Ethanol production using nuclear petite yeast mutants. *Appl Microbiol Biotechnol*, 49(5), pp.511–516.

Ibeas, J I & Jimenez, J, 1997. Mitochondrial DNA loss caused by ethanol in Saccharomyces flor yeasts. *Appl Environ Microbiol*, 63, pp.7–12.

Ibeas, Jose Ignacio & Jimenez, Juan, 1997. Mitochondrial DNA Loss Caused by Ethanol in Saccharomyces Flor Yeasts. *Appl Environ Microbiol*, 63(1), pp.7–12.

Illingworth, C.J.R. & Mustonen, V., 2011. A method to infer positive selection from marker dynamics in an asexual population. *Bioinformatics*, pp.1–7.

Inoue, T et al., 2000. Cloning and characterization of a gene complementing the mutation of an ethanol-sensitive mutant of sake yeast. *Biosci. Biotechnol. Biochem.*, 64(2), pp.229–36.

Izawa, S. et al., 2005. Characterization of Rat8 localization and mRNA export in Saccharomyces cerevisiae during the brewing of Japanese sake. *Appl Microbiol Biotechnol*, 69(1), pp.86–91.

Izawa, S. et al., 2007. Formation of Cytoplasmic P-Bodies in Sake Yeast during Japanese Sake Brewing and Wine Making. *Biosci. Biotechnol. Biochem.*, 71(11), pp.2800–2807.

Izawa, S. et al., 2008. Heat shock and ethanol stress provoke distinctly different responses in 3'-processing and nuclear export of HSP mRNA in Saccharomyces cerevisiae. *Biochem J.*, 414(1), pp.111–9.

Jain, K. & Krug, J., 2007. Deterministic and stochastic regimes of asexual evolution on rugged fitness landscapes. *Genetics*, 175, pp.1275–88.

Jansen, M.L.A. et al., 2004. Prolonged maltose-limited cultivation of Saccharomyces cerevisiae selects for cells with improved maltose affinity and hypersensitivity. *Appl Environ Microbiol*, 70(4), p.1956.

Jansen, M.L.A. et al., 2005. Prolonged selection in aerobic, glucose-limited chemostat cultures of Saccharomyces cerevisiae causes a partial loss of glycolytic capacity. *Microbiology*, 151(Pt 5), pp.1657–69.

Jazwinski, S.M., 1990. Aging and senescence of the budding yeast Saccharomyces cerevisiae. *Molecular microbiology*, 4(3), pp.337–43.

Jiménez, J. & Benitez, T, 1987. Adaptation of Yeast Cell Membranes to Ethanol. *Appl Environ Microbiol*, 53(5), pp.1196–1198.

Jiménez, J. & Benitez, Tahia, 1988. Selection of Ethanol-Tolerant Yeast Hybrids in pH-Regulated Continuous Culture. *Appl Environ Microbiol*, 54(4), pp.917–922.

Jiménez, J. & Van Uden, N., 1985. Use of extracellular acidification for the rapid testing of ethanol tolerance in yeasts. *Biotechnol Bioeng*, 27, pp.1596–1598.

Jimenez, Juan & Benitez, Tahia, 1988. Yeast cell viability under conditions of high temperature and ethanol concentrations depends on the mitochondrial genome. *Curr Genet*, 13, pp.461–469.

Joanes, D.N. & Gill, C. a., 1998. Comparing measures of sample skewness and kurtosis. *Statistician*, 47(1), pp.183–189.

Johnston, G C et al., 1979. Regulation of Cell Size in the Yeast Saccharomyces cerevisiae. *J. Bacteriol.*, 137(1), pp.1–5.

Jones, L.J., Carballido-López, R. & Errington, J., 2001. Control of cell shape in bacteria: helical, actin-like filaments in Bacillus subtilis. *Cell*, 104(6), pp.913–22.

Jones, R.P. & Greenfield, P.F., 1987. Ethanol and the fluidity of the yeast plasma membrane. *Yeast*, 3(4), pp.223–32.

Jorgensen, P. et al., 2002. Systematic identification of pathways that couple cell growth and division in yeast. *Science*, 297(5580), pp.395–400.

Jorgensen, P. et al., 2007. The size of the nucleus increases as yeast cells grow. *Mol Biol Cell*, 18(9), p.3523.

Joseph, S.B. & Hall, D.W., 2004. Spontaneous mutations in diploid Saccharomyces cerevisiae: more beneficial than expected. *Genetics*, 168(4), pp.1817–25.

Kajiwara, S. et al., 2000. Improved ethanol tolerance of Saccharomyces cerevisiae strains by increases in fatty acid unsaturation via metabolic engineering. *Biotechnol Lett*, 22, pp.1839–1843.

Kang, J.Y. et al., 2005. PIG-V involved in transferring the second mannose in glycosylphosphatidylinositol. *J Biol Chem*, 280(10), pp.9489–97.

Kao, K.C. & Sherlock, G., 2009. Molecular Characterization of Clonal Interference during Adaptive Evolution in Asexual Populations of Saccharomyces cerevisiae. *Nat Genet.*, 40(12), pp.1499–1504.

Katou, T. et al., 2008. Brewing characteristics of haploid strains isolated from sake yeast Kyokai No . 7. *Yeast*, 25, pp.799–807.

Katz Ezov, T. et al., 2010. Heterothallism in Saccharomyces cerevisiae isolates from nature: effect of HO locus on the mode of reproduction. *Mol Ecol*, 19(1), pp.121–31.

Kaushik, J.K. & Bhat, R., 2003. Why is trehalose an exceptional protein stabilizer? An analysis of the thermal stability of proteins in the presence of the compatible osmolyte trehalose. *J Biol Chem*, 278(29), pp.26458–65.

Kawahata, M. et al., 2006. Yeast genes involved in response to lactic acid and acetic acid: acidic conditions caused by the organic acids in Saccharomyces cerevisiae cultures induce expression of intracellular metal metabolism genes regulated by Aft1p. *FEMS Yeast Res*, 6(6), pp.924–36.

Kell, D B et al., 1998. Viability and activity in readily culturable bacteria: a review and discussion of the practical issues. *Antonie van Leeuwenhoek*, 73(2), pp.169–87.

Kellis, M., Birren, B.W. & Lander, E.S., 2004. Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. *Nature*, 428, pp.617–624.

Kent, W.J. et al., 2002. The Human Genome Browser at UCSC. *Genome Res*, 12(6), pp.996–1006.

Kim, J. et al., 1996. Disruption of the yeast ATH1 gene confers better survival after dehydration, freezing, and ethanol shock: potential commercial applications. *Appl Environ Microbiol*, 62, pp.1563–1569.

Kitagaki, H. et al., 2007. Ethanol-induced death in yeast exhibits features of apoptosis mediated by mitochondrial fission pathway. *FEBS Lett*, 581(16), pp.2935–42.

Klein, F. & Wintersberger, Ulrike, 1988. Determination of mating-type conversion rates of heterothallic Saccharomyces cerevisiae with the fluctuation assay. *Curr Genet*, 14(4), pp.355–362.

Koch, C. et al., 1999. A role for Ctr9p and Paf1p in the regulation G1 cyclin expression in yeast. *Nucleic Acids Res*, 27(10), pp.2126–34.

Kovacech, B., Nasmyth, K. & Schuster, T., 1996. EGT2 gene transcription is induced predominantly by Swi5 in early G1. *Mol Cell Biol*, 16(7), pp.3264–3274.

Kozarewa, I. et al., 2009. Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of GC-biased genomes. *Nat Methods*, 6(4), pp.291–295.

Kren, A. et al., 2003. War1p , a Novel Transcription Factor Controlling Weak Acid Stress Response in Yeast. *Mol Cell Biol*, 23(5), pp.1775–1785.

Kresnowati, M.T.A.P. et al., 2006. When transcriptome meets metabolome: fast cellular responses of yeast to sudden relief of glucose limitation. *Mol Syst Biol*, 2, p.49.

Kron, S.J., 1997. Filamentous growth in budding yeast. *Trends Microbiol.*, 5(11), pp.450–454.

Kubota, H. et al., 2001. Budding yeast GCN1 binds the GI domain to activate the eIF2alpha kinase GCN2. *J Biol Chem*, 276(20), pp.17591–6.

Kubota, S. et al., 2004. Effect of ethanol on cell growth of budding yeast: genes that are important for cell growth in the presence of ethanol. *Biosci. Biotechnol. Biochem.*, 68(4), pp.968–72.

Kuo, M. et al., 2000. Gcn4 Activator Targets Gcn5 Histone Acetyltransferase to Specific Promoters Independently of Transcription. *Mol Cell*, 6, pp.1309–1320.

Lachance, M. a et al., 2001. Biogeography of the yeasts of ephemeral flowers and their insects. *FEMS Yeast Res*, 1(1), pp.1–8.

Lahr, D.J.G. et al., 2011. The chastity of amoebae: re-evaluating evidence for sex in amoeboid organisms. *Proc. R. Soc. B*, 278(1715), pp.2081–90.

Laland, K.N., Odling-Smee, J. & Feldman, M.W., 2004. Causing a commotion. *Nature*, 429(10), p.609.

Lang, G.I., Botstein, D. & Desai, M.M., 2011. Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics*, 188(3), pp.647–61.

Lau, W.W., Schneider, K.R. & Shea, E.K.O., 1998. A Genetic Study of Signaling Processes for Repression of PHO5 Transcription in Saccharomyces cerevisiae. *Genetics*, 150(4), pp.1349–59.

Leao, C. & Van Uden, N., 1982. Effects of ethanol and other alkanols on the glucose transport system of Saccharomyces cerevisiae. *Biotechnol Bioeng.*, 24(11), pp.2601–2604.

Lenski, R E et al., 1998. Evolution of competitive fitness in experimental populations of E. coli: what makes one genotype a better competitor than another? *Antonie van Leeuwenhoek*, 73(1), pp.35–47.

Lenski, Richard E et al., 1991. Long-Term Experimental Evolution In Escherichia coli. I. Adaptation and Divergence During 2,000 Generations. *American Nat*, 138(6), pp.1315–1341.

Lenski, Richard E. & Mongold, J.A., 2000. Cell size, shape, and fitness in evolving populations of bacteria. In J. H. Brown & G. B. West, eds. *Scaling in Biology*. Oxford: Oxford University Press, pp. 221–235.

Lew, D J & Reed, S.I., 1995. A cell cycle checkpoint monitors cell morphogenesis in budding yeast. *J Cell Biol*, 129(3), pp.739–49.

Li, B.-Z. et al., 2010. Transcriptome analysis of differential responses of diploid and haploid yeast to ethanol stress. *J Biotechnol*, 148(4), pp.194–203.

Li, H. et al., 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, 25(16), pp.2078–9.

Li, H. & Durbin, R., 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 26(5), pp.589–95.

Li, H., Ruan, J. & Durbin, R., 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res*, 18, pp.1851–1858.

Li, Q. et al., 2012. Ethanol-induced yeast flocculation by the promoter of TPS1 encoding trehalose-6-phospahte synthase 2 fro efficient ethanol production. *Metab Eng*, 14(1), pp.1–8.

Liebman, S., Shalit, P. & Picologlou, S., 1981. Ty elements are involved in the formation of deletions in DEL1 strains of Saccharomyces cerevisiae. *Cell*, 26, pp.401–409.

Lillie, S.H. & Brown, S.S., 1994. Immunofluorescence localization of the unconventional myosin, Myo2p, and the putative kinesin-related protein, Smy1p, to the same regions of polarized growth in Saccharomyces cerevisiae. *J Cell Biol*, 125(4), pp.825–42.

Lillie, S.H. & Pringle, J.R., 1980. Reserve carbohydrate metabolism in Saccharomyces cerevisiae: responses to nutrient limitation. *J Bacteriol*, 143(3), pp.1384–94.

Lin, J. et al., 1996. Mechanisms of acid resistance in enterohemorrhagic Escherichia coli. *Appl Environ Microbiol*, 62(9), pp.3094–100.

Loewith, R et al., 2000. Three yeast proteins related to the human candidate tumor suppressor p33(ING1) are associated with histone acetyltransferase activities. *Mol Cell Biol*, 20(11), pp.3807–16.

Loewith, Robbie et al., 2001. Pho23 Is Associated with the Rpd3 Histone Deacetylase and Is Required for Its Normal Function in Regulation of Gene Expression and Silencing in Saccharomyces cerevisiae. *J Biol Chem*, 276(26), pp.24068–24074.

Lorincz, A. & Carter, B.L.A., 1979. Control of Cell Size at Bud Initiation in Saccharomyces cerevisiae. *J Gen Microbiol*, 113(2), pp.287–295.

Lucas, C. & Van Uden, N., 1985. The temperature profiles of growth, thermal death and ethanol tolerance of the xylose-fermenting yeast Candida shehatae. *J. Basic Microbiol.*, 8, pp.547–550.

Lynch, M. et al., 2008. A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proc Natl Acad Sci U S A.*, 105(27), pp.9272–7.

Mable, B K & Otto, S P, 2001. Masking and purging mutations following EMS treatment in haploid, diploid and tetraploid yeast (Saccharomyces cerevisiae). *Genet. Res. Camb.*, 77(1), pp.9–26.

Mable, B. K., 2001. Ploidy evolution in the yeast Saccharomyces cerevisiae: a test of the nutrient limitation hypothesis. *J.Evol.Biol.*, 14(1), pp.157–170.

MacLean, R.C. & Gudelj, I., 2006. Resource competition and social conflict in experimental populations of yeast. *Nature*, 441(7092), pp.498–501.

Malik, T. & Smith, H.L., 2008. Does dormancy increase fitness of bacterial populations in time-varying environments? *Bull Math Biol*, 70(4), pp.1140–62.

Manchester, S., Collinson, M. & Goth, K., 1994. Fruits of the juglandaceae from the Eocene of Messel, Germany and implications for early tertiary phytogeographic exchange between Europe and Western North America. *Int.J.Plant Sci.*, 155(3), pp.388–394.

Mansure, J. et al., 1994. Trehalose inhibits ethanol effects on intact yeast cells and liposomes. *Biochim Biophys Acta.*, 1191(2), pp.309–316.

Margalith, P.Z., 1981. *Flavor Microbiology*, Springfield, Ill: Thomas.

Marks, V.D. et al., 2008. Dynamics of the yeast transcriptome during wine fermentation reveals a novel fermentation stress response. *FEMS Yeast Res*, 8(1), pp.35–52.

Martinez, M. & Roy, S., 2004. Genomic analysis of stationary-phase and exit in Saccharomyces cerevisiae: gene expression and identification of novel essential genes. *Mol Biol Cell.*, 15(December), pp.5295–5305.

Marton, H.A. & Desiderio, S., 2008. The Paf1 complex promotes displacement of histones upon rapid induction of transcription by RNA polymerase II. *BMC Mol Biol*, 9, p.4.

Marton, M.J., Crouch, D. & Hinnebusch, a G., 1993. GCN1, a translational activator of GCN4 in Saccharomyces cerevisiae, is required for phosphorylation of eukaryotic translation initiation factor 2 by protein kinase GCN2. *Mol Cell Biol*, 13(6), pp.3541–56.

Matsuda, T. et al., 1998. Specific tandem GG to TT base substitutions induced by acetaldehyde are due to intra-strand crosslinks between adjacent guanine bases. *Nucleic Acids Res*, 26(7), pp.1769–74.

Mayer, V.W. & Aguilera, A., 1990. High levels of chromosome instability in polyploids of Saccharomyces cerevisiae. *Mutat Res*, 231(2), pp.177–86.

McKenna, A. et al., 2010. The Genome Analysis Toolkit : A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*, pp.1297–1303.

McNulty, J.J. & Lew, Daniel J, 2005. Swe1p responds to cytoskeletal perturbation, not bud size, in S. cerevisiae. *Curr Biol*, 15(24), pp.2190–8.

Meaden, P.G. et al., 1999. Endocytosis and Vacuolar Morphology in Saccharomyces cerevisiae Are Altered in Response to Ethanol Stress or Heat Shock. *Yeast*, 15, pp.1211–1222.

Medawar, W., Strehaiano, P. & Délia, M.-L., 2003. Yeast growth: lag phase modelling in alcoholic media. *Food Microbiol*, 20(5), pp.527–532.

Merico, A. et al., 2007. Fermentative lifestyle in yeasts belonging to the Saccharomyces complex. *FEBS J*, 274(4), pp.976–89.

Miller, C.R., Joyce, P. & Wichman, H., 2011. Mutational effects and population dynamics during viral adaptation challenge current models. *Genetics*, 187, pp.185–202.

Mills, D.A., Johannsen, E.A. & Cocolin, L., 2002. Yeast diversity and persistence in botrytis-affected wine fermentations. *Appl. Environ. Microbiol.*, 68(10), pp.4884–4893.

Mira, N P, Teixeira, M C & Sá-Correia, I, 2010. Adaptive response and tolerance to weak acids in Saccharomyces cerevisiae: a genome-wide view. *OMICS*, 14(5), pp.525–40.

Monteiro, G.A., Supply, P. & Goffeau, A., 1994. The in vivo Activation of Saccharomvces cerevisiae Plasma Membrane H+-ATPase by Ethanol depends on the Expression of the PMAI Gene, but not of the PMA2 Gene. *Yeast*, 10, pp.1439–1446.

Montooth, K.L., Siebenthall, K.T. & Clark, A.G., 2006. Membrane lipid physiology and toxin catabolism underlie ethanol and acetic acid tolerance in Drosophila melanogaster. *J Exp Biol*, 209(Pt 19), pp.3837–50.

Mortimer, R.K. & Johnston, J.R., 1986. Genealogy of principal strains of the yeast genetic stock center. *Genetics*, 113(1), pp.35–43.

Nakamura, Kensuke et al., 2011. Sequence-specific error profile of Illumina sequencers. *Nucleic Acids Res*, 39(13), pp.1–13.

Nei, M., 2005. Selectionism and Neutralism in Molecular Evolution. *Mol Biol Evol*, 22(12), pp.2318–42.

Ogawa, Y. et al., 2000. Tolerance mechanism of the ethanol-tolerant mutant of sake yeast. *J Biosci Bioeng*, 90(3), pp.313–20.

Ohno, S., 1970. *Evolution by Gene Duplication*, London: George Allen and Unwin.

Ormo, M., Cubitt, A. & Kallio, K., 1996. Crystal structure of the Aequorea victoria green fluorescent protein. *Science*, 273, pp.1392–1395.

Orr, H.A., 2000a. Adaptation and the cost of complexity. *Evolution*, 54(1), pp.13–20.

Orr, H.A., 1995. Somatic mutation favors the evolution of diploidy. *Genetics*, 139(3), pp.1441–7.

Orr, H.A., 2003. The distribution of fitness effects among beneficial mutations. *Genetics*, 163(4), p.1519.

Orr, H.A., 2005. The genetic theory of adaptation: a brief history. *Nat Rev Genet*, 6(2), pp.119–127.

Orr, H.A., 1998. The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. *Evolution*, 52(4), pp.935–949.

Orr, H.A., 2000b. The rate of adaptation in asexuals. *Genetics*, 155(2), pp.961–8.

Otterstedt, K. et al., 2004. Switching the mode of metabolism in the yeast Saccharomyces cerevisiae. *EMBO Rep.*, 5(5), pp.532–7.

Palmqvist, E. & Hahn-Hägerdal, B., 2000. Fermentation of lignocellulosic hydrolysates. I: inhibition and detoxification. *Bioresour Technol.*, 74(1), pp.17–24.

Paquin, C.E. & Adams, J., 1983. Frequency of fixation of adaptive mutations is higher in evolving diploid than haploid yest populations. *Nature*, 302, pp.495–500.

Park, M. et al., 2014. Comprehensive Analysis to Improve the Validation Rate for Single Nucleotide Variants Detected by Next- Generation Sequencing. *Plos ONE*, 9(1), pp.1–9.

Pastor, N., 2003. Induction of endoreduplication by topoisomerase II catalytic inhibitors. *Mutagenesis*, 18(2), pp.105–112.

Pavelka, N. et al., 2010. Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature*, 468(7321), pp.321–5.

Pel, H.J. & Grivell, L. a, 1993. The biology of yeast mitochondrial introns. *Mol Biol Rep*, 18(1), pp.1–13.

Peters, J.-M., 2006. The anaphase promoting complex/cyclosome: a machine designed to destroy. *Nat Rev Mol Cell Biol*, 7(9), pp.644–56.

Phaff, H.J., Miller, M.W. & Shifrine, M., 1956. The taxonomy of yeasts isolated from Drosophila in the Yosemite region of California. *Antonie Van Leeuwenhoek*, 22(1), pp.145–161.

Pinheiro, J. & Bates, Douglas, Self-Starting NlS Asymptotic Regression Model. Available at:

http://stat.ethz.ch/R-manual/R-patched/library/stats/html/SSasymp.html.

Piper, P W, 1995. The heat shock and ethanol stress responses of yeast exhibit extensive similarity and functional overlap. *FEMS Microbiol Lett*, 134, pp.121–127.

Piper, P. et al., 1998. The Pdr12 ABC transporter is required for the development of weak organic acid resistance in yeast. *EMBO J*, 17(15), pp.4257–4265.

Piper, P., Calderon, C.O. & Hatzixanthis, K., 2001. Weak acid adaptation : the stress response that confers yeasts with resistance to organic acid food preservatives. *Microbiology*, 147, pp.2635–2642.

Piskur, J et al., 1998. Structure and genetic stability of mitochondrial genomes vary among yeasts of the genus Saccharomyces. *Int J Syst Bacteriol*, 48, pp.1015–1024.

Piskur, Jure et al., 2006. How did Saccharomyces evolve to become a good brewer? *Trends Genet.*, 22(4), pp.183–186.

Piskur, Jure, 1994. Inheritance of the yeast mitochondrial genome. *Plasmid*, 31, pp.229–241.

Porro, D., Brambilla, L. & Alberghina, Lilia, 2003. Glucose metabolism and cell size in continuous cultures of Saccharomyces cerevisiae. *FEMS Microbiol Lett*, 229(2), pp.165–171.

Pouwels, L. et al., 2008. Kinetic isotope effect studies on the de novo rate of chromophore formation in fast- and slow-maturing GFP variants. *Biochemistry*, 47(38), pp.10111–10122.

Pretorius, I.S., 2000. Tailoring wine yeast for the new millennium: Novel approaches to the ancient art of winemaking. *Yeast*, 16, pp.675–729.

Puria, R. et al., 2009. Critical role of RPI1 in the stress tolerance of yeast during ethanolic fermentation. *FEMS Yeast Res*, 11, pp.1–11.

Ram, Y. & Hadany, L., 2012. the Evolution of Stress-Induced Hypermutation in Asexual Populations. *Evolution*, 66, pp.2315–28.

Ramsey, J. & Ripley, B., 2009. Package " pspline ". Available at: http://cran.r-project.org/web/packages/pspline/index.html.

Razeto-Barry, P. et al., 2011. Molecular Evolution, Mutation Size and Gene Pleiotropy: A Geometric Reexamination. *Genetics*, 187(2006), pp.877–885.

Reed, M., 1938. The olfactory reactions of Drosophila melanogaster meigen to the products of fermenting banana. *Physiol Zool.*, 11(3), pp.317–325.

Remington, S James, 2006. Fluorescent proteins : maturation , photochemistry and photophysics. *Curr Opin Struct Biol*, 16, pp.714–721.

Reuter, M., Bell, G. & Greig, Duncan, 2007. Increased outbreeding in yeast in response to dispersal by an insect vector. *Curr Biol*, 17(3), pp.81–83.

Rho, S.B. & Martinis, S.A., 2000. The bI4 group I intron binds directly to both its protein splicing partners , a tRNA synthetase and maturase , to facilitate RNA splicing activity The bI4 group I intron binds directly to both its protein splicing partners , a tRNA synthetase and maturase. *RNA*, 6, pp.1882–1894.

Ribereau-Gayon, P. et al., 2006. *Handbook of Enology* 2nd Ed., Chichester, UK: John Wiley and Sons.

Rieseberg, M. et al., 2001. Flow cytometry in biotechnology. *Appl Microbiol Biotechnol*, 56, pp.350–360.

Ristow, H., Seyfarth, A. & Lochmann, E.-R., 1995. Chromosomal damages by ethanol and acetaldehyde in Saccharomyces cerevisiae as studied by pulsed field gel electrophoresis. *Mutat Res.*, 326, pp.165–170.

Rizzoni, M. & Palitti, F., 1973. Regulatory mechanism of cell division. I. Colchicine-induced endoreduplication. *Exp. Cell Res.*, 77(450-458).

Robinson, J.T. et al., 2011. Integrative genomics viewer. *Nat Biotechnol.*, 29(1), pp.24–26.

Rokyta, D.R. et al., 2008. Beneficial fitness effects are not exponential for two viruses. *J Mol Evol.*, 67(4), pp.368–376.

Rolfes, R.J. & Hinnebusch, A G, 1993. Translation of the yeast transcriptional activator GCN4 is stimulated by purine limitation: implications for activation of the protein kinase GCN2. *Mol Cell Biol*, 13(8), pp.5099–111.

Rolland, F., Winderickx, J. & Thevelein, J.M., 2002. Glucose-sensing and -signalling mechanisms in yeast. *FEMS Yeast Res.*, 2(2), pp.183–201.

Roman, H. & Sands, S.M., 1953. Heterogeneity of Clones of Saccharomyces Derived from Haploid Ascospores. *Proc Natl Acad Sci U S A*, 39(3), pp.171–9.

Rosa, M.F. & Sá-Correia, I, 1991. In vivo activation by ethanol of plasma membrane ATPase of Saccharomyces cerevisiae. *Appl Environ Microbiol*, 57, pp.830–835.

Rozen, D.E., De Visser, J. & Gerrish, P.J., 2002. Fitness effects of fixed beneficial mutations in microbial populations. *Curr. Biol.*, 12(12), pp.1040–1045.

Rozen, Daniel E et al., 2007. Fitness costs of fluoroquinolone resistance in Streptococcus pneumoniae. *Antimicrob Agents Chemother*, 51(2), pp.412–6.

Rozen, Daniel E et al., 2008. Heterogeneous adaptive trajectories of small populations on complex fitness landscapes. *PLoS One.*, 3(3), p.e1715.

Rupeš, I., 2002. Checking cell size in yeast. *Trends Genet*, 18(9), pp.479–485.

Sa-correia, I. & Van Uden, N., 1983. Temperature profiles of ethanol tolerance: effects of ethanol on the minimum and the maximum temperatures for growth of the yeasts Saccharomyces cerevisiae and Kluyveromyces fragilis. *Biotechnol Bioeng.*, 25(6), pp.1665–1667.

Salathe, M., Ackermann, M. & Bonhoeffer, S., 2006. The Effect of Multifunctionality on the Rate of Evolution in Yeast. *Mol Biol Evol*, 23(4), pp.7221–722.

Sanchez, Y. et al., 1992. Hsp104 is required for tolerance to many forms of stress. *EMBO J*, 11, pp.2357–2364.

Santos, J. et al., 2008. Ethanol tolerance of sugar transport, and the rectification of stuck wine fermentations. *Microbiology*, 154(Pt 2), pp.422–30.

Sato, M et al., 2002. Analysis of an inactivated Lg-FLO1 gene present in bottom-fermenting yeast. *J Biosci Bioeng*, 93(4), pp.395–8.

Sato, M, Watari, J & Shinotsuka, K., 2001. Genetic instability in flocculation of bottom-fermenting yeast. *J Am Soc Brew Chem*, 59, pp.130–134.

Scannell, D.R. et al., 2011. The Awesome Power of Yeast Evolutionary Genetics: New Genome Sequences and Strain Resources for the Saccharomyces sensu stricto Genus. *G3*, 1(1), p.11.

Scannell, Devin R, Butler, G. & Wolfe, Kenneth H, 2007. Yeast genome evolution — the origin of the species. *Yeast*, 24, pp.929–942.

Schiestl, R. & Wintersberger, U, 1983. Induction of mating type interconversion in a heterothallic strain of Saccharomyces cerevisiae by DNA damaging agents. *Mol Gen Genet*, 191(1), pp.59–65.

Schiestl, R. & Wintersberger, U, 1982. X-ray enhances mating type switching in heterothallic strains of Saccharomyces cerevisiae. *Mol Gen Genet*, 186(4), pp.512–7.

Schoustra, S.E. et al., 2009. The properties of adaptive walks in evolving populations of fungus. *PLoS Biol*, 7(11), p.e1000250.

Schüller, H.-J., 2003. Transcriptional control of nonfermentative metabolism in the yeast Saccharomyces cerevisiae. *Current genetics*, 43(3), pp.139–60.

Di Segni, G. et al., 2011. Yeast pheromone receptor genes STE2 and STE3 are differently regulated at the transcription and polyadenylation level. *Proc Natl Acad Sci U S A*, 108(41), pp.17082–6.

Sena, E., Welch, J. & Fogel, S., 1976. Nuclear and mitochondrial DNA replication during zygote formation and maturation in yeast. *Science*, 194(4263), pp.433–5.

Sharma, S.C., 1997. A possible role of trehalose in osmotolerance and ethanol tolerance in Saccharomyces cerevisiae. *FEMS Microbiol Lett*, 152(1), pp.11–5.

Shoemaker, D., Lashkari, D. & Morris, D., 1996. Quantitative phenotypic analysis of yeast deletion mutants using a highly parallel molecular bar–coding strategy. *Nat Genet*, 14(4), pp.450–6.

Sia, R.A.L., Urbonas, B.L. & Sia, E.A., 2003. Effects of ploidy, growth conditions and the mitochondrial nucleoid-associated protein Ilv5p on the rate of mutation

of mitochondrial DNA in Saccharomyces cerevisiae. *Curr Genet*, 44(1), pp.26–37.

Slack, A. et al., 2006. On the mechanism of gene amplification induced under stress in Escherichia coli. *PLoS genetics*, 2(4), p.e48.

Smith, M., 1987. Molecular evolution of the Saccharomyces cerevisiae histone gene loci. *J Mol Evol*, 24(3), pp.252–259.

Smukalla, S. et al., 2008. FLO1 is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell*, 135(4), pp.726–737.

Snoek, I.S.I. & Steensma, H.Y., 2007. Factors involved in anaerobic growth of Saccharomyces cerevisiae. *Yeast*, 24, pp.1–10.

Sobieski, R.J. & Brewer, A.R., 1976. Toxicity of nalidixic acid on candida albicans, Saccharomyces cerevisiae, and Kluyveromyces lactis. *Antimicrob Agents Chemother*, 9(3), pp.485–92.

Stanley, Grant A et al., 2010. Generation and characterisation of stable ethanol-tolerant mutants of Saccharomyces cerevisiae. *J Ind Microbiol Biotechnol*, 37(2), pp.139–49.

Stanley, Grant A et al., 2009. Trehalose promotes the survival of Saccharomyces cerevisiae during lethal ethanol stress, but does not influence growth under sublethal ethanol stress. *FEMS Yeast Res.*, 1, pp.1208–1216.

Steffen, K. et al., 2008. Yeast life span extension by depletion of 60s ribosomal subunits is mediated by Gcn4. *Cell*, 133(2), pp.292–302.

Steinskog, D.J., Tjøstheim, D.B. & Kvamstø, N.G., 2007. A Cautionary Note on the Use of the Kolmogorov–Smirnov Test for Normality. *Monthly Weather Review*, 135(3), pp.1151–1157.

Storchová, Z. et al., 2006. Genome-wide genetic analysis of polyploidy in yeast. *Nature*, 443(7111), pp.541–7.

Suh, S. et al., 2005. The beetle gut: a hyperdiverse source of novel yeasts. *Mycol. Res.*, 109(3), pp.261–265.

Sun, Z. & Hampsey, M., 1999. A General Requirement for the Sin3-Rpd3 Histone Deacetylase Complex in Regulating Silencing in Saccharomyces cerevisiae. *Genetics*, 152(3), pp.921–932.

Suzuki, T. et al., 2012. Lactic-acid stress causes vacuolar fragmentation and impairs intracellular amino-acid homeostasis in Saccharomyces cerevisiae. *JBIOSC*, 113(4), pp.421–430.

Swaminathan, R., Hoang, C.P. & Verkman, a S., 1997. Photobleaching recovery and anisotropy decay of green fluorescent protein GFP-S65T in solution and cells:

cytoplasmic viscosity probed by green fluorescent protein translational and rotational diffusion. *Biophysical journal*, 72(4), pp.1900–7.

Swan, T.M. & Watson, K., 1998. Stress tolerance in a yeast lipid mutant: membrane lipids influence tolerance to heat and ethanol independently of heat shock proteins and trehalose. *Can J Microbiol*, 45, pp.472–479.

Swinnen, S. et al., 2012. Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis. *Genome Res*, 22(5), pp.975–84.

Teixeira, Miguel C et al., 2009. Genome-wide identification of Saccharomyces cerevisiae genes required for maximal tolerance to ethanol. *Appl Environ Microbiol.*, 75(18), pp.5761–72.

Tenreiro, S. et al., 2000. Expression of the AZR1 gene (ORF YGR224w), encoding a plasma membrane transporter of the major facilitator superfamily, is required for adaptation to acetic acid and resistance to azoles in Saccharomyces cerevisiae. *Yeast*, 16, pp.1469–1481.

Teste, M.A. et al., 2000. The Saccharomyces cerevisiae YPR184w gene encodes the glycogen debranching enzyme. *FEMS Microbiol Lett*, 193(1), pp.105–10.

Thomson, J.M. et al., 2005. Resurrecting ancestral alcohol dehydrogenases from yeast. *Nat Genet.*, 37(6), pp.630–5.

Tong, A.H. et al., 2001. Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science*, 294(5550), pp.2364–8.

Toulmay, A. & Schneiter, R., 2007. Lipid-dependent surface transport of the proton pumping ATPase: a model to study plasma membrane biogenesis in yeast. *Biochimie*, 89(2), pp.249–54.

Urbanczyk, H. et al., 2011. Sake yeast strains have difficulty in entering a quiescent state after cell growth cessation. *J Biosci Bioeng*, 112(1), pp.44–8.

Van Urk, H. et al., 1989. Glucose transport in crabtree-positive and crabtree-negative yeasts. *J Gen Microbiol*, 135(9), pp.2399–406.

Vasserot, Y., Mornet, F. & Jeandet, P., 2010. Acetic acid removal by Saccharomyces cerevisiae during fermentation in oenological conditions. Metabolic consequences. *Food Chem.*, 119(3), pp.1220–1223.

Vilela-Moura, A. et al., 2010. Effects of acetic acid, ethanol, and SO(2) on the removal of volatile acidity from acidic wines by two Saccharomyces cerevisiae commercial strains. *Appl Microbiol Biotechnol*, 87(4), pp.1317–26.

Visser, W. et al., 1990. Oxygen requirement of yeasts. *Appl Environ Microbiol*, 56, pp.3785–3792.

Volland, C. et al., 1994. Endocytosis and degradation of the yeast uracil permease under adverse conditions. *J Biol Chem*, 269(13), pp.9833–41.

Van Voorst, F. et al., 2006. Genome-wide identification of genes required for growth of Saccharomyces cerevisiae under ethanol stress. *Yeast*, 23(5), pp.351–9.

Wachter, R.M. et al., 1998. Structural basis of spectral shifts in the yellow-emission variants of green fluorescent protein. *Structure*, 6(10), pp.1267–77.

Wachter, R.M. & Remington, S J, 1999. Sensitivity of the yellow variant of green fluorescent protein to halides and nitrate. *Curr Biol.*, 9(17), pp.R628–9.

Wagner, A., 2005. Energy constraints on the evolution of gene expression. *Mol Biol Evol*, 22, pp.1365–1374.

Walker-Caprioglio, H.M. & Parks, L.W., 1987. Autoconditioning factor relieves ethanol-induced growth inhibition of Saccharomyces cerevisiae. *Appl Environ Microbiol.*, 53(1), pp.33–5.

Walker-Caprioglio, H.M., Rodriguez, R.J. & Parks, L.W., 1985. Recovery of Saccharomyces cerevisiae from ethanol-induced growth inhibition. *Appl Environ Microbiol.*, 50(3), pp.685–9.

Walkey, C.J. et al., 2011. The Saccharomyces cerevisiae fermentation stress response protein Igd1p/Yfr017p regulates glycogen levels by inhibiting the glycogen debranching enzyme. *FEMS yeast research*, 11(6), pp.499–508.

Watanabe, M. et al., 2009. Overexpression of MSN2 in a sake yeast strain promotes ethanol tolerance and increases ethanol production in sake brewing. *J Biosci Bioeng*, 107(5), pp.516–8.

Weiss, R.L., Kukora, J.R. & Adams, J., 1975. The Relationship Between Enzyme Activity, Cell Geometry, and Fitness in Saccharomyces cerevisiae. *Proc. Nat. Acad. Sci. USA*, 72(3), pp.794–798.

Westerbeek-Marres, C. a, Moore, M.M. & Autor, a P., 1988. Regulation of manganese superoxide dismutase in Saccharomyces cerevisiae. The role of respiratory chain activity. *Eur.J.Biochem*, 174(4), pp.611–20.

Wilcox, L.J. et al., 2002. Transcriptional profiling identifies two members of the ATP-binding cassette transporter superfamily required for sterol uptake in yeast. *J Biol Chem.*, 277(36), pp.32466–72.

Wilderjans, E. et al., 2010. Impact of potassium bromate and potassium iodate in a pound cake system. *J. Agric. Food Chem*, 58(10), pp.6465–71.

Williams, H.T.P., 2013. Phage-induced diversification improves host evolvability. *BMC evolutionary biology*, 13, pp.1–17.

Wilson, E.B., 1928. *The cell in development and heredity.*, New York: Macmillan.

Wilson, W.A., Wang, Zhong & Roach, P.J., 2002. Systematic Identification of the Genes Affecting Glycogen Storage in the Yeast Saccharomyces cerevisiae: Implication of the Vacuole as a Determinant of Glycogen Level. *Mol Cell Proteomics*, 1(3), pp.232–242.

Wolfe, K.H. & Shields, D.C., 1997. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature*, 387(6634), pp.708–713.

Wout, P.K. et al., 2009. Saccharomyces cerevisiae Rbg1 protein and its binding partner Gir2 interact on Polyribosomes with Gcn1. *Eukaryotic cell*, 8(7), pp.1061–71.

Wu, C.-Y. et al., 2010. Control of transcription by cell size. *PLoS Biol*, 8(11), p.e1000523.

Wu, H. et al., 2006. Global gene expression analysis of yeast cells during sake brewing. *Appl Environ Microbiol.*, 72(11), pp.7353–8.

Xie, C. & Tammi, M.T., 2009. CNV-seq, a new method to detect copy number variation using high-throughput sequencing. *BMC Bioinformatics*, 10, p.80.

Yang, R., Wek, S. a & Wek, R.C., 2000. Glucose limitation induces GCN4 translation by activation of Gcn2 protein kinase. *Mol Cell Biol*, 20(8), pp.2706–17.

Yazawa, H., Iwahashi, H & Uemura, H., 2007. Disruption of URA7 and GAL6 improves the ethanol tolerance and fermentation capacity of Saccharomyces cerevisiae. *Yeast*, 24, pp.551–560.

Ye, K. et al., 2009. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics*, 25(21), pp.2865–71.

Yeh, J. & Haarer, B.K., 1996. Response to thermal stress. *FEBS Lett*, 398, pp.303–307.

Yeong, F.M., 2005. Severing all ties between mother and daughter: cell separation in budding yeast. *Mol Microbiol*, 55(5), pp.1325–31.

Yoshikawa, K. et al., 2009. Comprehensive phenotypic analysis for identification of genes affecting growth under ethanol stress in Saccharomyces cerevisiae. *FEMS Yeast Res*, 9(1), pp.32–44.

You, K.M., Rosen, C. & Knipple, D.C., 2003. Ethanol Tolerance in the Yeast Saccharomyces cerevisiae Is Dependent on Cellular Oleic Acid Content. *Appl Environ Microbiol.*, 69(3), pp.1499–1503.

Zancan, P. & Sola-Penna, M., 2005. Trehalose and glycerol stabilize and renature yeast inorganic pyrophosphatase inactivated by very high temperatures. *Arch Biochem Biophys.*, 444(1), pp.52–60.

Zeng, J., Smith, K.E. & Chong, L.-G., 1993. Effects of Alcohol-induced Lipid Interdigitation on Proton Permeability in L-a-Dipalmitoylphosphatidylcholine Vesicles. *Biophys J.*, 65, pp.1404–1414.

Zeyl, C, 2004. Experimental studies of ploidy evolution in yeast. *FEMS Microbiol Lett*, 233(2), pp.187–192.

Zeyl, Clifford, Vanderford, T. & Carter, M., 2003. An evolutionary advantage of haploidy in large yeast populations. *Science*, 299(5606), pp.555–8.

Zhang, L. et al., 2006. Reaction progress of chromophore biogenesis in green fluorescent protein. *J Am Chem Soc*, 128, pp.4766–4772.

Zheng, D.-Q. et al., 2013. Comparative functional genomics to reveal the molecular basis of phenotypic diversities and guide the genetic breeding of industrial yeast strains. *Appl Microbiol Biotechnol*, 97(5), pp.2067–76.

9. Appendices

9.1 Appendix I: Code for population genetic simulation.

This R script simulates two strains growing in a single population which repeatedly grows from a bottleneck size to a carrying capacity. Each bottleneck is achieved by random sampling without replacement. Rounds of cell division increase the population until it reaches the carrying capacity. The likelihood of each cell dividing in a round of division is given by a probability

```
setwd("…/PopulationGenetics/Rsimulations")

headerLab<-"Evolution6"

numPlates<-10

carryCapacity<-3E7

divideProbability<-0.9

bottleNeck<-seq(from=100,to=5000,by=50)

for(h in c(1:length(bottleNeck)))

        {

         print(bottleNeck[h])

         for (i in c(1:50))

                {

                 pop<-sample(c(rep(0,bottleNeck[h]/2),rep(1,bottleNeck[h]/2)))

                 ratio<-log((sum(pop))/(length(pop)-sum(pop)))/log(10)

#Log10 of the starting ratioCounts –

#the number of 1s/number of zeros -

                #NB If 1s go to fixation this will throw an error
```

```
		for (x in c(1:numPlates))
			{#Iterate for a number of plates
			 tmpPop<-sample(pop,size=bottleNeck[h], replace=FALSE)
			 #Grow from bottleNeck[h] to carryCapacity
			 while(length(tmpPop)< carryCapacity)
				{#Grow population to carrying capacity
				 divides<-runif(length(tmpPop))>divideProbability
				 tmpPop<-c(tmpPop,tmpPop[divides])
				}
			 #Calculate ratio and store in ratio vector
			 ratio<-round(c(ratio,(log((sum(tmpPop))/(length(tmpPop)-
sum(tmpPop)))/log(10))),3)
			 pop<-tmpPop
			 rm(tmpPop)
			 gc()
			}

		ratioRange<-max(ratio)-min(ratio)
		maxDeviation<-
ifelse((abs(min(ratio))>max(ratio)),min(ratio),max(ratio))
		mainLab<-paste(headerLab," Simulations.
Bottleneck:",bottleNeck[h]," carrying capacity:",carryCapacity," division
probability:",divideProbability," range:",ratioRange,sep="")

#####################################################################
		#Plot
```

```
                outputString<-
paste("PopSimulation_Neck_Div1_",bottleNeck[h],"_",i,".png",sep="")


png(filename=outputString,width=1500,height=1000,units="px",pointsize=26)

                plot(ratio,ylim=c(-1,1),xlab="Plate",ylab="Log10(Population
Ratio)",main=mainLab)

                abline(h=0)

                points(ratio,pch=16)

                lines(ratio)

                dev.off()

#########################################################################

                #Write Out

                ratioText<-paste(ratio,collapse=",")

                outputText<-
paste((numPlates+1),bottleNeck[h],carryCapacity,divideProbability,ratioRange,maxD
eviation,ratioText,sep="\t")

                write(outputText,file="Rsimulations6.txt",append=TRUE,sep="")

               }

        }
```

9.2 Appendix II: Biolog GenIII well contents

Negative Control

Dextrin

D-Maltose

D-Trehalose

D-Cellobiose

Gentiobiose

Sucrose

D-Turanose

Stachyose

Positive Control

pH 6

pH 5

D-Raffinose

α-D-Lactose

D-Melibiose

β-Methyl-DGlucoside

D-Salicin

N-Acetyl-DGlucosamine

N-Acetyl-β-DMannosamine

N-Acetyl-DGalactosamine

N-Acetyl Neuraminic Acid

1% NaCl

4% NaCl

8% NaCl

α-D-Glucose

D-Mannose

D-Fructose

D-Galactose

3-Methyl Glucose

D-Fucose

L-Fucose

L-Rhamnose

Inosine

1% Sodium

Fusidic Acid

D-Serine

D-Sorbitol

D-Mannitol