**Supplementary Information**

|  |
| --- |
| **Table S1. *Neurospora* species and strains investigated in this study.** |
| **Species****(Lineage)** | **Straina** | **Mating type** | **Sample IDe** | **Genome reference** |
| *N. tetrasperma*(Lineage 1) | FGSC9033b | A | Ntet-L1\_A | Sun et al. (2017) |
| FGSC9034b | a | Ntet-L1\_a | Sun et al. (2017) |
| *N. tetrasperma*(Lineage 6) | FGSC2508c | A | Ntet-L6\_A | Ellison et al. (2011) |
| FGSC2509c | a | Ntet-L6\_a | Ellison et al. (2011) |
| *N. tetrasperma*(Lineage 10) | FGSC10707d | A | Ntet-L10\_A | Sun et al. (2017) |
| FGSC10708d | a | Ntet-L10\_a | Sun et al. (2017) |
| a FGSC; Fungal Genetics Stock Centre. For *N. tetrasperma*, the homokaryotic strains originate from heterokaryotic (n+n) tissue of strains isolated from natureb-e, and were isolated in the laboratory by using asexual spores (conidia) of single mating type (either *mat A* or *mat a*) occasionally produced by the heterokaryon (Raju 1992; Corcoran et al., 2012). b Originally isolated from P44992 heterokaryon.c Originally isolated from P581 heterokaryon.d Originally isolated from UK33 heterokaryon.e Fungal tissues harvested at two developmental stages: vegetative growth and sexual development.  |

|  |
| --- |
| **Table S2. Number and proportion of genes with different modes of expression in heterokaryons as compared to homokaryons during the vegetative stage of development.** The total amount of expression (i.e., *mat A* + *mat a*) of each gene in heterokaryons is compared to the amount in each homokaryon. Percentages are given as proportion out of the number of genes in the group that is one level higher. |
|  | **L1** | **L6** | **L10** |
| Total number of genes expressed | 7846 | 7800 | 7767 |
|  | Genes with conserved expression |  | 4596 (58.6%) |  | 7534 (96.6%) |  | 1899 (24.4%) |
|  | Genes without conserved expression |  | 3250 (41.4%) |  | 266 (3.4%) |  | 5868 (75.6%) |
|  |  | Genes differing in expression between *mat A* and *mat a* |  |  | 2114 (65.0%) |  |  | 244 (91.7%) |  |  | 4550 (77.5%) |
|  |  |  | Additive expressiona |  |  |  | 2102 (99.4%) |  |  |  | 229 (93.8%) |  |  |  | 2273 (50.0%) |
|  |  |  | *Mat a* dominance in heterokaryonsb |  |  |  | 0 (0.0%) |  |  |  | 0 (0.0%) |  |  |  | 25 (0.5%) |
|  |  |  | *Mat A* dominance in heterokaryonsb |  |  |  | 0 (0.0%) |  |  |  | 0 (0.0%) |  |  |  | 2 (0.0%) |
|  |  |  | Overdominantc |  |  |  | 0 (0.0%) |  |  |  | 0 (0.0%) |  |  |  | 3 (0.1%) |
|  |  |  | Underdominantc |  |  |  | 0 (0.0%) |  |  |  | 0 (0.0%) |  |  |  | 6 (0.1%) |
|  |  |  | Other |  |  |  | 12 (0.6%) |  |  |  | 15 (6.1%) |  |  |  | 2241 (49.3%) |
|  |  | Genes with another pattern of expression difference |  |  | 1136 (35.0%) |  |  | 22 (8.3%) |  |  | 1318 (22.5%) |
| a Genes for which *mat A* and *mat a* expression differ in homokaryons and for which the amount of expression in heterokaryons is in accordance with the ratio of *mat A* and *mat a* nuclei (see also Methods)b Genes for which *mat A* and *mat a* expression differ in homokaryons, whereby the expression is equal to the level in one homokaryon and differs from the level in the other, and for which the expression also differs from additive expressionc Genes for which *mat A* and *mat a* expression differ in homokaryons, whereby the expression is significantly different from the level in both homokaryons, respectively higher or lower. |

|  |
| --- |
| **Table S3. 95% confidence interval for coefficients of linear regression analysis.** |
| L1 | Variable(Intercept) | 95% CI (sex)[0.052, 0.068] | 95% CI (veg)[0.018, 0.022] |  |  |
| slope | [0.83, 0.84] | [0.97, 0.98] |  |  |
| L6 | (Intercept) | [0.241, 0.251] | [-0.008, -0.002] |  |  |
| slope | [0.52, 0.54] | [0.93, 0.94] |  |  |
| L10 | (Intercept) | [-1.225, -1.187] | [0.106, 0.112] |  |  |
| slope | [2.18, 2.22] | [0.88, 0.89] |  |  |

|  |
| --- |
| **Table S4. Number of genes in the vegetative stage of development exhibiting a slope of the regressed interaction between nuclear ratio and gene expression ratio confidence interval is not overlapping 1.** Numbers only include genes with positional information. Numbers in brackets represent percentages out of the number of genes in the group that is one level higher. |
| **Groupa** | **L1** | **L6** | **L10** |
| Total no. genes in the regression analysis | 550 | 1002 | 2078 |
|  No. genes with slopeveg-hetero < 1 |  53 (10%) |  227 (23%) |  1103 (53%) |
|  on autosomes |  1 (2%) |  2 (1%) |  615 (56%) |
|  on *mat* chromosome PAR |  0 (2%) |  0 (0%) |  44 (4%) |
|  on *mat* chromosome non-recombining region |  52 (98%) |  225 (99%) |  444 (40%) |
|  No. genes with slopeveg-hetero > 1 |  0 (0%) |  5 (0%) |  22 (1%) |
|  on autosomes |  0 (0%) |  0 (0%) |  9 (41%) |
|  on *mat* chromosome PAR |  0 (0%) |  0 (0%) |  2 (9%) |
|  on *mat* chromosome non-recombining region |  0 (0%) |  5 (100%) |  11 (50%) |
|  No. genes with slopeveg-hetero ≅ 1  |  497 (90%) |  770 (77%) |  953 (46%) |
|  on autosomes |  2 (0%) |  1 (0%) |  551 (58%) |
|  on *mat* chromosome PAR |  15 (3%) |  0 (0%) |  51 (5%) |
|  on *mat* chromosome non-recombining region |  480 (97%) |  769 (100%) |  351 (37%) |
| aAutosomes refer to genes located on any of the chromosomes not harboring the *mat*-locus. *Mat* chromosome PAR refers to the recombining region at the freely recombining ends of the chromosome harboring the *mat*-locus. |

|  |
| --- |
| **Table S5. Number of genes in the sexual stage of development exhibiting a slope of the regressed interaction between nuclear ratio and gene expression ratio confidence interval is not overlapping 1.** Numbers only include genes with positional information. Numbers in brackets represent percentages out of the number of genes in the group that is one level higher. |
| **Groupa** | **L1** | **L6** | **L10** |
| Total no. genes in the regression analysis | 550 | 1002 | 2078 |
|  No. genes with slopesex-hetero < 1  |  467 (85%) |  975 (97%) |  15 (0%) |
|  on autosomes |  3 (1%) |  3 (0%) |  10 (67%) |
|  on *mat* chromosome PAR |  11 (2%) |  0 (0%) |  1 (7%) |
|  on *mat* chromosome non-recombining region |  453 (97%) |  972 (100%) |  4 (26%) |
|  No. genes with slopesex-hetero > 1 |  0 (0%) |  0 (0%) |  1857 (89%) |
|  on autosomes |  0 (0%) |  0 (0%) |  1046 (56%) |
|  on *mat* chromosome PAR |  0 (0%) |  0 (0%) |  85 (5%) |
|  on *mat* chromosome non-recombining region |  0 (0%) |  0 (0%) |  726 (39%) |
|  No. genes with slopesex-hetero ≅ 1 |  83 (15%) |  27 (3%) |  206 (10%) |
|  on autosomes |  0 (0%) |  0 (0%) |  119 (58%) |
|  on *mat* chromosome PAR |  4 (5%) |  0 (0%) |  11 (5%) |
|  on *mat* chromosome non-recombining region |  79 (95%) |  27 (100%) |  76 (37%) |
| aAutosomes refer to genes located on any of the chromosomes not harboring the *mat*-locus. *Mat* chromosome PAR refers to the recombining region at the freely recombining ends of the chromosome harboring the *mat*-locus. |

|  |
| --- |
| **Table S6. Number of genes in the sexual stage of development exhibiting a slope of the regressed interaction between nuclear ratio and gene expression ratio that is significantly different from the slope in the vegetative stage.** Numbers only include genes with positional information. Numbers in brackets represent percentages out of the number of genes in the group that is one level higher. |
| **Groupa** | **L1** | **L6** | **L10** |
| Total no. genes in the regression analysis | 550 | 1002 | 2078 |
|  No. genes with slopesex < slopeveg |  414 (75%) |  958 (96%) |  12 (0%) |
|  on autosomes |  2 (1%) |  3 (0%) |  9 (75%) |
|  on MT chromosome PAR |  9 (2%) |  0 (0%) |  1 (8%) |
|  on MT chromosome non-recombining region |  403 (97%) |  955 (100%) |  2 (17%) |
|  No. genes with slopesex > slopeveg |  0 (0%) |  0 (0%) |  2014 (97%) |
|  on autosomes |  0 (0%) |  0 (0%) |  1135 (56%) |
|  on MT chromosome PAR |  0 (0%) |  0 (0%) |  93 (5%) |
|  on MT chromosome non-recombining region |  0 (0%) |  0 (0%) |  786 (39%) |
|  No. genes with slopesex ≅ slopeveg |  136 (25%) |  44 (4%) |  52 (3%) |
|  on autosomes |  1 (1%) |  0 (0%) |  31 (59%) |
|  on MT chromosome PAR |  6 (4%) |  0 (0%) |  3 (6%) |
|  on MT chromosome non-recombining region |  129 (95%) |  44 (100%) |  18 (35%) |
| aAutosomes refer to genes located on any of the chromosomes not harboring the *mat*-locus. *Mat* chromosome PAR refers to the recombining region at the freely recombining ends of the chromosome harboring the *mat*-locus. |

|  |
| --- |
| **Table S7. Number of genes in the sexual stage of development exhibiting an intercept of the regressed interaction between nuclear ratio and gene expression ratio that is significantly different from 0.** Numbers only include genes with positional information. Numbers in brackets represent percentages out of the number of genes in the group that is one level higher. |
| **Groupa** | **L1** | **L6** | **L10** |
| Total no. genes in the regression analysis | 83 | 27 | 206 |
|  No. genes with interceptsex < 0 |  0 (0%) |  0 (0%) |  26 (13%) |
|  on autosomes |  0 (0%)  |  0 (0%) |  18 (69%) |
|  on MT chromosome PAR |  0 (0%)  |  0 (0%) |  2 (8%) |
|  on MT chromosome non-recombining region |  0 (0%)  |  0 (0%) |  6 (23%) |
|  No. genes with interceptsex > 0 |  21 (25%) |  17 (63%) |  0 (0%) |
|  on autosomes |  0 (0%) |  0 (0%) |  0 (0%) |
|  on MT chromosome PAR |  0 (0%) |  1 (6%) |  0 (0%) |
|  on MT chromosome non-recombining region |  21 (100%) |  16 (94%) |  0 (0%) |
|  No. genes with interceptsex ≅ 0 |  62 (75%) |  10 (37%) |  180 (87%) |
|  on autosomes |  0 (0%) |  0 (0%) |  101 (56%) |
|  on MT chromosome PAR |  4 (6%) |  1 (10%) |  9 (5%) |
|  on MT chromosome non-recombining region |  58 (94%) |  9 (90%) |  70 (39%) |
| aAutosomes refer to genes located on any of the chromosomes not harboring the *mat*-locus. *Mat* chromosome PAR refers to the recombining region at the freely recombining ends of the chromosome harboring the *mat*-locus. |

|  |
| --- |
| **Table S8. Number of genes with *mat*-biased and developmental stage-biased expression in each lineage.** |
| **Lineage** | **Vegetative-biased genes** | **Sexual-biased genes** | ***mat A*-biased genes** | ***mat a*-biased genes** |
| L1 | 294 | 633 | 81 | 174 |
| L6 | 602 | 964 | 115 | 105 |
| L10 | 198 | 505 | 253 | 397 |



**Fig. S1. Principal component analysis of orthologs gene expression in *N. tetrasperma* L1 (A), L6 (B) and L10 (C).** Points represent samples for each developmental stage and nuclear ratio. The first two principal components are plotted, with the proportion of variance explained by each component shown on the axes labels. The 500 most diverging genes were used in this figure.



**Fig. S2. Principal component analysis of orthologs gene expression in *N. tetrasperma* L1 (A), L6 (B) and L10 (C).** Points represent samples for each developmental stage and nuclear ratio. The first and third principal components are plotted, with the proportion of variance explained by each component shown on the axes labels. The 500 most diverging genes were used in this figure.



**Fig. S3. Principal component analysis of orthologs gene expression in the vegetative (A) and sexual (B) stages of development across all *N. tetrasperma* lineages.** Points represent samples for each lineage and nuclear ratio. The first two principal components are plotted, with the proportion of variance explained by each component shown on the axes labels. The 500 most diverging genes were used in this figure.

**Fig S4**

****

**Fig. S4. Distribution of slope and intercept of linear models of individual genes.** For each gene in the three lineages (Lineage 1, Lineage 6 and Lineage 10) and in the two tissues (sex and veg) two linear models (Heterokaryons and Homokaryons) were created for the correlation between the relative fraction of RNA and DNA levels from the mat a and mat A orthologs of the gene. Homokaryons models were calculated using only values from the two Homokaryons mat a and mat A. Heterokaryons models were calculated using only the mixed samples with different concentration of mat a and mat A. (A) visualises the distribution of intercept values and (B) the slope distribution. Each color and the 2x2 grid identify which of the distribution it represents. Lineage 1 is represented by the blue distribution, Lineage 6 by the red and Lineage 10 by the green. Upper row represents the distribution from the sex tissue. The lower row represents the distribution in the vegetative tissue. The left column represents the Heterokaryons samples and the right columns represents the Homokaryons samples. Black vertical line shows the outcome of a one to one correlation between DNA concentration levels and RNA concentrations levels, i.e intercept equals 0 and slope equals 1. Note that for narrow peaks we do not show the upper part of the peak.

**References**

Ellison CE, Stajich JE, Jacobson DJ, Natvig DO, Lapidus A, Foster B, Aerts A, Riley R, Lindquist EA, Grigoriev IV et al. 2011. Massive changes in genome architecture accompany the transition to self-fertility in the filamentous fungus *Neurospora tetrasperma*. Genetics 189: 55-69.

Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma L-J, Smirnov S, Purcell S et al. 2003. The genome sequence of the filamentous fungus *Neurospora crassa*. Nature 422: 859-868.

Sun Y, Svedberg J, Hiltunen M, Corcoran P, Johannesson H. 2017. Large-scale suppression of recombination predates genomic rearrangements in *Neurospora tetrasperma*. Nat Commun 8: 1140