Additional file 4: identification of inverted repeats in *T. reesei* and *F. graminearum* centromeres

In 4 cases (scaffolds 51 (chr I), 56 (chr IV), 57 and 58), we observed in centromere scaffolds an inverted repeat structure with a central core region of 1 to 2 kb surrounded by an inverted repeat of 2.5 to 5 kb (Figure S2A). This structure seems quite similar to the centromere structure of *S. pombe* [1–3], *Candida tropicalis* [4] and *Komagataella phaffii* (formerly *Pichia partoris*) [5] (Figure S2B).

We annotated these sequences “mid” for the central cores and , “LR” for the left repeat, and “RR” for the right repeat, consistently with *C. tropicalis* and *K. phaffii* [4, 5], followed by the chromosome or scaffold number (Figure S2A below). The LR4 and RR4 sequences of the inverted repeat of chr IV centromere (scaffold 56) share 92% identity on 4kb without any gaps. In the 3 other cases, the LR and RR sequences share ≈58% identity but with large gaps (identity reaches 84 to 92% while excluding gaps).

While these observations could result from a misassembly of these AT-rich regions, they suggest that centromere structure in *Trichoderma* is significantly different from what is described in *Neurospora* and other filamentous fungi [6], and share some similarities with structures observed in *Taphrinomycotina* and *Saccharomycetales*.

Moreover, using the latest *Fusarium graminearum* genome release [7], we observed undescribed similar inverted repeats in the centromeres of *F. graminearum* chromosomes 1 and 2 (Figure S2B).

---

**Figure S2A: Inverted repeats found on centromere-involved scaffolds**

Four similar structures with a central core (mid) region surrounded by an inverted repeat (LR and RR) sequences were identified on 4 scaffolds involved in *T. reesei* centromeres (scaffold 56 in chr IV centromere, scaffold 51 in chr VI centromere, and scaffolds 57 and 58 with centromere signature but not assembled).
Figure S2B: Sequence alignment of centromeres on themselves

Core centromere sequences (containing LR, mid and RR sequences) have been aligned against themselves using the LASTZ software [8, 9] with default parameters, in order to show the inverted repeats. This figure includes the 4 sequences from *T. reesei*, and arbitrary chosen sequences from *S. pombe*, *K. phafii*, *C. tropicalis* and *F. graminearum*.

<table>
<thead>
<tr>
<th>Strain and location</th>
<th>Sequence alignment on itself using LASTZ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T. reesei CEN4</strong></td>
<td>9,130 bp</td>
</tr>
<tr>
<td>Scaffold_56:19,652-28,781</td>
<td></td>
</tr>
<tr>
<td><strong>T. reesei CEN6</strong></td>
<td>7,287 bp</td>
</tr>
<tr>
<td>Scaffold_51:8,571-15,857</td>
<td></td>
</tr>
<tr>
<td><strong>T. reesei CEN57</strong></td>
<td>10,548 bp</td>
</tr>
<tr>
<td>Scaffold_57:4,556-15,103</td>
<td></td>
</tr>
</tbody>
</table>
**T. reesei CEN58**
6,812 bp
Scaffold_58:8,532-15,343

---

**Schizosaccharomyces pombe CEN2**
41,139 bp
Chromosome II:
1,602,264-1,6447,747

---

**Komagataella phaffii (Pichia pastoris) CEN2**
6,655 bp
Chromosome 2 (FR839629.1):
843,845-850,499

---

**Candida tropicalis CEN5**
10b113 bp
Supercontig3.5 (GG692399.1):
718,785-728,897
Fusarium graminearum CEN1
9,818 bp
Chromosome 1:
8,976,756-8,986,573

Fusarium graminearum CEN2
11,796 bp
Chr2:
3,288,357-3,330,152

References
