Master's Thesis

Mutation rate of ribosomal DNA copy number in *Neurospora crassa*

Suvi Lehtonen

University of Jyväskylä

Department of Biological and Environmental Science

Cell and molecular biology

18.05.2023

JYVÄSKYLÄN YLIOPISTO, Matemaattis-luonnontieteellinen tiedekunta
Bio- ja ympäristötieteiden laitos
Solu- ja molekyylibiologia

Suvi Lehtonen: **ribosomaalisen DNA:n toistojaksojen mutaatiovauhti *Neurospora crassa* -sienellä**

Pro gradu -tutkielma: 30 s., 2 liitettä (3 s.)
Työn ohjaajat: Dosentti Ilkka Kronholm ja tutkijatohtori Mariana Villalba de la Peña
Tarkastajat: Varpu Marjomäki ja Nerea Abrego
05/2023

Hakusanat: efektiivinen populaatiokoko, kopiolukumäärä, mutaatioakkumulaatio, mutaatiovauhti, rDNA

Ribosomaalinen DNA (rDNA) koodaa rRNA:ta, joka on ribosomien tärkeä rakenneosa. Eukaryooteilla rDNA esiintyy lukuisien identtisten toistojaksojen sarjoina. rDNA:n toistojaksojen tiedetään olevan yksi genomin epävakaimmista alueista, ja niissä esiintyy paljon kopiolukumäärän vaihtelua. Vaikka vaihtelun mekanismeista on tehty tutkimusta, syitä runsaalle kopiomäärälle ja sen muutosalttiudelle ei täysin tunneta, On esitetty, että se saattaa mahdollistaa lokuksen pienenemisen tai laajenemisen myötä paremman adaptaation muuttuviin olosuhteisiin. Tässä työssä tutkittiin *Neurospora crassa* -sienen rDNA:n kopiolukumäärän spontaania mutaatiovauhtia mutaatioakkumulaatiokokeen avulla ja estimoitiin valinnan vaikutusta mutaatiovauhtiin. Tutkimuksessa käytettiin mutaatioakkumulaatiokantoja ja luonnonkantoja. rDNA:n kopiolukumäärä analysoitiin droplet digital PCR-menetelmällä. rDNA-mutaatioita havaittiin useissa linjoissa, ja mutaatiovauhdiksi saatiin 18S rDNA:lle 1.10 [0.64, 1.72] x 10^{-5} ja 26S rDNA:lle 1.30 [0.09, 1.95] x 10^{-5} mutaatiota / mitoosi, mikä on melko korkea mutaatiovauhti. Deleetioita tapahtui huomattavasti enemmän kuin insertioita, mikä viittaa siihen, että mutaatiopaine pienentää rDNA:n kopiolukumäärää silloin kun muut mekanismit eivät vaikuta siihen. Koska mutaatioita tapahtui melko vähän, mutaatiomallia ei pystytty muodostamaan.

UNIVERSITY OF JYVÄSKYLÄ, Faculty of Mathematics and Science
Department of Biological and Environmental Science
Cell and molecular biology

Suvi Lehtonen: **Ribosomal DNA copy number mutation rate in *Neurospora crassa***

MSc thesis: 30 p., 2 appendices (3 p.)
Supervisors: Docent Ilkka Kronholm and post-doctoral researcher Mariana Villalba de la Peña
Reviewers: Varpu Marjomäki and Nerea Abrego
May 2023

Keywords: copy number variation, effective population size, mutation accumulation, mutation rate rDNA

Ribosomal DNA (rDNA) encodes ribosomal RNA, which is an essential structure of ribosomes. In eukaryotes rDNA loci contain multiple identical tandem repeat sequences and are susceptible to copy number variation (CNV). Although the mechanisms of copy number increase and decrease have been studied, the reasons underlying CNV are not yet completely understood. It has been suggested that environmental factors may change the copy number, and therefore it may be one of the mechanisms that helps the organism to adapt to changes in the environment. In this study, spontaneous mutation rate of 18S and 26S rDNA copy number of *Neurospora crassa* was studied using droplet digital PCR. *N. crassa* lines from a mutation accumulation experiment and from natural populations were used. The mutation accumulation lines had been cultured for 40 generations in a way that minimized the effect of selection. Mutation rate was estimated by comparing rDNA copy numbers between ancestors and different generations. This information was then used to estimate how selection affects mutation rate. The mutation rate in 18S rDNA was 1.10 [0.64, 1.72] x 10^{-5} and in 26S rDNA 1.30 [0.09, 1.95] x 10^{-5} mutations per mitosis, which is a relatively high mutation rate. There were significantly more deletions than insertions, suggesting that mutation pressure reduces rDNA copy number when other mechanisms are not involved. However, there were not enough mutations to form a mutation model.

TABLE OF CONTENTS

ABBREVIATIONS

CNV	Copy number variation
ddPCR	Droplet digital polymerase chain reaction
MA	mutation accumulation
NOR	Nucleolar organization region
rDNA	Ribosomal DNA
rRNA	Ribosomal RNA

1 INTRODUCTION

1.1 Ribosomal DNA (rDNA)

Ribosomal RNA molecules are essential structural and catalytic elements of ribosomes, the machineries of protein synthesis. Eukaryotic ribosomes contain four different rRNA molecules which are highly conserved throughout species and different cell types. These molecules are coded in ribosomal DNA (rDNA) loci, which are typically located in the nucleolar organizer region in one or several chromosomes depending on the species, although the rDNA loci can also be located in other regions or even outside the genome as an extra chromosomal DNA. rDNA is crucial for biosynthesis of ribosomes, and it is also suggested to have other key functions such as organizing the nucleus, and it may also have some other, non-coding functions (Ganley and Kobayashi 2011; Gibbons *et al.* 2014; Bughio and Maggert 2019).

Despite some species-specific differences in the organization and location of rDNA sequences, the structure of rDNA loci is similar in all eukaryotic cells. There are two different types of rDNA loci. One of them is 45S, and it comprises tandemly arranged copies of 5.8S, 18S and 25—28S rRNA genes, which are transcribed as a single precursor and later processed into mature rRNA segments. The other locus contains tandem repeats of 5S rRNA genes (Gibbons *et al.* 2014; Bughio and Maggert 2019). In addition to coding sequences, the 45S locus has two internal transcribed sequences between the coded 18S, 5.8S and 28S sequences, and both 45S and 5S loci have large non-transcribed intergenic spacers between the repeat units. Unlike most other rDNA regions, these sequences are hypervariable and can be used as a phylogenetic marker **(Li *et al.*** 2016; Guo *et al.* 2019). The structure of 45S and 5S loci is visualized in figure 1.

To ensure that there is enough rRNA, there often exists several gene copies for genes that encode rRNA products. RNA molecules do not go through translation and their production cannot be amplified in this step, and therefore multiple gene copies are needed to produce the necessary amount of RNAs for the demands of the cells (Ide *et al.* 2010). Demonstrating the pivotal role of rRNA, there are between ~40 to ~20 000 tandemly arranged rDNA gene copies in a locus in eukaryotic cells, each species having its typical copy number. The copy number correlates positively with genome size (Prokopowich *et al.* 2003).

rRNA accounts for over 60% of the total RNA of the cells, and transcription of rDNA genes has been indicated to be nearly 50% of all transcription in eukaryotes (Moss and Stefanovsky 2002; Ide *et al.* 2010; Aldrich and Maggert 2015). However, not all rDNA genes are actively transcribed. It has been estimated that approximately half of the rDNA copies are inactivated with histone modifications, and when more RNA is needed, it is more likely that the cell increases activity of one gene copy than activates more copies (Aldrich and Maggert 2015; Bughio and Maggert 2019).

1.2 rDNA Copy number variation

rDNA loci are highly dynamic, and it has been suggested that these loci have more copy number variation (CNV) and polymorphism containing segments than any other parts of the genome. Changes in copy number can evolve rapidly especially in multicellular organisms with multiple rDNA loci. For example, even when an organism has only a few loci, the copy number variation can be as high as 6-fold (Lyckegaard and Clark 1989; Lofgren *et al.* 2019; Bughio and Maggert 2019). Despite differences in species-specific rDNA copy number, there seems to be a somewhat similar amount of variation in the number of rDNA copies between species.

Probable causes underlying this kind of instability in the rDNA is high transcription activity taking place even during the S-phase of cell division, and the presence of multiple identical tandem repeats, which are difficult to replicate (Gibbons *et al.* 2014; Salim *et al.* 2017; Lu *et al.* 2018). Because transcription and replication occur

simultaneously, it is possible that the transcription and replication machineries collide and cause breakages to the double stranded DNA (Takeuchi *et al.* 2003; Helmrich *et al.* 2013). When these breakages are fixed using homologous recombination repair (HR repair), the gene copies might be attached in a wrong way causing either insertions or deletions. It has been estimated that in every mitosis, approximately one rDNA gene copy is either lost or gained, and deletions and insertions seem to be equally likely (Ganley and Kobayashi 2011; Lu *et al.* 2018). In addition, some other mechanisms using HR repair, such as meiotic magnification, have been shown to impact the copy number (Salim *et al.* 2017; Bughio and Maggert 2019).

The reasons for high copy number variation in rDNA are not completely understood. One possibility is that because many of the gene copies are silenced and thus changes in the copy number do not show direct effects on the phenotype, therefore maintaining a certain copy number would not be vital for the organism (Lyckegaard and Clark 1989; Moss and Stefanovsky 2002). However, several studies have reported opposite results. It is suggested that rDNA copy number is important for the organism, and that too low or high rDNA copy number might be harmful or even deleterious. To support this, the rDNA copy number variation has been linked for example to some diseases and aging, and to nutritional and metabolic changes (Aldrich and Maggert 2015; Salim *et al.* 2017; Lu *et al.* 2018), In addition, there seems to be a connection between insufficient rDNA transcription and some diseases such as cancer in humans (Gibbons *et al.* 2014; Aldrich and Maggert 2015; Bughio and Maggert 2019).

rDNA copy number variation has been suggested to be one of the mechanisms for an organism to adapt to its environment: high number of gene copies and the hypervariability of the copy number in the loci might make fast adaptation to environmental changes possible. Spontaneous duplications are quite common in general, and they are known to occur rapidly under selection pressure. In multicellular organisms, duplications in parts of the genome are crucial for

evolution ((Okazaki *et al.* 2021), and this might also be the case with unicellular organisms. What comes to deletions, reducing copy number might for example release resources for the replication of rest of the genome and thus speed up cell division (Salim *et al.* 2017).

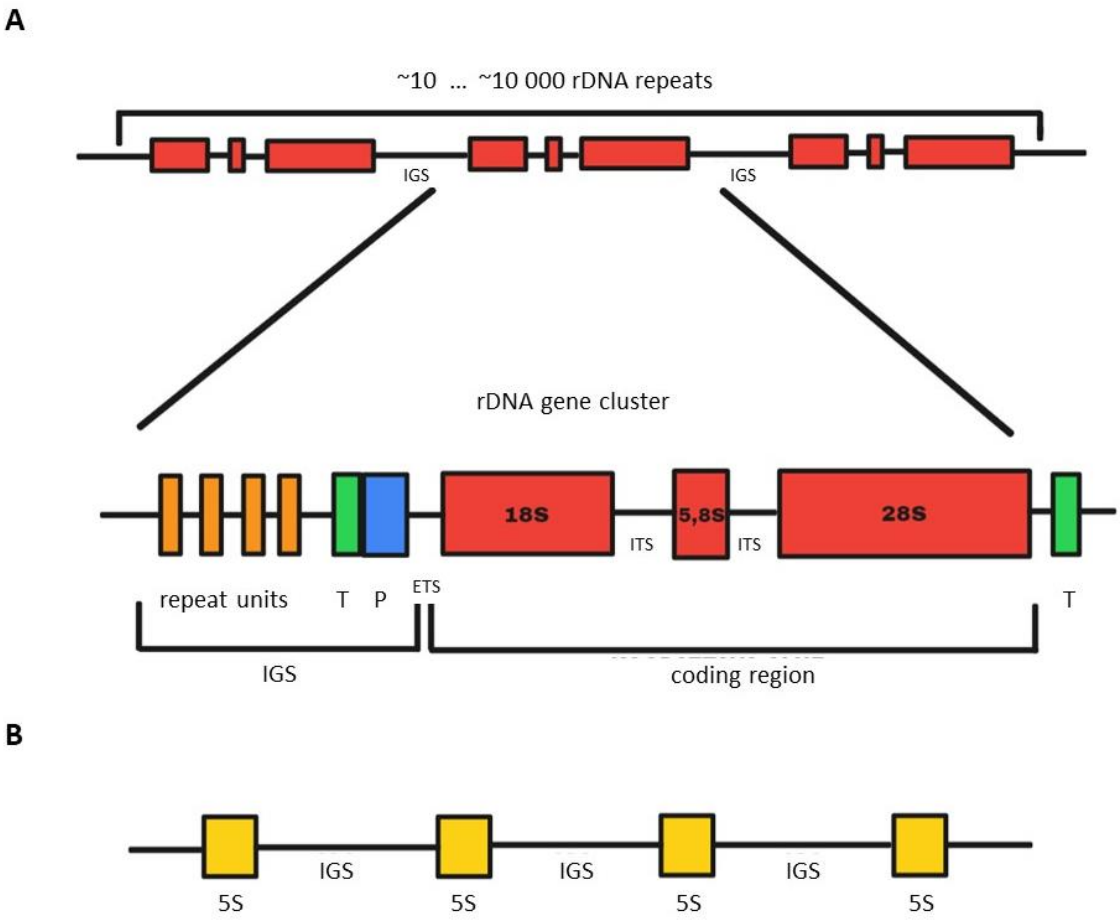


Figure 1. A) Structure of 45S rDNA gene cluster coding 18S, 5,8S, and 28S rDNAs. B) Structure of 5S rDNA. IGS = intergenic spacer, ETS = external transcribed spacer, ITS = internal transcribed spacer, T = transcription end point, P = promoter.

1.3 Study species: *Neurospora crassa*

Neurospora crassa is a filamentous fungus naturally growing on burned wood after a forest fire. It has been used as an experimental organism for nearly a hundred years, and it has had a central role in the development of modern genetics and molecular biology (Galagan *et al.* 2003). The organism is relatively easy to grow, and

because manipulating it genetically is somewhat simple, it is considered a model genus for genetic evolution.

N. crassa has a genome size slightly over 40 megabases including ~10—11 000 protein coding genes, which is larger than for example that of other often used model organisms *S. pombe* and *Saccharomyces cerevisiae*. The genes, spread in 7 chromosomes, make up at least 44% of the whole genome. It has been estimated that 10% of the *N.* crassa genome contains repeat sequences (Galagan *et al.* 2003). This includes the 8,6 kbp 45S rDNA locus in the nucleolar organizing region of chromosome V (Cox and Peden 1979), where genes for 26S-, 5,8S- ja 18S -rRNA are located. In *N. crassa*, the locus carries approximately ~170—200 rDNA gene copies per haploid genome. For 5S rRNA coding, the species has 20—30 loci spread in an un-clustered fashion in the seven chromosomes and outside the nucleolar organizing region (Free *et al.* 1979; Radford and Parish 1997; Galagan *et al.* 2003; Rodriguez *et al.* 2022).

The mutation rate of the *N. crassa* genome has been studied quite extensively. A recent study by Villalba de la Peña et al. (2022) showed that the mutation rate in the whole genome of the species during asexual reproduction is 0.03 [0.03, 0.04] mutations / genome / mitosis, which is somewhat typical for microbes. However, there seems to be a lot of variation in the mutation rate, repeat sequences having the highest rate. For example, the rate for single nucleotide mutations was estimated to be 6.7 [6.32, 7.11] × 10^{-10} / bp / mitosis, whereas the mutation rate of repeat units was 0.0014 [0.0010, 0.0017] mutations / genome / mitosis for deletions and 0.0030 [0.0025, 0.0036] mutations / genome / mitosis for insertions. Furthermore, it seems that areas containing heterochromatin have higher mutation rate than those containing euchromatin. (Villalba de la Peña *et al.* 2022). What is the mutation rate of *N. crassa* rDNA, containing repeat units and transcribed and silenced genes, is yet to be investigated.

1.4 Study questions and hypotheses

The aim of this study is to examine mutation rate of rDNA copy number in the fungus *N. crassa,* and to find out what kind of mutation model it follows. Mutation models are mathematical models which aim to characterize patterns of the mutational process. To this end, 18S and 26S rDNAs from a mutation accumulation (MA) experiment, and a sample of strains from natural populations were analyzed. The study focuses only on copy number (repeat unit) mutations, meaning that for example point mutations are excluded from the study.

Majority of the knowledge about genetic variation caused by spontaneous mutations has been gained from MA experiments. They offer ideal and somewhat unbiased means to study accumulation of spontaneous mutations, and mutation rate (Katju and Bergthorsson 2019). In a typical MA experiment, several lines are cultured for generations from one ancestor line. In these lines the effect of natural selection is minimized, so that most of the mutations appear or disappear randomly. These mutations are allowed to accumulate to the whole genome for several generations, and after this comparison between the ancestors and the different generations is carried out (Halligan and Keightley 2009).

Further, natural variation occurring in the rDNA repeat units of *N. crassa* is examined. This information is used for estimating how natural selection impacts the rDNA copy number: when the mutation rate of MA lines is known, it is possible to find out whether genetic mutations observed in natural populations can be explained solely by the mutation rate, or does natural selection also have an effect on the accumulation of mutations. In previous studies, rDNA copy number has been indicated to remain somewhat stable in a population level, and some species have been shown to have various mechanisms for the maintenance of copy number (Ide *et al.* 2010; Ganley and Kobayashi 2011; Gibbons *et al.* 2014; Salim *et al.* 2017). It can be hypothesized that for each species there is an optimal rDNA copy number, and thus it is suggested to be subject to balancing selection.

The study will be carried out using droplet digital PCR (ddPCR). It is a high-throughput method capable of absolute quantification of DNA molecules. In ddPCR the reagents containing fluorescence probes are divided randomly to ~20 000 water-oil droplets, and DNA concentration is calculated from the fraction of positive and negative droplets based on their fluorescence amplitude. (Hindson *et al.* 2011; Salim *et al.* 2017). The strength of this method is that it can reliably detect even minor changes in the DNA concentration. Because of its accuracy and sensitivity, ddPCR is considered a suitable method for studying CNV.

The study is basic research, and its primary focus is in achieving a better understanding of rDNA mutation rate in *N. crassa*. The aim is to explain what kind of effects spontaneous mutations, genetic variation, and natural selection have on the rDNA copy number in natural populations. rDNA copy number variation and the reasons underlying it are poorly understood, and there are only a few MA experiments carried out in fungi (Halligan and Keightley 2009; Katju and Bergthorsson 2019).

2 MATERIALS AND METHODS

2.1 Overview

In this study, the aim was to examine mutation rate of ribosomal DNA copy number in *N. crassa* fungus from mutation accumulation experiment lines and natural population strains. To this end, *N. crassa* MA lines had been cultured for 40 generations. To be able to compare the copy number of the lines rDNA was extracted and analyzed from the samples using droplet digital PCR. In addition, the effect of selection on copy number was estimated from the data by comparing the MA results to those from natural population strains.

2.2 Materials

Total of 39 *N. crassa* MA lines were used in the study. The lines had been cultured in a recent experiment (Villalba de la Peña *et al.* 2022), and they consisted of mating types *mat A* (n = 20) and *mat a* (n = 19). Mutations had been allowed to accumulate for 40 generations in the MA-lines. For this experiment, ancestors and generation 40 samples from each line were analyzed. In addition, samples from generations 5 and 20 from several lines were used in the analysis.

Mutation accumulation lines had been created by transferring a colony from a single spore descent from a plate to a test tube and allowing the cells to propagate asexually for five days, before making another similar transfer. The transfers had been repeated 40 times, each representing one generation. The outline of the MA experiment is visualized in figure 2. It was estimated that during 40 generations the

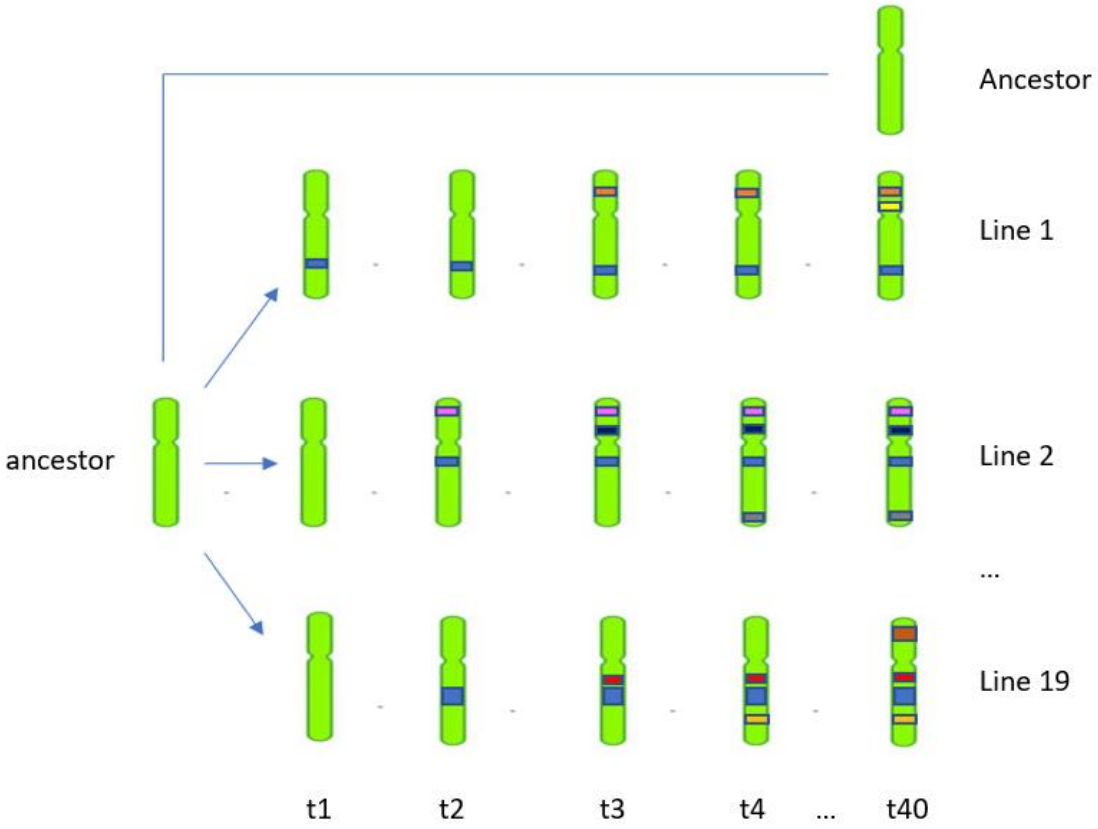


Figure 2. In a typical MA experiment, multiple lines are split from a single spore ancestor and cultured for several generations. Differences between the genomes of ancestor and generations can then be compared.

cells went through 1015 [1003, 1026] (mean [95% confidence interval]) mitosis on average (Villalba de la Peña et al. 2022). The natural population samples (n = 92) were ordered from Fungal Genetics Stock Center and had been originally collected from the Caribbean and from Louisiana. All *N. crassa* natural population strains used in this experiment are listed in appendix 1.

2.3 Methods

2.2.1 Culturing of *Neurospora crassa* samples and DNA extraction

DNA was extracted from the MA-samples and natural population samples using Phenol Chloroform extraction. Cultures with conidia were first grown in Vogel's Medium N (Metzenberg 2003) with 1.5% agar and 1.5% sucrose for two days at 25 ℃. The mycelium was harvested, and the samples were freeze-dried in cryotubes overnight using a lyophilizer. The sample tissues were frozen in liquid nitrogen and then ground twice using a Qiagen shaker (frequency 25 / sec., time 20 sec.). 500 µl of extraction buffer (10 mM Tris pH 8, 0.1 M EDTA, 150 mM NaCl, 2% SDS) was added to the tubes, which were then vortexed and shaken to dissolve the tissue. The samples were transferred to new tubes, 750 µl of Phenol Chloroform (25 Phenol: 24 Chloroform: 1 isoamyl alcohol) was added to the samples and they were centrifuged five minutes at 20 000 rounds per minute.

To degrade RNAs, 2 µl of RNAse A (10mg / ml) and 1 µl of RNAse I (50 U / µl) was added to the samples, and they were allowed to incubate for 1 h at 37 ˚C. After this, 750 µl of chloroform was added, and the samples were again centrifuged for five minutes at maximum speed. The aqueous phases of the samples were put into new tubes, and 1 ml of 100% ethanol was added to each sample. The samples were mixed by inverting the tubes and then allowed to precipitate for 1 h at -20 ˚C. To pellet DNA, the samples were centrifuged for 20 minutes at maximum speed. Ethanol was aspirated from the tubes and the samples were washed with 70% ethanol, which was then aspirated. The samples were air dried for 5—10 minutes and eluted to 77.5

µl of TE-buffer, and then incubated at 37 ˚C until the pellets were dissolved. Then, 12.5 µl of 4M NaCl was added, and the samples were vortexed.

PEG precipitation was used to remove small DNA fragments from the samples, because with this method only large DNA samples would precipitate. Therefore, 12.5 µl of 50% PEG (polyethylene glycol P3350) was added to the tubes, the tubes were mixed, and the samples were allowed to precipitate overnight in the fridge (4 ˚C). Next day, the samples were centrifuged at maximum speed for 10 minutes to pellet DNA, and the liquid was aspirated from the tubes. To remove salt, the sample pellets were washed twice with 70% ethanol, and then they were eluted to 55 µl of TE-buffer. To check the quality of DNA after the pellets had dissolved, 2 µl samples were run on a 0.8% agarose gel, and DNA concentration was measured on Qubit Fluorometer (Qubit ™ dsDNA BR assay, #Q32850, Invitrogen).

2.2.2 Droplet digital PCR

To estimate the rDNA copy number from *N. crassa* MA-lines and natural population strains, ddPCR-samples were prepared. The copy number was analysed for both 18S and 26S rDNA segments, and therefore each line was sampled twice. The ddPCR reaction had been optimized and the primers and probes had been tested earlier for 18S rDNA. The reactions were made accordingly for 26S rDNA. To prepare the samples, 100 ng of DNA / sample was restricted using NcoI-HF enzyme (R3193L VIAL, Biolabs, New England) in 37 ˚C for 45 minutes. The DNA concentration was then diluted to 0.025—0.0025 ng/ml, and 20 µl ddPCR reactions were prepared following the manufacturers (Bio-Rad) instructions. Primers and probes for either 18S or 26S rDNA and for a reference gene (Tba1) were added to the reactions. Primers and probes are presented in appendix 2. After droplet generation, PCR reaction was initiated by heating the samples in 95 ˚C for 10 minutes, after which the temperature was decreased to 94 ˚C for 30 seconds. The annealing was made in 62 ˚C for 1 minute, and the extension in 98 ˚C for ten minutes,

and the cycle was repeated 40 times. The samples were then read using Bio-Rad QX200 Droplet reader and QuantaSoft Software.

2.2.3 Copy number analysis

After ddPCR, rDNA copy numbers were estimated from the data by comparing DNA concentrations of target and reference samples using the Poisson distribution. First, the concentrations of reference and target molecules were calculated based on the fractions of positive and negative droplets (Hindson *et al.* 2011). Concentration (copies /µl) was derived from a formula

$$C = -\ln(N_{neg} / N) / V_{droplet}$$

where N_{neg} is the number of negative droplets in the sample, N is the total number of accepted droplets, and $V_{droplet}$ is the computational volume of one droplet (= 1 nl). After this, the copy number for each sample was defined from the concentrations of reference and target molecules using a formula

$$CN = A / B * N_b$$

where A is concentration of target molecule, B is concentration of reference molecule, and N_b is the number of reference loci Tba1 (= 1 in haploid *N. crassa*). The copy numbers and variation (95% confidence interval) were calculated using common statistical software (Excel, R).

2.2.4 Quantification of mutations

The number of mutations, defined as a difference in the copy number between generations, was estimated from the copy number data of MA lines. This was done for each line separately, by comparing results from ancestors and generation 40. For several lines, the copy number had been calculated also for generations 5 and 20, and these were included in the analysis as well. The criterion for a mutation was that the range (95% confidence interval) of copy numbers of two generations from the same line should not be overlapping, or the overlap should be < 10 copies.

In *mat a* lines, the copy number of the ancestor was first compared to that of generation 40. For those lines that had copy numbers calculated from generations 5 and 20, the next step was to carry out a comparison between all consecutive generations. In each line, the number of mutations was summed. For example, if a line had one deletion in generation 5 and another one in generation 20, this was counted as two mutations. Insertion-deletion pairs were counted as two mutations as well.

The *mat A* line ancestor was excluded from the mutation rate analysis, because an abnormally low copy number was detected in all the ancestor samples. Therefore, for *mat A* lines, the mutations were calculated by using the generation 5 as a starting point and comparing it to generation 40. After excluding the ancestors, some of the samples had results from generations 5, 20 and 40 whereas the rest had results only from generation 40. Thus, it was decided that the *mat A* samples lacking generation 5 results should be compared to the median of generation 5 copy numbers of all other *mat A* lines (one outlier (line 13A) was excluded when calculating the median). Otherwise, the procedure was similar to that of *mat a* samples.

2.2.5 Estimation of mutation rate and mutation model

From the number of mutations, a mutation rate for rDNA copy number was estimated using Poisson regression (R studio). The mutation rate was calculated together for the two mating types but separately for 18S and 26S genes. Poisson regression returned a natural logarithmic result of the number of mutations which had occurred in the MA-lines. To obtain the number of mutations, the result was inversed with an exponential function by using a formula

$$x=\ln(y) \leftrightarrow y=e^{x}$$

where x is the intercept from the Poisson regression, y is the number of mutations, and e is the mathematical constant e. Last, to calculate mutation rate, the number of mutations was divided by the number of mitoses (1015) and lines (18S = 38, 26S

= 39). The results are represented as means [95% confidence interval]. The effect of selection on mutation rate was then hypothesized based on these results.

3 RESULTS

3.1 rDNA copy number

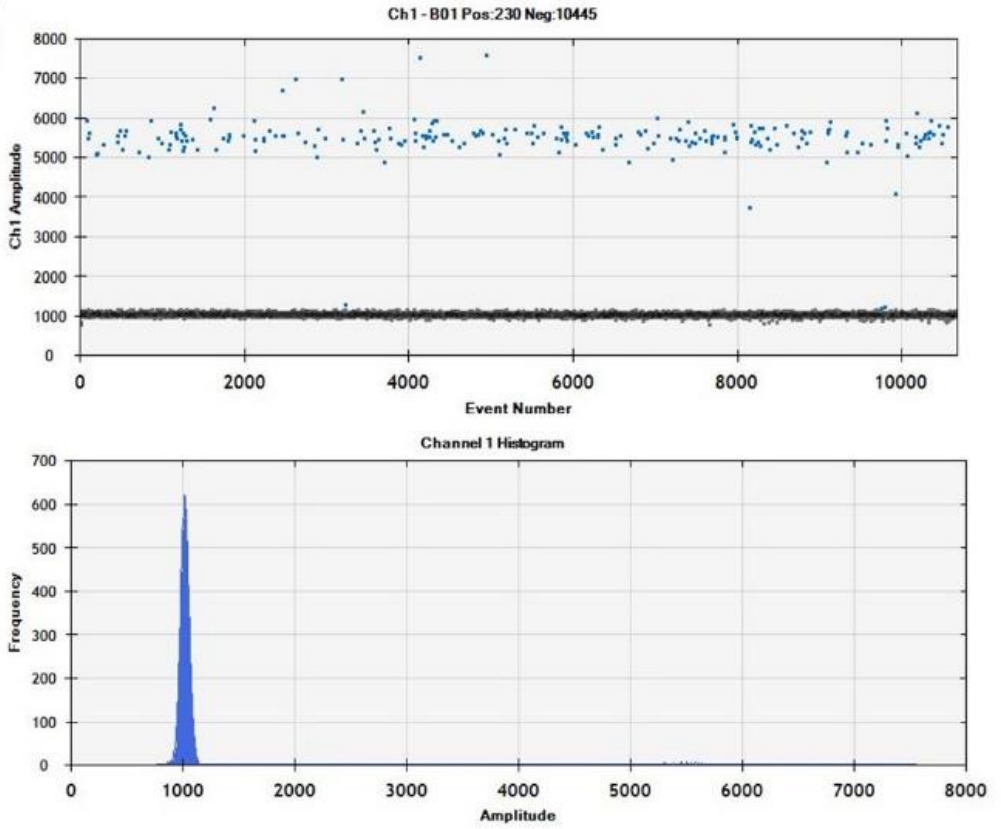
3.1.1 18S rDNA

For the 18S rDNA, the generation 40 ddPCR samples were successfully run and analyzed from 38 lines (*mat a*, n = 19, *mat A*, n = 19). In addition, generation 5 and 20 samples were run and analyzed from 10 lines (four lines had a missing value in one of the generations. See complete list of successfully analyzed samples from table 1.

Table 1. *N. crassa MA samples and generations used in the analysis.*

		n		
Line	Ancestors	generation 5	Generation 20	Generation 40
18S *mat a*	1	9	10	19
18S *mat A*	(5)	9	9	19
total		18	19	38
26S *mat a*	1	8	9	19
26S *mat A*	(5)	9	9	20
total		17	18	29

A


Ch1 - B01 Pos:230 Neg:10445
Ch1 Amplitude
Event Number
Channel 1 Histogram
Frequency
Amplitude

B

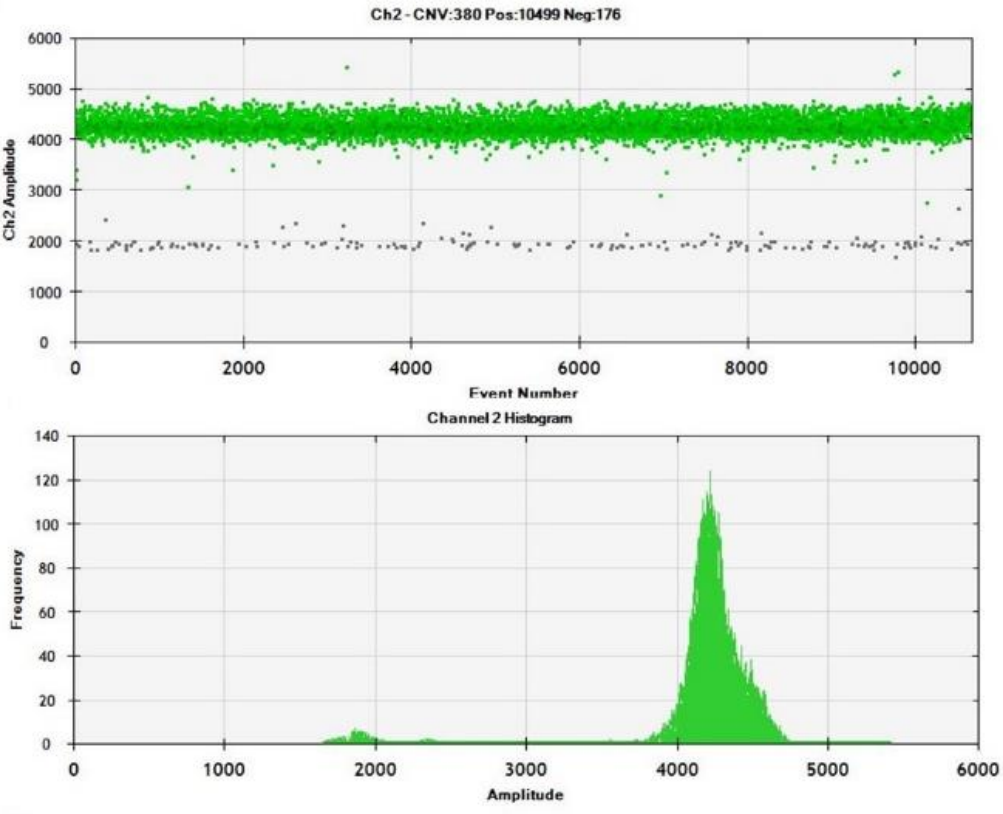

Ch2 - CNV:380 Pos:10499 Neg:176
Ch2 Amplitude
Event Number
Channel 2 Histogram
Frequency
Amplitude

Figure 3. ddPCR results from a representative MA sample. A) Above, the positive (blue) and negative (black) droplets of the reference (Tba1) gene, having only one copy in the genome. Below, the fluorescence amplitude and frequency of negative (amplitude ~1000–2000) and positive (amplitude ~4000–5000) droplets in Tba1 gene. B) Above, in 18S rDNA, most of the droplets are positive (green) due to high copy number. Negative droplets are gray in the picture. Below: the fluorescence amplitude and frequency of negative and positive droplets in 18S.

Two 18S *mat a* ancestors were also quantified in ddPCR, and because the results were expectedly almost similar (119[132,106] and 125[139,111]), the one with a lower copy number was selected as a reference point. From *mat A* line, 5 ancestor samples were quantified, but because they all had an abnormally low copy number (mean = 48), the *mat A* ancestors were excluded from further analyses. Thus, in 18S *mat A* lines, the generation 5 samples were used as a reference point to which all other *mat A* generations were compared.

The average copy number of 18S rDNA in 5th generation was 119, in 20th generation 109, and in 40th generation 104 copies. Although repeated measures ANOVA revealed a statistically significant difference ($F(1.566, 23{,}486) = 4.878$, $p = 0.023$) in copy number between generations in those lines (n = 16) that had results from all three generations, further pairwise comparison (Bonferroni) showed no differences between them. 18S gene copy numbers are presented in figure 4.

3.1.2 26S rDNA

For 26S rDNA, generation 40 lines (n = 39), as well as generation 5 and 20 *mat a* (n = 10) and *mat A* (n = 10) lines were analyzed (complete list of samples, see table 1). The *mat a* ancestor that was selected for the analysis had a copy number 123 [134,112]. The mean copy number of 26S rDNA in 5th generation was 117, in 20th generation the mean was 121, and in 40th generation 104 copies. As was the case with 18S samples, the *mat A* ancestors were excluded from the analyses, because the copy number was estimated to be too low (mean = 58). When comparing generations 5, 20, and 40 (repeated measures ANOVA), there was a statistically significant difference ($F(2, 30) = 9.136$, $p < 0.001$) in the copy numbers between the

generations in those lines that had results from all three generations (n = 11). Further, pairwise comparison (Bonferroni) revealed a statistically significant difference ($p < 0.001$) between generations 20 and 40. The 26S copy number results are presented in figure 5.

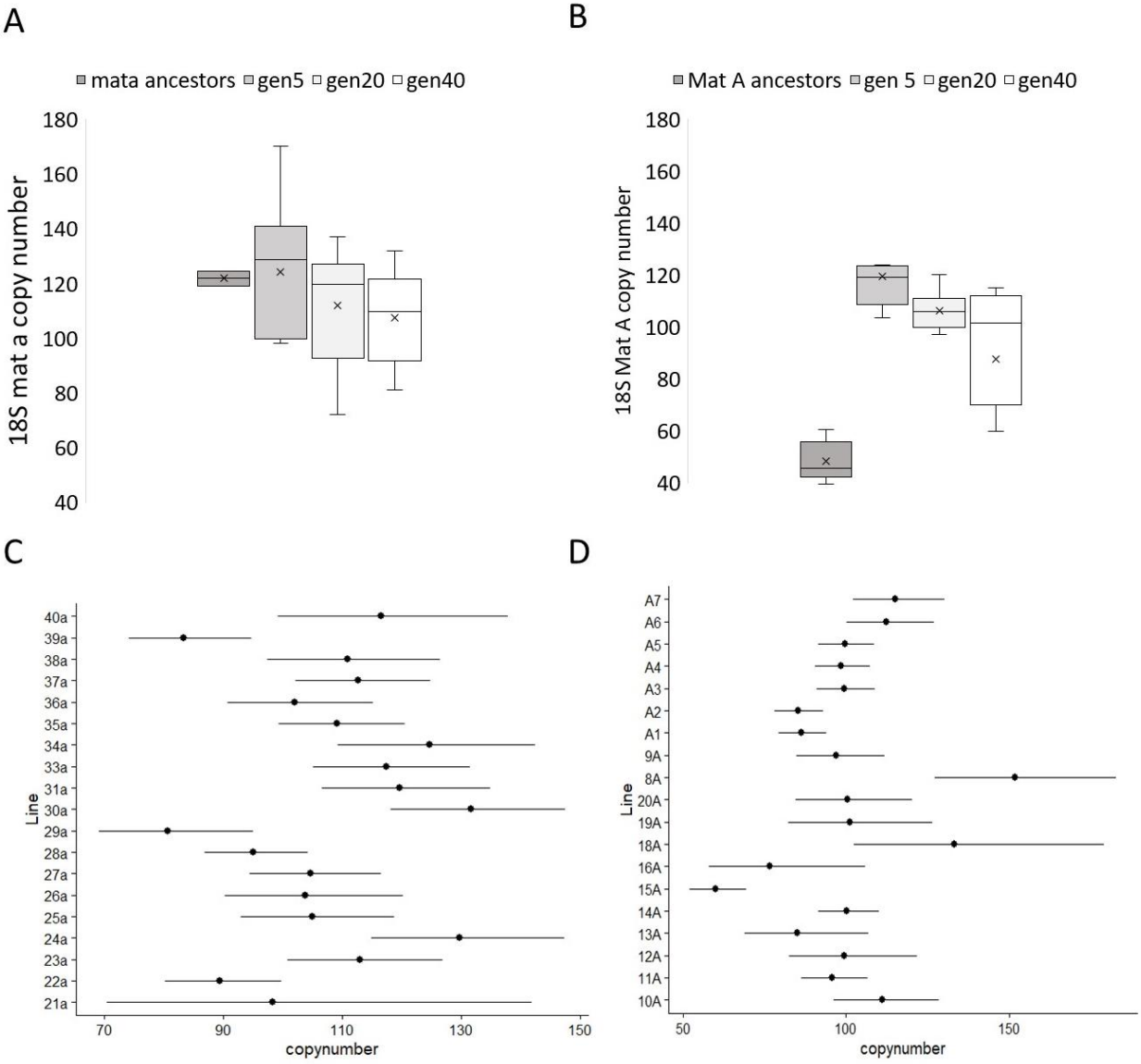


Figure 4. 18S rDNA copy numbers A) of *mat a* samples in ancestors (n = 2) and generations 5, 20, and 40. B) Similar results from *mat A* samples (ancestors, n = 5). The horizontal lines represent quartiles and the median, and x is the mean copy number. C) Copy numbers (95% confidence interval) of each *mat a* line in generation 40 D) Similar results from *mat A* lines.

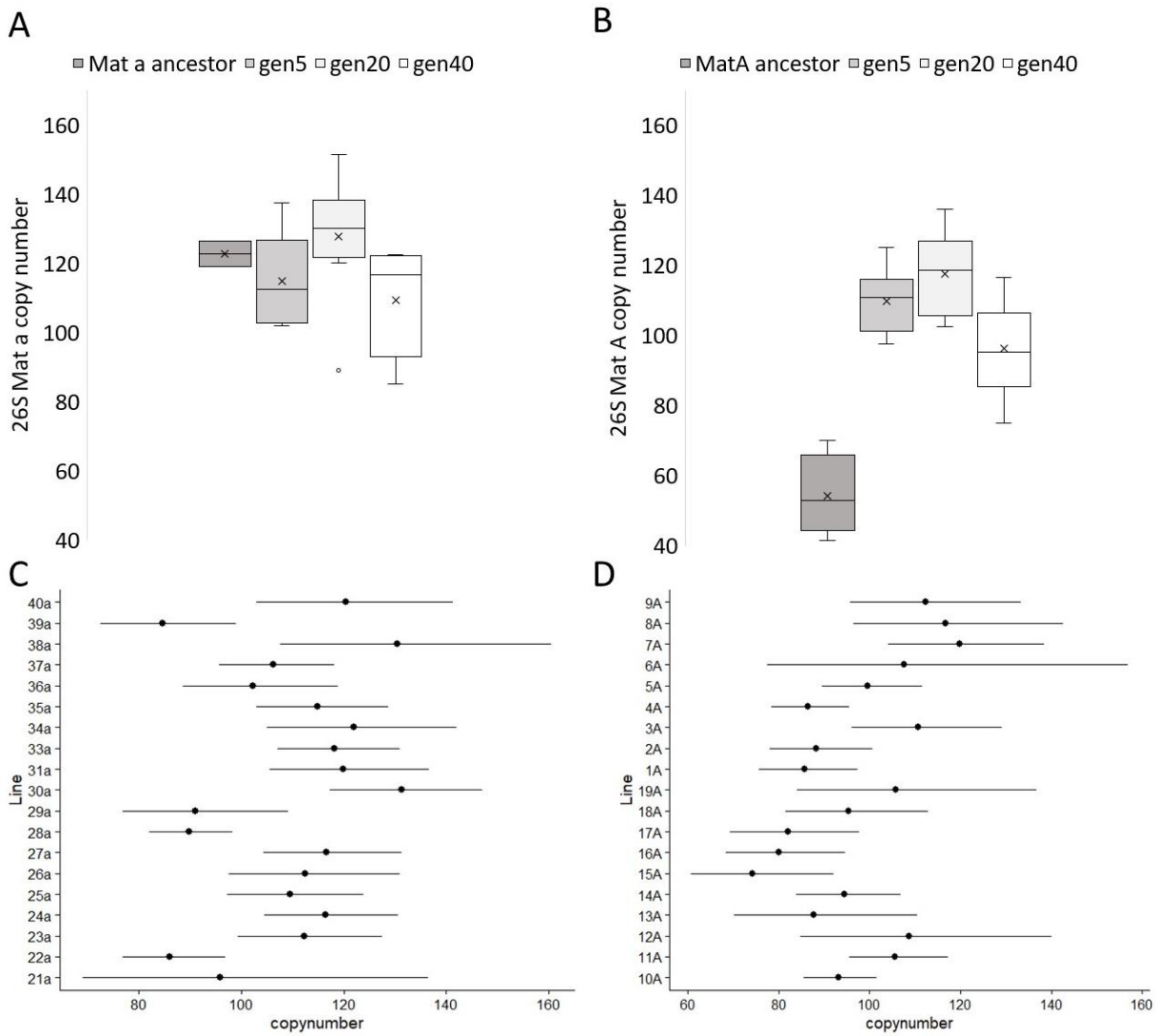


Figure 5. 26S rDNA copy numbers A) of mating type a samples from ancestors (n = 2) and generations 5, 20, and 40. B) Similar results from mating type A (ancestors, n = 5). The horizontal lines represent quartiles and the median, and x is the mean copy number. C) Copy numbers (95% confidence interval) of each *mat a* line in generation 40 D) Similar results from *mat A* lines.

3.1.4 Natural populations

DNA was extracted from 92 natural population samples, and the samples were run in ddPCR. Copy number quantification was carried out successfully to 18S (n = 92) and 26S (n = 89) samples. Because there was no information available of the mating types, those were ignored in the analyses. After quantifying the copy numbers, one

sample (strain 10894) was excluded from further analyses, because it had a relatively high copy number, and the 18S and 26S results were unexpectedly divergent (873 copies in 18S and 2765 copies in 26S). These results were likely caused by contamination or other technical issue. After this, the copy number for 18S samples was 128 ± 27 (mean ± SD), and for 26S rDNA it was 129 ± 34.

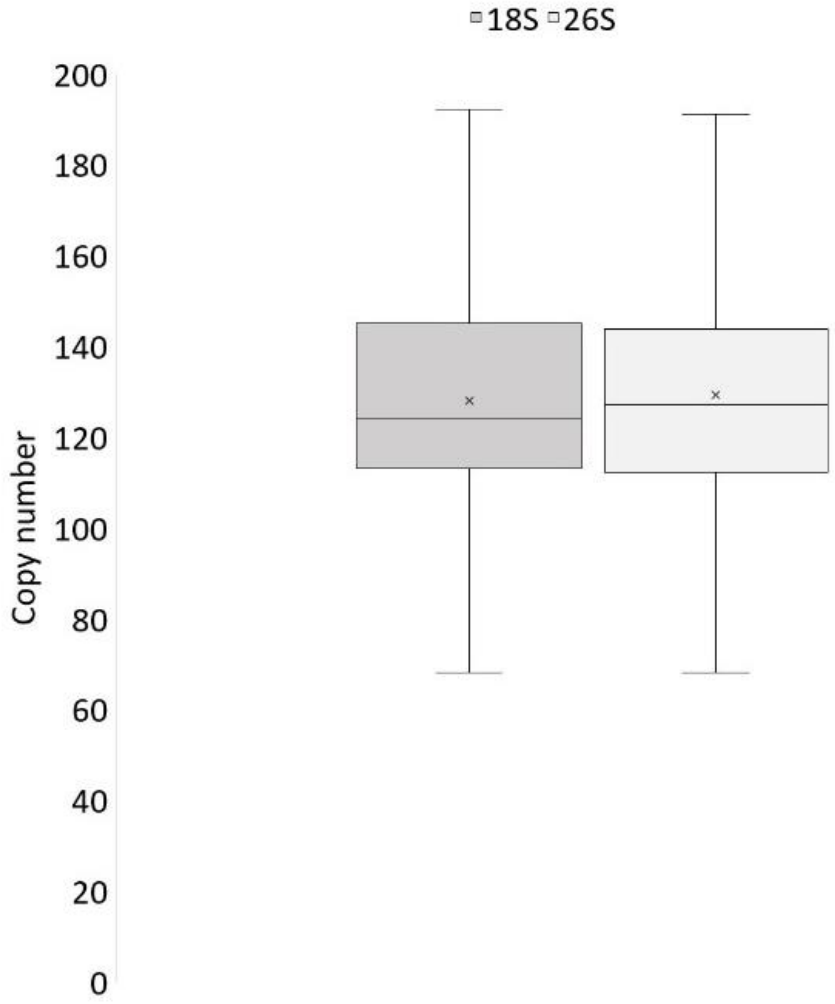


Figure 6. Copy numbers from natural populations. Left: 18S samples, right: 26S samples. The horizontal lines represent quartiles and the median, and x is the mean copy number.

3.2 Mutations and mutation rate in MA-lines

To assess whether the copy number changes were similar in 18S and 26S rDNA, correlation coefficient for 18S and 26S rDNA copy number was calculated for each

generation. In generation 5, there was no significant correlation between the samples ($r = 0.123$, $p = 0.649$). In contrast, a relatively high positive correlation was found in generation 20 ($r = 0.675$, $p = 0.002$) and in generation 40 ($r = 0.752$, $p < 0.001$), suggesting that mutations in copy number have a tendency to occur in complete repeat units.

Next, the number of mutations in each MA-line was calculated. This was carried out for 18S and 26S rDNA samples separately. In 18S samples (n = 38), including both mat a *and* mat A *lines, total of 16 copy number mutations could be detected in 13 lines. In 10 lines there was one mutation in each, and three of the lines had two mutations. 26S samples (n = 39) had a total of 20 mutations, and they occurred in 17 lines. In 14 of the lines only one mutation was detected, and 3 lines had two mutations. For more information about the lines and mutations, see tables 2 and 3.*

Table 2. *N. crassa rDNA mutations in mating type a*

	18S *mat a*		26S *mat a*	
sample	n	type	n	type
21a	2	ins/del	2	ins/del
22a	1	del	1	del
23a	0		0	
24a	0		0	
25a	0		0	
26a	0		0	
27a	0		0	
28a	1	del	1	del
29a	2	ins/del	1	del
30a	0		0	
31a	0		0	
33a	0		0	
34a	0		0	
35a	2	del/ins	0	
36a	0		2	ins/del
37a	0		1	del
38a	0		0	
39a	1	del	1	del
40a	0		0	
total	9		9	

When comparing *mat a* ancestors to generation 40, there were mutations in four lines in both 18S and 26S rDNA. These mutations were all deletions (figures 7A and 8A). In *mat A* lines, there were eight mutations in 18S and 11 mutations in 26S samples (figures 7B and 8B). In addition, in 10 lines mutations occurred simultaneously in 18S and 26S rDNAs when comparing all three (5, 20, 40) generations. In one case there was an insertion in 18S rDNA and a deletion in 26S rDNA, but all the other cases of simultaneous mutations were deletions. In 9 lines, mutations occurred only in either 18S or 26S rDNA. Overall, deletions were more common than insertions: there were 27 deletions and only 6 insertions. In all the six lines that had more than

Table 3. *N. crassa rDNA mutations in mating type A*

	18S *mat A*		26S *mat A*	
sample	n	type	n	type
1A	1	del	1	del
2A	1	del	1	del
3A	0		0	
4A	1	del	1	del
5A	0		0	
6A	0		1	del
7A	0		1	del
8A	0		0	
9A	0		0	
10A	0		1	del
11A	-		0	
12A	0		0	
13A	0		0	
14A	1	del	0	
15A	1	del	1	del
16A	1	del	1	del
17A	0		1	del
18A	0		0	
19A	1	del	0	
20A	0		2	ins/del
total	7		11	

one mutation, there was always both an insertion and a deletion. In five of these lines the insertion occurred first and the deletion after it, and only in one of the six lines an insertion was preceded by a deletion. Mutation rate was then calculated for 18S and 26S rDNA from the number of mutations. The mutation rate was 1.09 [0.64, 1.72] x 10^{-5} mutations/ mitosis in 18S rDNA and 1.30 [0.09, 1.95] x 10^{-5} mutations / mitosis in 26S rDNA. Because of a relatively low number of mutations in the MA lines, it was not possible to form a mutation model. The effects of selection on the mutation rate are shortly discussed in the discussion section.

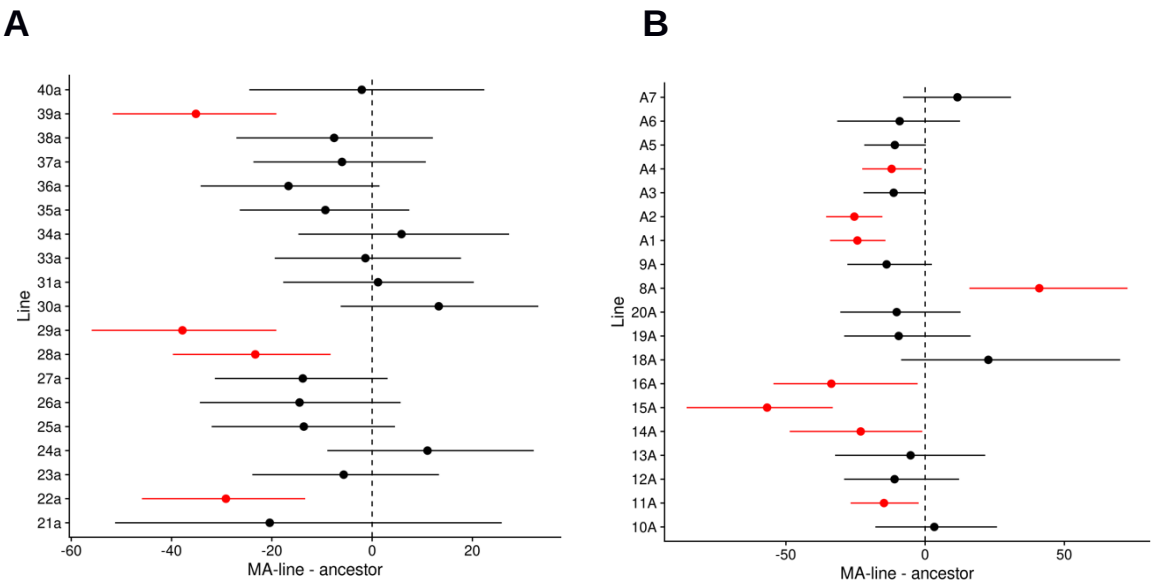


Figure 7. Relative mutation rate of 18S rDNA repeats. A) In *mat a* lines, comparison of copy number between ancestor and generation 40 lines revealed mutations in 4 lines. B) In *mat A* lines, 8 mutations were detected. Lines with mutations are shown in red.

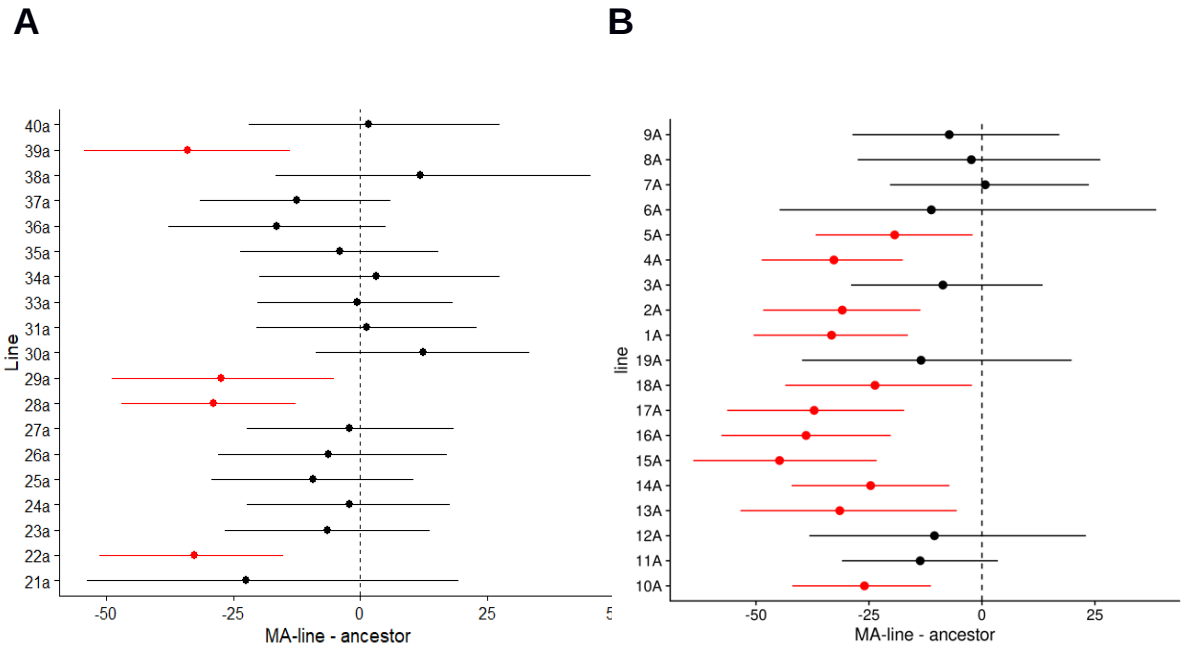


Figure 8. Relative mutation rate of 26S rDNA repeats. A) In *mat a* lines, comparison of copy number between ancestor and generation 40 lines revealed mutations in 4 lines. These are the same lines as in 18S rDNA B) In *mat A* lines, 11 mutations were detected. Lines with mutations are shown in red.

4 DISCUSSION

In this study, 18S and 26S rDNA copy numbers in *N. crassa* fungus were quantified, and mutation rate of rDNA was then estimated. The aim was to estimate the effect of selection on mutation rate by comparing results from mutation accumulation lines to natural population lines. Overall, there was a lot of copy number variation in both mutation accumulation lines and natural population lines. This was expected, because rDNA locus is one of the most unstable areas in the genome and has been suggested to have more copy number variation than any other part of the genome. Interestingly, in almost all the lines that were used in this study, the average copy number was somewhat lower than that in the literature (Free *et al.* 1979; Radford and Parish 1997; Galagan *et al.* 2003; Rodriguez *et al.* 2022). It is not likely that this was caused by a sampling error, because it occurred throughout the

lines. This poses a question how reliable are the methods which are used to quantify genome areas consisting of repeat sequences.

To verify the accuracy of the ddPCR method in copy number quantification, rDNA copy number was analyzed from two genes (18S and 26S) from the same locus, and these results were compared. On average, copy number seemed to be similar in 18S and 26S rDNAs, and this was observed in both MA lines and natural populations. Thus, ddPCR seems to be a valid method for copy number quantification, unless a same bias was present in both measurements. However, there was a considerable amount of variation in the results in individual measurements, and thus it was not possible to quantify the copy number very accurately, which made the further analyses somewhat difficult. In future studies other, more accurate copy number quantification methods should be considered.

The similarity in copy number can be explained by the fact that 18S and 26S rDNAs are in the same locus and almost next to each other. There is a very high probability that when a breakage in DNA strand occurs during cell division, both rDNAs will end up to the same segment and will then be either deleted or inserted together. Thus, one could expect that deletions and insertions take place simultaneously in both genes, which was indeed the case with all the generation 40 samples. However, our data from generations 5 and 20 showed that this is not always true. In half of the lines, deletions and insertions occurred somewhat synchronously, but in the other half they occurred only in one gene and not in the other. These results suggest that there is a slight tendency for the rDNA locus to go through mutations as a complete unit, but also incomplete deletions/insertions occur quite often. Despite this, the copy number of 18S and 26S rDNAs seem to remain somewhat similar in the two rDNAs, which might be explained by the stochastic nature of deletions and insertions.

The spontaneous mutation rate of 18S and 26S rDNA copy number was almost similar, 26S having a marginally higher rate. Overall, the rate was relatively high

compared to that of the whole *N. crassa* genome. This is in accordance with the recent study by Villalba de la Peña et al. (2022), in which they used short-read sequencing data from the same MA lines as in this experiment. They indicated that repeat units have a higher mutation rate than other genome areas (Villalba de la Peña *et al.* 2022). The repeats are suggested to be more vulnerable to deletions and insertions because they consist of repetitive DNA segments. When this type of DNA is broken, homologous recombination repair cannot necessarily reattach the segments to their correct locations. This is even more so with rDNA, which is transcriptionally exceptionally active and thus might be more vulnerable to breakages (Takeuchi *et al.* 2003; Helmrich *et al.* 2013). Moreover, there is evidence that non-transcribed genome regions might have a higher mutation rate than transcribed regions (Villalba de la Peña *et al.* 2022). Approximately half of the rDNA is silenced, and therefore the mutation rate could be hypothesized to be even higher in those regions. However, investigating differences in mutation rate between silenced and transcribed regions was out of scope of this study.

Mutation analysis of the MA-lines revealed that there were significantly more deletions than insertions throughout the lines. The rate of deletions was nearly five times as high as that of insertions. Villalba de la Peña et al. (2022) detected this as well, as they reported that in general the rate of deletions in repeat units was three times higher than that of insertions (Villalba de la Peña *et al.* 2022). Especially large deletions have been shown to be quite common in rDNA, which is suggested to be important for concerted evolution of rDNA (Lu et al. 2018). It has been indicated that aging or going through multiple mitotic cell divisions can cause a decrease in rDNA copy number in fruit fly *Drosophila melanogaster* (Lu *et al.* 2018) and in budding yeast *Saccharomyces cerevisiae* (Sinclair and Guarente 1997; Kobayashi 2011), and now it seems evident that this is the case also with *N. crassa*. Statistical analysis of *N. crassa* data revealed a decrease in copy number in 26S rDNA lines between generations 20 and 40, and although not statistically significant, the tendency was seen also in 18S rDNA. To sum up, in MA-lines the effect of selection

had been minimized and mutations were allowed to accumulate spontaneously, which caused a genetic drift towards a lower rDNA copy number.

Moreover, insertions were in most cases followed by a deletion, and only in one case was a deletion followed by an insertion. Restoring of rDNA copy number was thus detected in this study only in one direction: it seems that the increase of rDNA copy number was slowed down by deletions that occur soon after insertions. These results need to be confirmed, because the data used in this study was relatively small, and the missing of mating type A ancestors might have made the quantification of mutations somewhat inaccurate.

Contrary to the results of this study, some have reported that insertions occur as likely as deletions in rDNA (Lu *et al.* 2018). Molecular level amplification mechanisms have been shown to exist in some species, and so far there is evidence of this phenomenon for example from germ line cells in *Drosophila melanogaster* (Lu *et al.* 2018), but also from *Saccharomyces cerevisiae* (Gangloff et al. 1996; Kobayashi 2014). To my knowledge, there are no studies showing that rDNA amplification occurs in *N. crassa.* Yet, too low or too high copy number is known to be harmful for the organism, and it is reasonable to expect that some type of mechanism for the maintenance of the copy number should exist in *N. crassa* as well.

In natural population stains, maintenance of rDNA copy number was observed in this study. Based on these results it can be hypothesized that rDNA amplification in *N. crassa* could occur either a) during mitosis under selection pressure, or b) in meiosis, possibly as a consequence of some sort of "meiotic magnification" (Lu et al. 2018). The results support both of these suggestions: the rDNA copy number in natural populations was higher (mean ~129 copies) than that of generation 40 samples of MA lines (mean 104 copies), and closer to the copy number of generation 5 samples (mean ~118 copies). Thus, from MA-lines it was observed that without other mechanisms being involved, mutation pressure has a tendency to reduce the rDNA copy number. This type of tendency would have possibly been harmful in natural populations and it was balanced towards an optimal copy number in some

way. Because natural population strains may have been reproducing both asexually or sexually, it is equally likely that the amplification took place in either meiosis, or in mitosis under selection pressure. However, based on the results it is not possible to confirm which type of balancing mechanisms was in action.

In conclusion, the spontaneous mutation rate of rDNA in *N. crassa* is higher than the overall mutation rate of the whole genome, as expected. There seems to be more deletions than insertions, causing a genetic drift, in which rDNA copy number decreases relatively rapidly. However, because of a low number of mutations, it was not possible to form an accurate mutation model. Since the copy number in natural populations was higher than in the generation 40 MA-lines, and it is expected that rDNA is subject to balancing selection, it was hypothesized that some type of rDNA amplification likely occurs under selection pressure. In the future, it would be interesting to find out if this is really the case, and to explore the molecular mechanisms of detecting and balancing the changes in rDNA copy number.

There were several limitations in this study that should be considered when interpreting the results. First, only approximately half of the MA lines had DNA extracted from all three generations, and the rest had DNA only from generation 40. Because insertions and deletions often occurred in same lines, resulting first to an increase and right after it to a decrease in the copy number, these types of mutations might have been unnoticed when comparing only ancestors and generation 40 samples. Thus, some information about mutations was probably lost. Second, the problems with the copy number of mating type A ancestor made it difficult to estimate mutations in these lines. Third, the low number of mutations limited the possibility to make a mutation model. Therefore, the effects of selection on the mutation rate could only be hypothesized.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Associate professor Ilkka Kronholm and PhD Mariana Villalba de la Peña for the opportunity. Special thanks to Jussi Lehtonen for his support and patient guidance with statistical analyses.

REFERENCES

Aldrich J.C. & Maggert K.A. 2015. Transgenerational Inheritance of Diet-Induced Genome Rearrangements in Drosophila Cavalli G. (ed.). *PLOS Genetics* 11: e1005148.

Bughio F. & Maggert K.A. 2019. The peculiar genetics of the ribosomal DNA blurs the boundaries of transgenerational epigenetic inheritance. *Chromosome Research* 27: 19–30.

Cox R.A. & Peden K. 1979. A study of the organisation of the ribosomal ribonucleic acid gene cluster of Neurospora crassa by means of restriction endonuclease analysis and cloning in bacteriophage lambda. *Molecular and General Genetics MGG* 174: 17–24.

Free S.J., Rice P.W. & Metzenberg R.L. 1979. Arrangement of the genes coding for ribosomal ribonucleic acids in Neurospora crassa. *Journal of Bacteriology* 137: 1219–1226.

Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D., Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B., Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M., Qui D., Ianakiev P., Bell-Pedersen D., Nelson M.A., Werner-Washburne M., Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U., Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D., Roy A., Foley K., Naylor J., Stange-Thomann N., Barrett R., Gnerre S., Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D., Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S., Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A., DeSouza C.P.C., Glass L., Orbach M.J., Berglund J.A., Voelker R., Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R., Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M., Paulsen I., Sachs M.S., Lander E.S., Nusbaum C. & Birren B. 2003. The genome sequence of the filamentous fungus Neurospora crassa. *Nature* 422: 859–868.

Gangloff S., Zou H. & Rothstein R. 1996. Gene conversion plays the major role in controlling the stability of large tandem repeats in yeast. *The EMBO Journal* 15: 1715–1725.

Ganley A.R.D. & Kobayashi T. 2011. Monitoring the Rate and Dynamics of Concerted Evolution in the Ribosomal DNA Repeats of Saccharomyces cerevisiae Using Experimental Evolution. *Molecular Biology and Evolution* 28: 2883–2891.

Gibbons J.G., Branco A.T., Yu S. & Lemos B. 2014. Ribosomal DNA copy number is coupled with gene expression variation and mitochondrial abundance in humans. *Nature Communications* 5: 4850.

Guo Z., Han L., Liang Z. & Hou X. 2019. Comparative analysis of the ribosomal DNA repeat unit (rDNA) of *Perna viridis* (Linnaeus, 1758) and *Perna canaliculus* (Gmelin, 1791). *PeerJ* 7: e7644.

Halligan D.L. & Keightley P.D. 2009. Spontaneous Mutation Accumulation Studies in Evolutionary Genetics. *Annual Review of Ecology, Evolution, and Systematics* 40: 151–172.

Helmrich A., Ballarino M., Nudler E. & Tora L. 2013. Transcription-replication encounters, consequences and genomic instability. *Nature Structural & Molecular Biology* 20: 412–418.

Hindson B.J., Ness K.D., Masquelier D.A., Belgrader P., Heredia N.J., Makarewicz A.J., Bright I.J., Lucero M.Y., Hiddessen A.L., Legler T.C., Kitano T.K., Hodel M.R., Petersen J.F., Wyatt P.W., Steenblock E.R., Shah P.H., Bousse L.J., Troup C.B., Mellen J.C., Wittmann D.K., Erndt N.G., Cauley T.H., Koehler R.T., So A.P., Dube S., Rose K.A., Montesclaros L., Wang S., Stumbo D.P., Hodges S.P., Romine S., Milanovich F.P., White H.E., Regan J.F., Karlin-Neumann G.A., Hindson C.M., Saxonov S. & Colston B.W. 2011. High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Analytical Chemistry* 83: 8604–8610.

Ide S., Miyazaki T., Maki H. & Kobayashi T. 2010. Abundance of Ribosomal RNA Gene Copies Maintains Genome Integrity. *Science* 327: 693–696.

Katju V. & Bergthorsson U. 2019. Old Trade, New Tricks: Insights into the Spontaneous Mutation Process from the Partnering of Classical Mutation Accumulation Experiments with High-Throughput Genomic Approaches Makova K. (ed.). *Genome Biology and Evolution* 11: 136–165.

Kobayashi T. 2011. How does genome instability affect lifespan?: Roles of rDNA and telomeres. *Genes to Cells* 16: 617–624.

Kobayashi T. 2014. Ribosomal RNA gene repeats, their stability and cellular senescence. *Proceedings of the Japan Academy, Series B* 90: 119–129.

Li X., Xu J., He Y., Shen S., Zhu J. & Shen Z. 2016. The complete nuclear ribosomal DNA (nrDNA) cistron sequence of Pyropia yezoensis (Bangiales, Rhodophyta). *Journal of Applied Phycology* 28: 663–669.

Lofgren L.A., Uehling J.K., Branco S., Bruns T.D., Martin F. & Kennedy P.G. 2019. Genome-based estimates of fungal rDNA copy number variation across phylogenetic scales and ecological lifestyles. *Molecular Ecology* 28: 721–730.

Lu K.L., Nelson J.O., Watase G.J., Warsinger-Pepe N. & Yamashita Y.M. 2018. Transgenerational dynamics of rDNA copy number in Drosophila male germline stem cells. *eLife* 7: e32421.

Lyckegaard E.M. & Clark A.G. 1989. Ribosomal DNA and Stellate gene copy number variation on the Y chromosome of Drosophila melanogaster. *Proceedings of the National Academy of Sciences* 86: 1944–1948.

Metzenberg R.L. 2003. Vogel's Medium N salts: avoiding the need for ammonium nitrate. *Fungal Genetics Reports* 50: 14–14.

Moss T. & Stefanovsky V.Y. 2002. At the Center of Eukaryotic Life. *Cell* 109: 545–548.

Okazaki A., Yamazaki S., Inoue I. & Ott J. 2021. Population genetics: past, present, and future. *Human Genetics* 140: 231–240.

Prokopowich C.D., Gregory T.R. & Crease T.J. 2003. The correlation between rDNA copy number and genome size in eukaryotes. *Genome* 46: 48–50.

Radford A. & Parish J.H. 1997. The Genome and Genes of Neurospora crassa. *Fungal Genetics and Biology* 21: 258–266.

Rodriguez S., Ward A., Reckard A.T., Shtanko Y., Hull-Crew C. & Klocko A.D. 2022. The genome organization of *Neurospora crassa* at high resolution uncovers principles of fungal chromosome topology Dekker J. (ed.). *G3 Genes|Genomes|Genetics*: jkac053.

Salim D., Bradford W.D., Freeland A., Cady G., Wang J., Pruitt S.C. & Gerton J.L. 2017. DNA replication stress restricts ribosomal DNA copy number Hopper A.K. (ed.). *PLOS Genetics* 13: e1007006.

Sinclair D.A. & Guarente L. 1997. Extrachromosomal rDNA Circles— A Cause of Aging in Yeast. *Cell* 91: 1033–1042.

Takeuchi Y., Horiuchi T. & Kobayashi T. 2003. Transcription-dependent recombination and the role of fork collision in yeast rDNA. *Genes & Development* 17: 1497–1506.

Villalba de la Peña M., Summanen P.A.M., Liukkonen M. & Kronholm I. 2022. *Variation in spontaneous mutation rate and spectrum across the genome of* Neurospora crassa. Genetics.

APPENDIX 1. *Neurospora crassa* natural population strains

10948	10906	10915
10886	10897	10916
10932	10896	10918
1165	10917	10919
4498	10931	10920
8816	10934	10921
3223	10909	10908
8845	10911	10925
10908	10892	10936
847	10891	10929
10904	10894	10930
851	10895	10938
1131	10982	10943
8850	10983	10941
8819	10887	10935
4712	10882	10926
6203	10889	10928
4824	10888	10942
8783	10890	10893
8790	10901	8848
3975	10900	8851
10928	2489	8787
10912	10907	1132
4494	10881	4713
3210	10883	8789
10923	10844	3200
10950	10922	5914
10951	10905	4715
10946	10903	8829
3211	10954	3968
3943		8784

APPENDIX 2. Primers and probes

Name	Sequence (5′-3′)	Notes
Tba1_v1_F	TCTTGTTCGTCTGGAGTCTG	
Tba1_v1_R	CTGATGGGATTTCGAGGTTG	
Tba1_v2_F	ACGGGTGATCTCCAAGTATG	
Tba1_v2_R	CCCGCATCTTTCTACATGTG	
Ncv1_18S_F	TCCAGACACGATGAGGATTG	
Ncv1_18S_R	ACTCCACCAACTAAGAACGG	
Ncv2_18S_F	AACTCACCAGGTCCAGACAC	
Ncv2_18S_R	CCACCACCCACGAAATCAAG	
Ncv1_26S_F	GTAGTTTGGGAATGCTGCTC	

Ncv1_26S_R	CACTCTACTTGTGCGCTATC	
Ncv2_26S_F	CAAGGATGCTGGCGTAATGG	
Ncv2_26S_R	TTTACACCCAAACACTCGCG	
Tba1_probe	6-FAM/ACCGCTTGGCTTGGCATCTTC/BHQ-1	A fluorescent 6-FAM stamp in 5′-end and a Black Hole Quencher 1 in 3′-end
Nc_18S_probe	Hex/TGCACCACCACCCACGAAATC/BHQ-1	A fluorescent HEX stamp in 5′-end and Black Hole Quencher 1 in 3′-end
Nc_26S_probe	Hex/AACCTTGACTCCTTGGTCCGTG/BHQ-1	A fluorescent HEX stamp in 5′-end and Black Hole Quencher 1 in 3′-end