**Supplemental information**
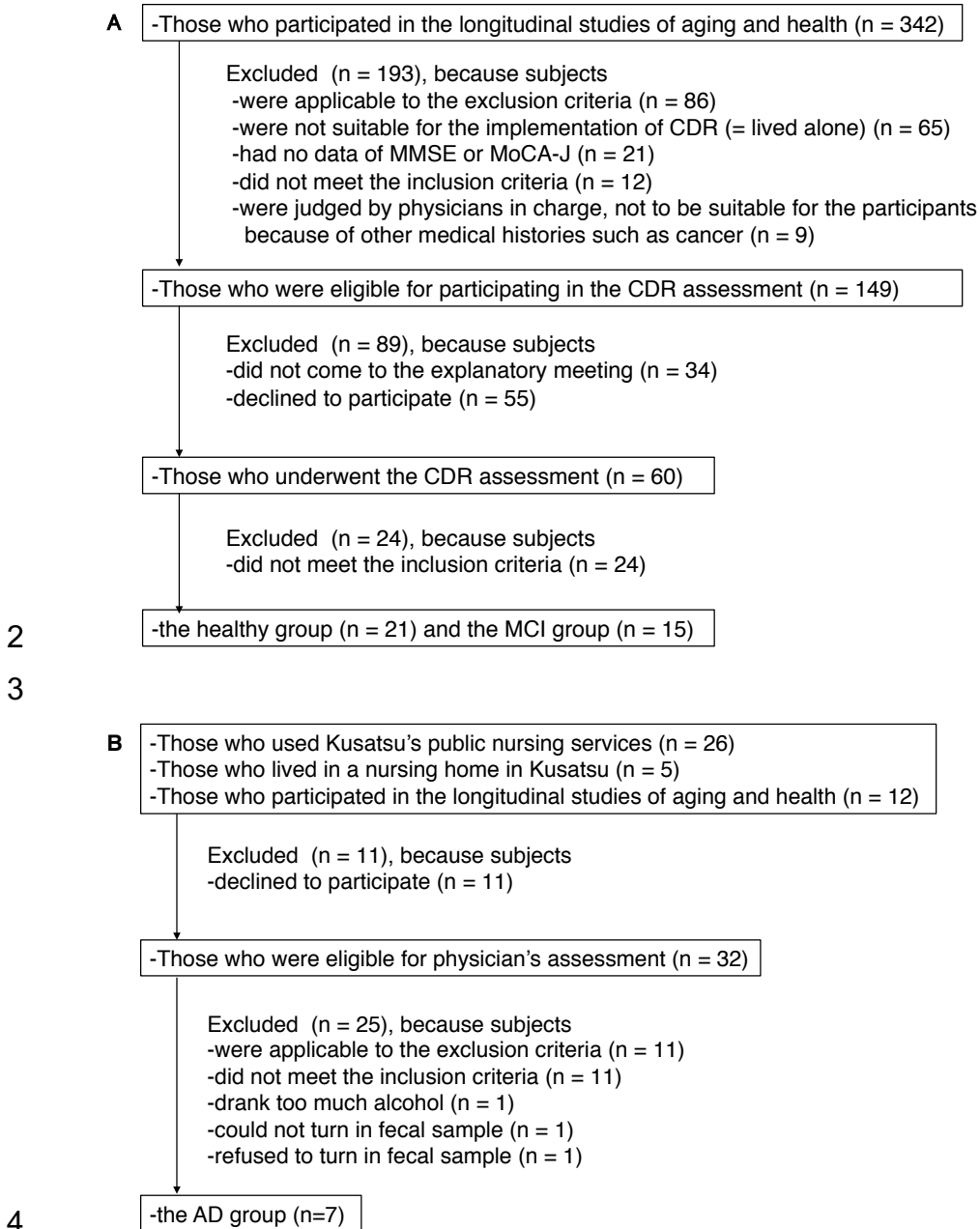
# Identification of *Faecalibacterium prausnitzii*

# strains for gut microbiome-based intervention

# in Alzheimer's-type dementia

Atsushi Ueda, Shoji Shinkai, Hirotsugu Shiroma, Yu Taniguchi, Sayaka Tsuchida, Takahiro Kariya, Tomohiro Kawahara, Yodai Kobayashi, Noriyuki Kohda, Kazunari Ushida, Akihiko Kitamura, and Takuji Yamada

1    **Supplemental information:**

A    -Those who participated in the longitudinal studies of aging and health (n = 342)

    Excluded  (n = 193), because subjects
    -were applicable to the exclusion criteria (n = 86)
    -were not suitable for the implementation of CDR (= lived alone) (n = 65)
    -had no data of MMSE or MoCA-J (n = 21)
    -did not meet the inclusion criteria (n = 12)
    -were judged by physicians in charge, not to be suitable for the participants
      because of other medical histories such as cancer (n = 9)

    -Those who were eligible for participating in the CDR assessment (n = 149)

    Excluded  (n = 89), because subjects
    -did not come to the explanatory meeting (n = 34)
    -declined to participate (n = 55)

    -Those who underwent the CDR assessment (n = 60)

    Excluded  (n = 24), because subjects
    -did not meet the inclusion criteria (n = 24)

2    -the healthy group (n = 21) and the MCI group (n = 15)

3

B    -Those who used Kusatsu's public nursing services (n = 26)
    -Those who lived in a nursing home in Kusatsu (n = 5)
    -Those who participated in the longitudinal studies of aging and health (n = 12)

    Excluded  (n = 11), because subjects
    -declined to participate (n = 11)

    -Those who were eligible for physician's assessment (n = 32)

    Excluded  (n = 25), because subjects
    -were applicable to the exclusion criteria (n = 11)
    -did not meet the inclusion criteria (n = 11)
    -drank too much alcohol (n = 1)
    -could not turn in fecal sample (n = 1)
    -refused to turn in fecal sample (n = 1)

4    -the AD group (n=7)

5    **Figure S1. Flow diagram of the selection of the subjects. Related to Figure 2.**

6    (A) Flow diagram showing the selection of subjects in the healthy group and the MCI group. (B) Flow
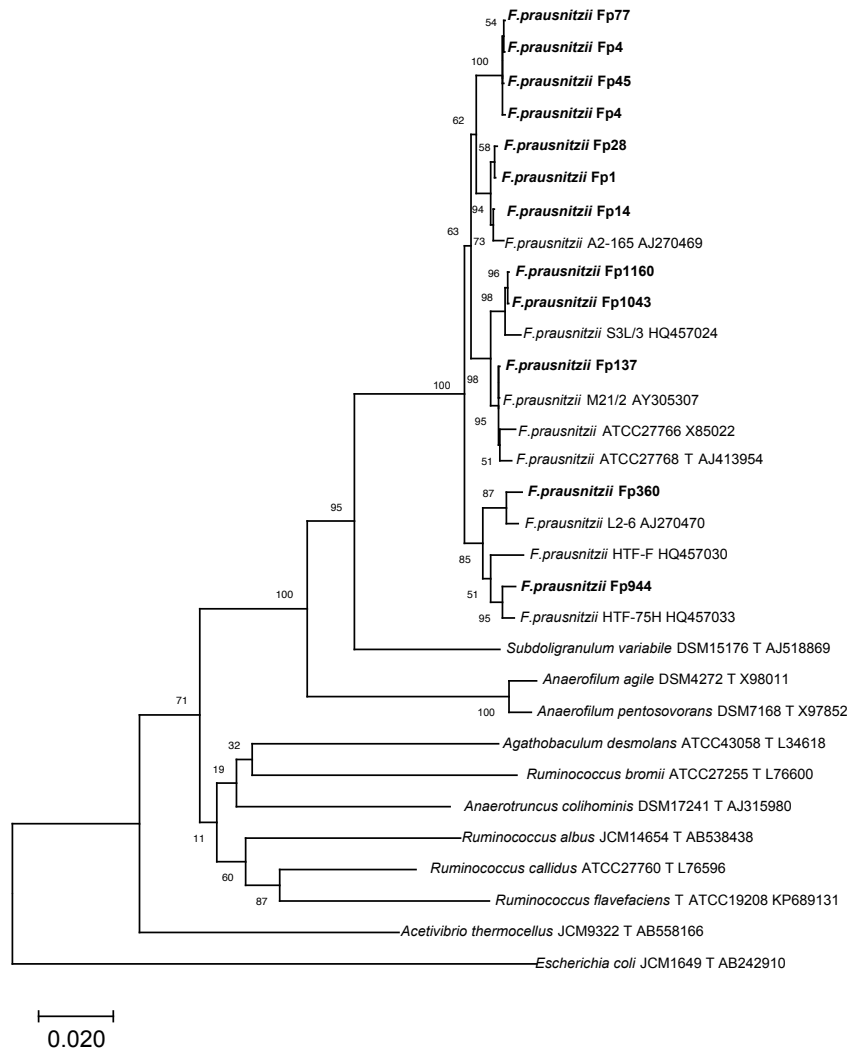
7    diagram showing the selection of subjects in the AD group.

8
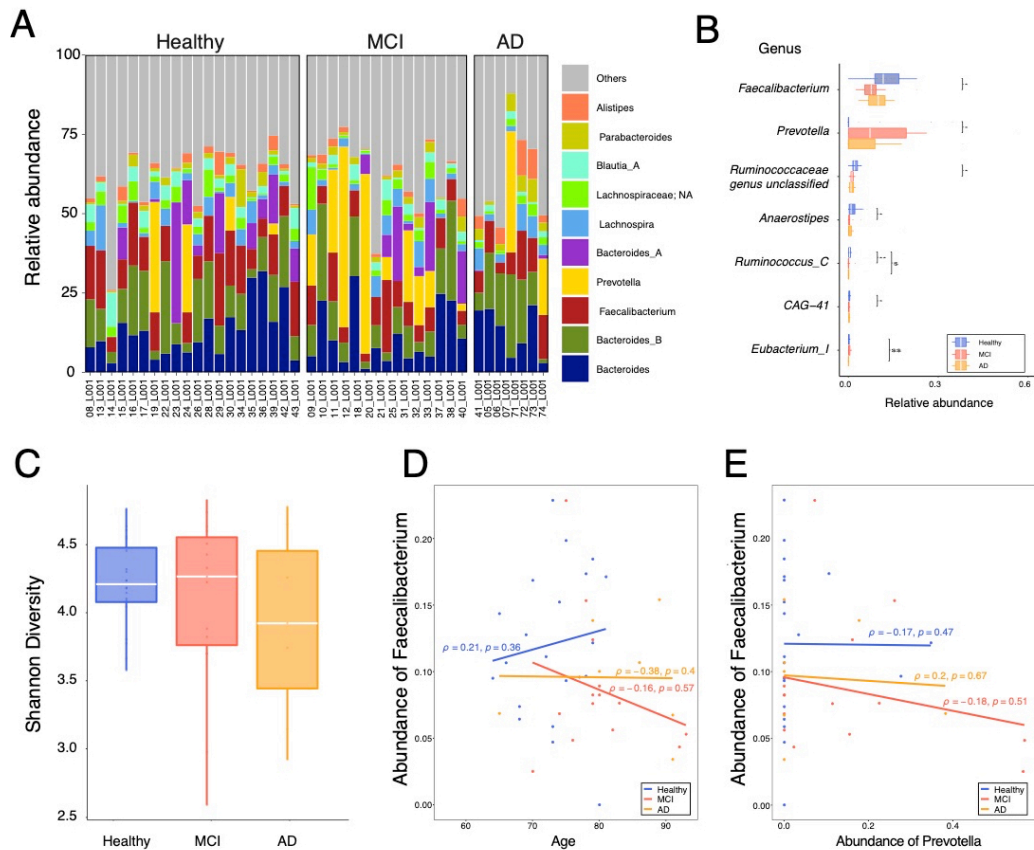
**Figure S2. Plot showing the relative abundance of genera, alpha-diversity, and the correlation between the relative abundance of *F. prausnitzii* and age or the relative abundance of *Prevotella*. Related to Figure 2 and Table S4.**

(A) Stacked bar plot showing the overall genus composition in each individual. (B) Box plot showing the interquartile range (IQR) of the relative abundance of genera. Seven genera whose abundance was significantly different either between the healthy (n=20) and the MCI group (n=15) or between the healthy and the AD group (n=7) are shown (Wilcoxon rank sum test, the healthy vs the MCI: $^{**}p < 0.01$, $^{*}p < 0.05$, the healthy vs the AD: $^{\$\$}p < 0.01$, $^{\$}p < 0.05$). (C) Box plot showing the IQR of the Shannon-Wiener alpha-diversity index, which did not differ between groups. (D, E) Scatter plot showing Spearman's ρ and p value between the relative abundance of *F. prausnitzii* and age or the relative abundance of *Prevotella* of the healthy (blue), the MCI (red), and the AD group (orange). There was no significant correlation between the abundance of *Faecalibacterium* and age or the abundance of *Prevotella*.

```
              54 ┌ F.prausnitzii Fp77
                 ├ F.prausnitzii Fp4
             100 │
                 ├── F.prausnitzii Fp45
          62 │   └─ F.prausnitzii Fp4
             58 ┌ F.prausnitzii Fp28
                ├ F.prausnitzii Fp1
             94 │
                ├ F.prausnitzii Fp14
          63 73 └─ F.prausnitzii A2-165 AJ270469
             96 ┌ F.prausnitzii Fp1160
             98 ├ F.prausnitzii Fp1043
                └── F.prausnitzii S3L/3 HQ457024
             98 ┌ F.prausnitzii Fp137
        100     ├ F.prausnitzii M21/2 AY305307
             95 ┌ F.prausnitzii ATCC27766 X85022
                └─ F.prausnitzii ATCC27768 T AJ413954
             87 ┌ F.prausnitzii Fp360
                └ F.prausnitzii L2-6 AJ270470
        85      ├ F.prausnitzii HTF-F HQ457030
             51 ┌ F.prausnitzii Fp944
             95 └ F.prausnitzii HTF-75H HQ457033
    95          Subdoligranulum variabile DSM15176 T AJ518869
  100           ┌ Anaerofilum agile DSM4272 T X98011
                100 Anaerofilum pentosovorans DSM7168 T X97852
71              32 Agathobaculum desmolans ATCC43058 T L34618
             19 ┌ Ruminococcus bromii ATCC27255 T L76600
                Anaerotruncus colihominis DSM17241 T AJ315980
             11 Ruminococcus albus JCM14654 T AB538438
             60 Ruminococcus callidus ATCC27760 T L76596
             87 Ruminococcus flavefaciens T ATCC19208 KP689131
                Acetivibrio thermocellus JCM9322 T AB558166
                Escherichia coli JCM1649 T AB242910
```

├─┤ 0.020

**Figure S3. *F. prausnitzii* strains were isolated as candidates for gut microbiome-based intervention in MCI. Related to Figure 2.**

Phylogenetic tree showing the relationships of the 12 *F. prausnitzii* isolates with other members of *Ruminococcaceae* based on 16S rRNA gene sequences. A neighbor-joining phylogenetic tree was constructed using MEGA. Bootstrap values (expressed as a percentage of 1,000 replications) are shown at branching points. The scale bar indicates the number of substitutions per nucleotide position. *F. prausnitzii* isolates in this study are highlighted in bold. Sequence accession numbers are shown at the last part of the strain name. "T" is the abbreviation for "type strain".

30
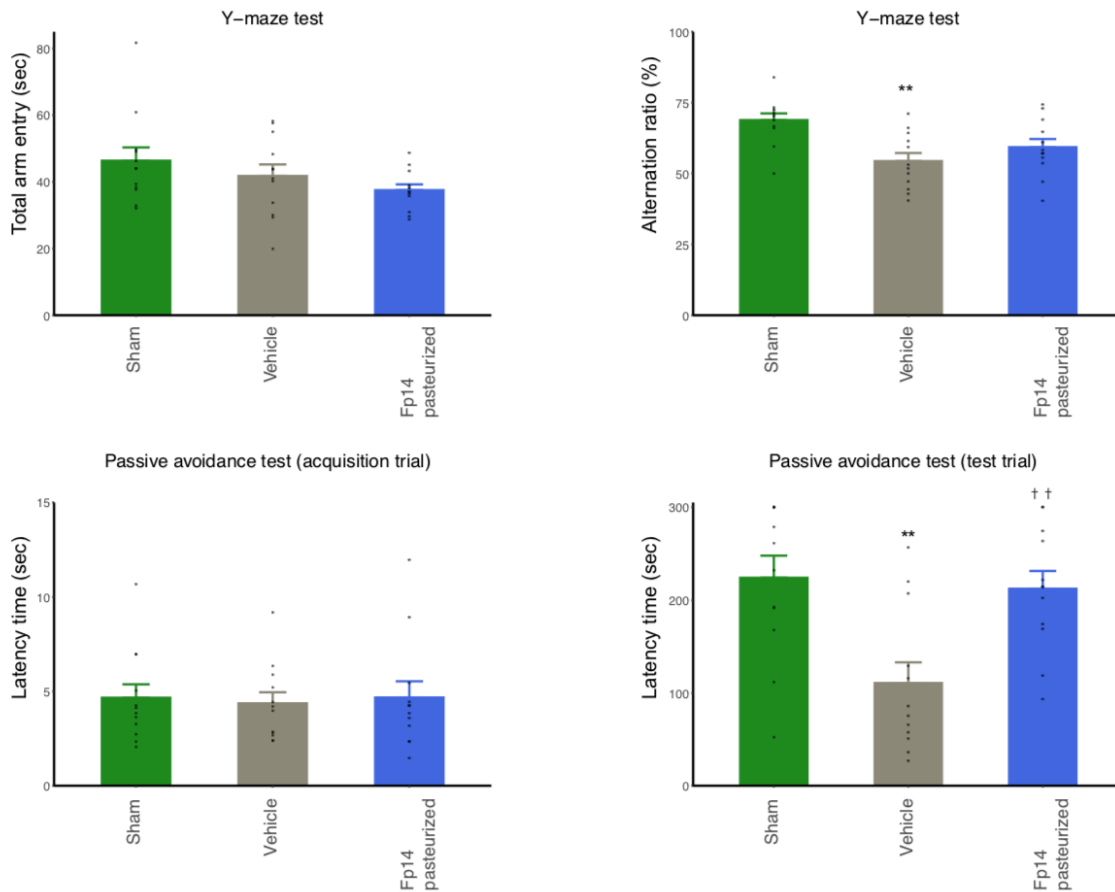
**Figure S4. Pasteurized Fp14 reproducibly improved Aβ-induced cognitive impairment. Related to Figure 4.**

Pasteurized Fp14 were orally administered to mice injected i.c.v. with Aβ25-35. Cognitive performance was evaluated by the Y-maze test and the passive avoidance test. (A) Bar plot showing the total entry time in the Y-maze test. (B) Bar plot showing the alternation ratio in the Y-maze test. (C) Bar plot showing the latency time of the acquisition trial in the passive avoidance test. (D) Bar plot showing the latency time of the test trial in the passive avoidance test. All values are expressed as the mean + S.E. (n=12, biological replicates). [**]$p <$ 0.01 by two-sided unpaired Student's $t$-test (sham-operation vs vehicle); [††]$p < 0.01$ by two-sided unpaired Student's $t$-test (vehicle vs Fp14 pasteurized).

40

41
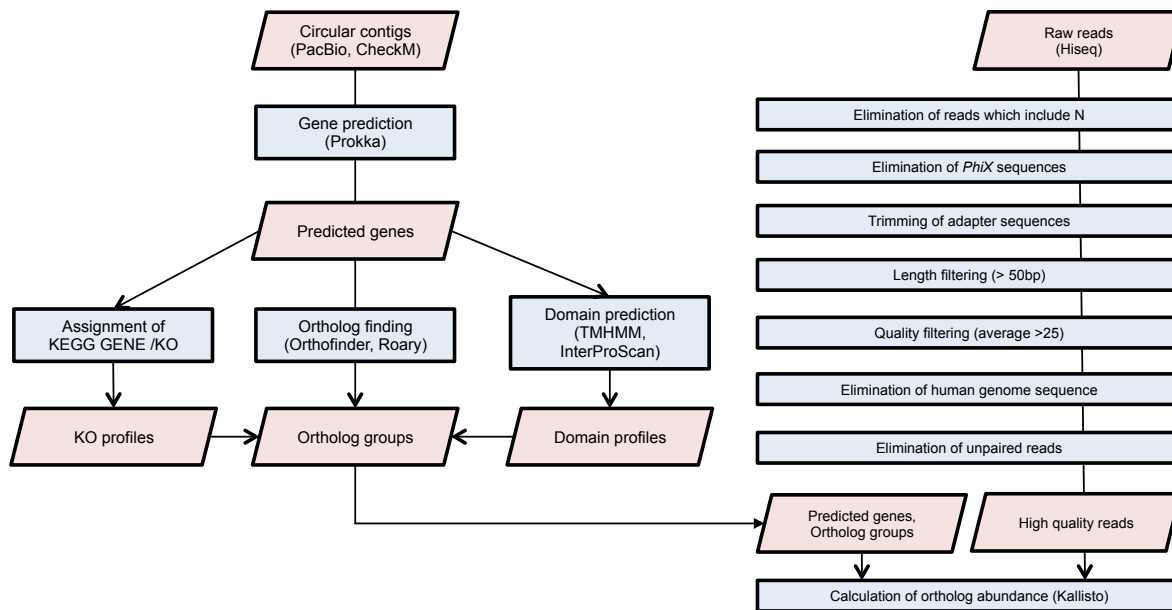
**Figure S5. Pipeline for whole-genome comparison and whole-metagenome shotgun sequencing analysis.**

43 **Related to Figure 5 and 6.**

44 Regarding the whole-genome comparison, circular contigs from Pacbio underwent a quality control check by

45 CheckM and gene prediction by Prokka. Protein domains were predicted using InterProScan, and TMHMM.

46 Genes were annotated with KEGG genes and KEGG Orthologs (KOs) using DIAMOND. Orthologs were

47 identified by using two different ortholog finding tools, Roary and Orthofinder, to enhance the precision of

48 the analysis. KOs and domain profiles were assigned to each ortholog. Regarding the whole-metagenome

49 shotgun sequencing analysis, raw reads first underwent a quality control to generate high-quality reads. The

50 abundance of each gene from the isolated strains was evaluated by Kallisto. Each gene was categorized into

51 an ortholog group, which was identified by Roary or Orthofinder, and the abundance of the orthologs was

52 calculated (see also Methods Details).

53
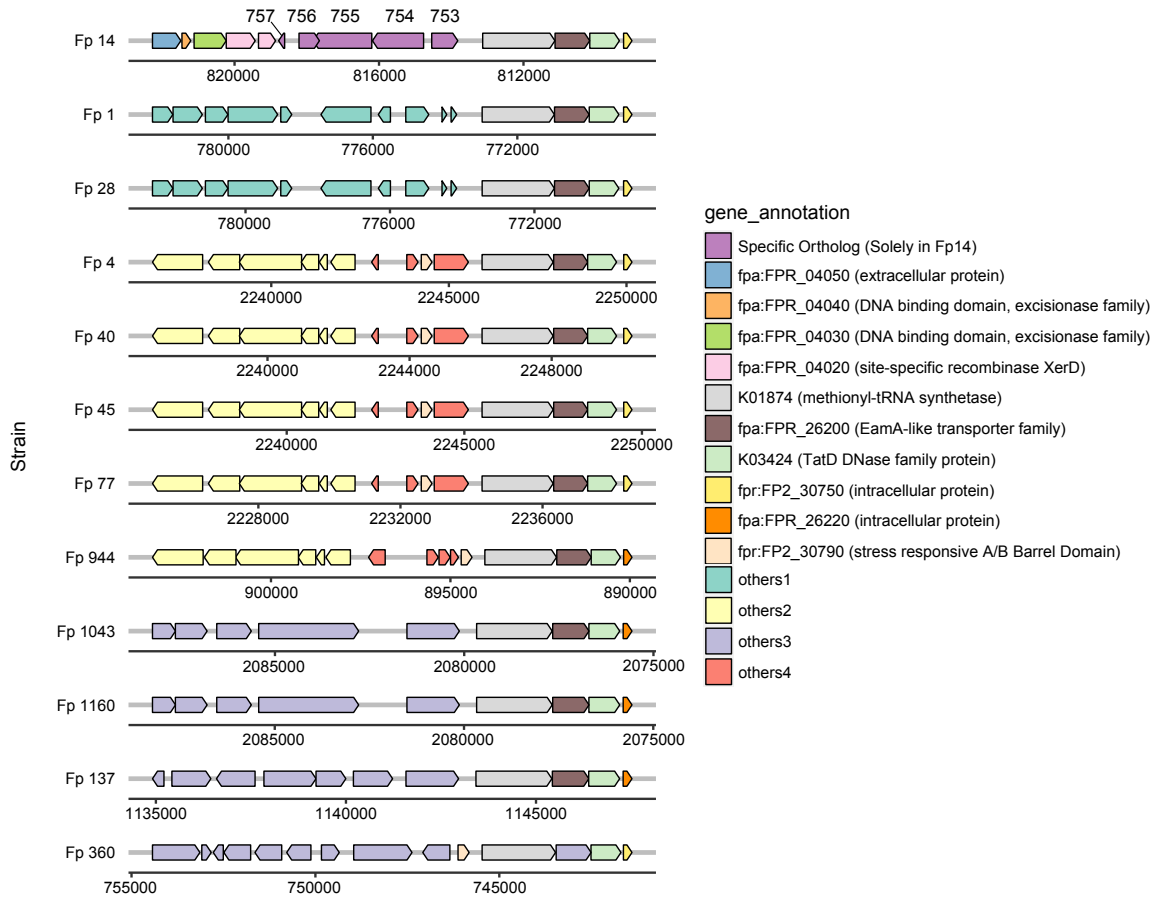
**Figure S6. Whole-genome comparison revealed specific orthologs in the Fp14. Related to Figure 5.**

Complete genomes of the 12 *F. prausnitzii* isolates were obtained using a PacBio sequencer. Specific orthologs were identified by using different ortholog finding tools, Roary and Orthofinder. Orthologs found only in specific strains or orthologs that contained KEGG Orthologs (KOs) found only in specific strains were defined as "specific orthologs". KOs and domain profiles were assigned to each ortholog using DIAMOND, InterProScan, and TMHMM. 5 of 150 specific orthologs found only in Fp14 are shown.

61 **Table S1. The characteristics of the selected subjects. Related to Figure 2 and 6.**

62 Table showing the characteristics of the selected subjects in this study. All values except gender are expressed

63 as the mean ± S.D. *P* value except for gender was calculated by Wilcoxon rank sum test (vs the healthy group).

64 *P* value for gender was calculated by Fisher's exact test (vs the healthy group).

65

66

| Number | Healthy (n = 21) | MCI (n = 15) | AD (n = 7) | *p* value (HvsM) | *p* value (HvsA) |
|---|---|---|---|---|---|
| Age (years) | 72.7±5.3 | 79.8±6.1 | 83.0±9.3 | 0.0014 | 0.0092 |
| Gender (n, % Female) | 13, 61.9 | 9, 60.0 | 5, 71.4 | 1 | 1 |
| BMI (kg/m$^2$) | 22.7±3.7 | 24.4±2.3 | 22.9±2.7 | 0.1828 | 0.8735 |
| MMSE (score) | 29.5±0.7 | 26.9±2.3 | 15.6±4.3 | 0.0001 | <0.0001 |
| MoCA-J (score) | 27.6±1.5 | 19.8±2.7 | 8.4±4.0 | <0.0001 | <0.0001 |
| Education (years) | 11.2±1.7 | 9.8±3.6 | 8.3±2.0 | 0.2146 | 0.0025 |

67 **Table S2. The completeness and the contamination of the genomes of the 12 isolated *F. prausnitzii*.**

68 **Related to Figure 5.**

69 All genomes were estimated to be ≥95% complete with ≤5% contamination by CheckM, making them

70 excellent data sets for whole-genome comparison(Parks et al., 2015).

71

| Strain | Completeness (%) | Contamination (%) | Strain heterogeneity (%) |
|--------|------------------|-------------------|--------------------------|
| Fp1 | 100 | 0 | 0 |
| Fp4 | 100 | 0 | 0 |
| Fp14 | 100 | 0 | 0 |
| Fp28 | 100 | 0 | 0 |
| Fp40 | 100 | 0 | 0 |
| Fp45 | 100 | 0 | 0 |
| Fp77 | 100 | 0 | 0 |
| Fp137 | 100 | 0 | 0 |
| Fp360 | 100 | 0 | 0 |
| Fp944 | 100 | 0 | 0 |
| Fp1043 | 100 | 0 | 0 |
| Fp1160 | 100 | 0.34 | 100 |
| Fp1233 | 100 | 0 | 0 |

72