Table 5
Clusters of white matter loss (14 left-side dominant onset MSA-P vs. 30 controls).

Voxel size 2 × 2 × 2 mm³. Clusters of gray matter SPM analysis with FDR-corrected at p < .05 are shown. The coordinates refer to the Talairach reference space.

Region volume (mm³)	Z score	Talairach coordinates (x, y, z)	Location of local maxima
13,448	5.51	32, −8, 4	Right external capsule
	4.93	18, 0, −5	Right lateral globus pallidus

Table 6
Clusters of gray matter loss (6 right-side dominant onset MSA-P vs. 30 controls).

Region volume (mm³)	Z score	Talairach coordinates (x, y, z)	Location of local maxima
11,832	5.76	−27, 6, 5	Left putamen
	3.76	−61, −5, 8	Left superior temporal gyrus
	3.70	−36, −9, 13	Left insula
11,216	4.84	12, −7, 45	Right cingulate gyrus
	4.52	2, 17, 34	Right cingulate gyrus
	4.18	−12, −8, 43	Left cingulate gyrus
5536	3.76	20, −70, −8	Right lingual gyrus
	3.62	24, −76, −11	Right fusiform gyrus
	3.55	38, −54, −26	Right cerebellum
4504	3.86	−18, −86, −11	Left fusiform gyrus
	3.50	−36, −58, −24	Left cerebellum

Voxel size 2 × 2 × 2 mm³. Clusters of gray matter SPM analysis with FDR-corrected at p < .05 are shown. The coordinates refer to the Talairach reference space.

Table 7
Clusters of white matter loss (6 right-side dominant onset MSA-P vs. 30 controls).

Region volume (mm³)	Z score	Talairach coordinates (x, y, z)	Location of local maxima
101,384	6.05	−34, −10, 2	Left external capsule
	5.68	−22, −6, −6	Left lateral globus pallidus
	4.79	22, −49, −19	Right cerebellum
4912	4.06	22, −32, 0	Right external capsule
	3.88	34, −12, 2	Right external capsule
	3.65	24, −8, −10	Right parahippocampal area
4288	4.08	30, 5, 26	Right frontal lobe
	3.97	46, 5, 20	Right frontal lobe
	3.71	28, −2, 41	Right frontal lobe

Voxel size 2 × 2 × 2 mm³. Clusters of gray matter SPM analysis with FDR-corrected at p < .05 are shown. The coordinates refer to the Talairach reference space.

than in MSA-C (Wenning et al., 1996, 1997). Atrophy of the cerebellar vermis and lobes was detected, which is in line with previous VBM studies (Chang et al., 2009; Minnerop et al., 2010; Tzarouchi et al., 2010) and pathological findings (Wenning et al., 1997). White matter cerebellar atrophy was also detected in addition to gray matter atrophy. The pons, bilateral middle cerebellar peduncles and the left cerebellar lobe were involved, a finding that was also consistent with previous VBM studies (Minnerop et al., 2010; Tzarouchi et al., 2010).

The white matter loss was detected in bilateral globus pallidi and external capsules extending to the midbrain which has never been reported before in a study using VBM. The white matter volume loss of the globus pallidus might reflect the fact that the globus pallidus has many efferent fibers and many bundles of myelinated fibers from the striatum traverse globus pallidus (Mamata et al., 2002; Nieuwenhuys et al., 2008). It is supported by the fact that the glucose metabolism of the globus pallidus is generally almost equal to white matter metabolism.

The areas connecting the globus pallidus with each structure were also significantly atrophic in this study. The neuronal loss in globus pallidus is pathologically proven (Jellinger et al., 2005; Ozawa et al., 2004; Wenning et al., 1997, 2002) and it plays an important role in parkinsonism as well as substantia nigra. Thus, those findings could result from the degeneration of the motor pathway such as striopallidal

fiber, strionigral fiber and pallidonigral fiber (Nieuwenhuys et al., 2008). On the right side, the white matter volume loss extended upward to the subcortical to precentral area. The involvement of the precentral area could be explained by the fact that the motor cortex, the supplementary motor cortex and the premotor area were also involved in the pathways mentioned above; the afferent fibers of the striatum originate mainly from the motor and premotor areas, and then go back to the area (Nieuwenhuys et al., 2008). Another possibility is that the corticospinal tract itself is affected by this disease. The recent VBM studies of Minnerop et al. (2010) revealed white matter reduction along the corticospinal tract in the bilateral internal capsules and subcortical to left precentral gyrus, findings that agree with our study. However, they did not detect white matter atrophy around the deep gray matter. We believe that this was because they evaluated MSA patients including both MSA-C and MSA-P patients, and MSA-C patients made up more than 70% of their patient groups.

This better detection of deep gray matter and white matter volume loss in our study probably resulted from the application of the SPM8 plus DARTEL algorithm (Matsuda et al., 2012). The latest version of SPM8 enabled more accurate segmentation of the MRIs into gray matter, white matter, and cerebrospinal fluid images compared to older versions of SPM. In addition, DARTEL provides improved registration accuracy compared with conventional VBM.

In this study, we also evaluated both 14 left-side dominant and 6 right-side dominant onset MSA-P patients and detected the contralateral putaminal atrophy. Our results agreed with the previous MR imaging and pathological findings (Fearnley and Lees, 1990; Kato et al., 1992; Kume et al., 1992). The reduced white matter was also predominant in the contralateral globus pallidus and external capsule extending to the level of the midbrain. To our knowledge, this is also the first report to mention the laterality of the clinical findings and of white matter volume loss using VBM. The detection of volume loss only on the right side in the corona radiate of 20 MSA-P patients might be due to the large number of left-side dominant patients.

This study has several limitations. First, the number of patients in this study was not large. This might be the main reason for the failure
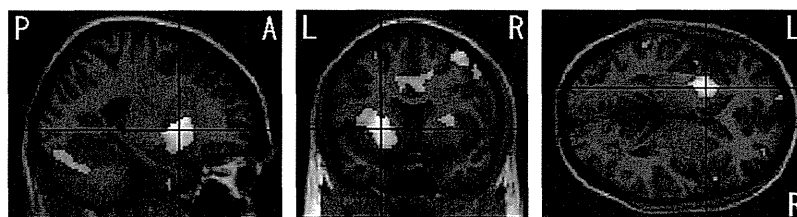

Fig. 4. Comparison of gray and white matter volumes by VBM using SPM8 plus DARTEL among 6 patients with right-side dominant onset MSA-P and 30 control subjects. Significant gray matter atrophy which is shown in a yellow color is observed in the left putamen, bilateral cerebellums and several cortical regions (FDR-corrected at p < .05). Bilateral cerebellums and several cortical regions are partly shown. Significant white matter atrophy which is shown in a red color is observed in the left globus pallidus, bilateral external capsules, right frontal lobe, parahippocampal area and cerebellum (FDR-corrected at p < .05).

of detecting the correlation with disease duration and severity. A few previous VBM studies with a small number of patients have reported inconsistent results on correlation with disease duration and severity (Brenneis et al., 2007; Minnerop et al., 2007, 2010). A further study with a large number of MSA-P patients would be necessary. Second, this study lacked the pathologic confirmation. However, it was difficult to select data only for pathologically confirmed patients, and the diagnosis in this study was clinically evaluated by an experienced neurologist based on consensus criteria. Third, we failed to detect the gray matter volume loss of the substantia nigra as well as in previous VBM studies, although the neuronal loss of substantia nigra is a hallmark of MSA-P (Mamata et al., 2002; Minnerop et al., 2010; Tir et al., 2009). Even though the software used for the analysis and the evaluation of spatial resolution have greatly improved, the segmentation of MRIs of nuclei located in brainstem remains a weakness.

5. Conclusions

In conclusion, VBM using SPM8 plus DARTEL detected significant volume loss not only in the gray but also in the white matter of the area affected by MSA-P. Significant structural atrophic change of the areas connecting the globus pallidus with each structure which plays a crucial role in parkinsonism was detected for the first time using VBM. Further prospective investigations involving a larger number of MSA-P patients combined with DTI techniques are required to confirm our findings. VBM using SPM8 plus DARTEL could also be a useful tool for evaluating other neurodegenerative diseases as well as MSA-P.

References

Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. NeuroImage 38, 95–113.

Ashburner, J., Friston, K.J., 2000. Voxel-based morphometry: the methods. NeuroImage 11, 805–821.

Bergfield, K.L., Hanson, K.D., Chen, K., Teipel, S.J., Hampel, H., Rapoport, S.I., Moeller, J.R., Alexander, G.E., 2010. Age-related networks of regional covariance in MRI gray matter: reproducible multivariate patterns in healthy aging. NeuroImage 49, 1750–1759.

Brenneis, C., Seppi, K., Schocke, M.F., Müller, J., Luginger, E., Bösch, S., Löscher, W.N., Büchel, C., Poewe, W., Wenning, G.K., 2003. Voxel-based morphometry detects cortical atrophy in the Parkinson variant of multiple system atrophy. Movement Disorders 18, 1132–1138.

Brenneis, C., Seppi, K., Schocke, M., Benke, T., Wenning, G.K., Poewe, W., 2004. Voxel based morphometry reveals a distinct pattern of frontal atrophy in progressive supranuclear palsy. Journal of Neurology, Neurosurgery, and Psychiatry 75, 246–249.

Brenneis, C., Egger, K., Scherfler, C., Seppi, K., Schocke, M., Poewe, W., Wenning, G.K., 2007. Progression of brain atrophy in multiple system atrophy. A longitudinal VBM study. Journal of Neurology 254, 191–196.

Burton, E.J., Karas, G., Paling, S.M., Barber, R., Williams, E.D., Ballard, C.G., McKeith, I.G., Scheltens, P., Barkhof, F., O'Brien, J.T., 2002. Patterns of cerebral atrophy in dementia with Lewy bodies using voxel-based morphometry. NeuroImage 17, 618–630.

Chang, C.C., Chang, Y.Y., Chang, W.N., Lee, Y.C., Wang, Y.L., Lui, C.C., Huang, C.W., Liu, W.L., 2009. Cognitive deficits in multiple system atrophy correlate with frontal atrophy and disease duration. European Journal of Neurology 16, 1144–1150.

Fearnley, J.M., Lees, A.J., 1990. Striatonigral degeneration. A clinicopathological study. Brain 113, 1823–1842.

Gilman, S., Wenning, G.K., Low, P.A., Brooks, D.J., Mathias, C.J., Trojanowski, J.Q., Wood, N.W., Colosimo, C., Dürr, A., Fowler, C.J., Kaufmann, H., Klockgether, T., Lees, A., Poewe, W., Quinn, N., Revesz, T., Robertson, D., Sandroni, P., Seppi, K., Vidailhet, M., 2008. Second consensus statement on the diagnosis of multiple system atrophy. Neurology 71, 670–676.

Jellinger, K.A., Seppi, K., Wenning, G.K., 2005. Grading of neuropathology in multiple system atrophy: proposal for a novel scale. Movement Disorders 20 (Suppl. 12), S29–S36.

Kato, T., Kume, A., Ito, K., Tadokoro, M., Takahashi, A., Sakuma, S., 1992. Asymmetrical FDG-PET and MRI findings of striatonigral system in multiple system atrophy with hemiparkinsonism. Radiation Medicine 10, 87–93.

Klockgether, T., Ludtke, R., Kramer, B., Abele, M., Bürk, K., Schöls, L., Riess, O., Laccone, F., Boesch, S., Lopes-Cendes, I., Brice, A., Inzelberg, R., Zilber, N., Dichgans, J., 1998. The natural history of degenerative ataxia: a retrospective study in 466 patients. Brain 121, 589–600.

Kume, A., Shiratori, M., Takahashi, A., Kato, T., Ito, K., Tadokoro, M., Sakuma, S., 1992. Hemi-parkinsonism in multiple system atrophy: a PET and MRI study. Journal of Neurological Sciences 110, 37–45.

Mamata, H., Mamata, Y., Westin, C.F., Shenton, M.E., Kikinis, R., Jolesz, F.A., Maier, S.E., 2002. High-resolution line scan diffusion tensor MR imaging of white matter fiber tract anatomy. AJNR. American Journal of Neuroradiology 23, 67–75.

Matsuda, H., Mizumura, S., Nemoto, K., Yamashita, F., Imabayashi, E., Sato, N., Asada, T., 2012. Automatic voxel-based morphometry of structural MRI by SPM8 plus diffeomorphic anatomic registration through exponentiated lie algebra improves the diagnosis of probable Alzheimer Disease. AJNR. American Journal of Neuroradiology 33, 1109–1114.

Messina, D., Cerasa, A., Condino, F., Arabia, G., Novellino, F., Nicoletti, G., Salsone, M., Morelli, M., Lanza, P.L., Quattrone, A., 2011. Patterns of brain atrophy in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy. Parkinsonism & Related Disorders 17, 172–176.

Minnerop, M., Specht, K., Ruhlmann, J., Schimke, N., Abele, M., Weyer, A., Wüllner, U., Klockgether, T., 2007. Voxel-based morphometry and voxel-based relaxometry in multiple system atrophy-a comparison between clinical subtypes and correlations with clinical parameters. NeuroImage 36, 1086–1095.

Minnerop, M., Lüders, E., Specht, K., Ruhlmann, J., Schimke, N., Thompson, P.M., Chou, Y.Y., Toga, A.W., Abele, M., Wüllner, U., Klockgether, T., 2010. Callosal tissue loss in multiple system atrophy-a one-year follow-up study. Movement Disorders 25, 2613–2620.

Nieuwenhuys, R., Voogd, J., Huijzen, C.V., 2008. Telencephalon: Basal Ganglia, The Human Central Nervous Systemfourth ed. Springer 427–489.

Ozawa, T., Paviour, D., Quinn, N.P., Josephs, K.A., Sangha, H., Kilford, L., Healy, D.G., Wood, N.W., Lees, A.J., Holton, J.L., Revesz, T., 2004. The spectrum of pathological involvement of the striatonigral and olivopontocerebellar systems in multiple system atrophy: clinicopathological correlations. Brain 127, 2657–2671.

Papp, M.I., Lantos, P.L., 1994. The distribution of oligodendroglial inclusions in multiple system atrophy and its relevance to clinical symptomatology. Brain 117, 235–243.

Quinn, N., 1989. Multiple system atrophy: the nature of the beast. Journal of Neurology, Neurosurgery, and Psychiatry 52, 78–89.

Schulz, J.B., Skalej, M., Wedekind, D., Luft, A.R., Abele, M., Voigt, K., Dichgans, J., Klockgether, T., 1999. Magnetic resonance imaging-based volumetry differentiates idiopathic Parkinson's syndrome from multiple system atrophy and progressive supranuclear palsy. Annals of Neurology 45, 65–74.

Tir, M., Delmaire, C., le Thuc, V., Duhamel, A., Destée, A., Pruvo, J.P., Defebvre, L., 2009. Motor-related circuit dysfunction in MSA-P: usefulness of combined whole-brain imaging analysis. Movement Disorders 24, 863–870.

Tzarouchi, L.C., Astrakas, L.G., Konitsiotis, S., Tsouli, S., Margariti, P., Zikou, A., Argyropoulou, M.I., 2010. Voxel-based morphometry and voxel-based relaxometry in parkinsonian variant of multiple system atrophy. Journal of Neuroimaging 20, 260–266.

Wenning, G.K., Tison, F., Elliott, L., Quinn, N.P., Daniel, S.E., 1996. Olivopontocerebellar pathology in multiple system atrophy. Movement Disorders 11, 157–162.

Wenning, G.K., Tison, F., Ben Shlomo, Y., Daniel, S.E., Quinn, N.P., 1997. Multiple system atrophy: a review of 203 pathologically proven cases. Movement Disorders 12, 133–147.

Wenning, G.K., Seppi, K., Tison, F., Jellinger, K., 2002. A novel grading scale for striatonigral degeneration (multiple system atrophy). Journal of Neural Transmission 109, 307–320.

Hindawi

Research Article

Differential Diagnosis Tool for Parkinsonian Syndrome Using Multiple Structural Brain Measures

Miho Ota,[1] Yasuhiro Nakata,[2] Kimiteru Ito,[2] Kouhei Kamiya,[2] Masafumi Ogawa,[3] Miho Murata,[3] Satoko Obu,[1] Hiroshi Kunugi,[1] and Noriko Sato[2]

[1] *Department of Mental Disorder Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, 4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8502, Japan*
[2] *Department of Radiology, National Center of Neurology and Psychiatry Hospital, 4-1-1, Ogawa-Higashi, Kodaira, Tokyo 187-8551, Japan*
[3] *Department of Neurology, National Center of Neurology and Psychiatry Hospital, 4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8551, Japan*

Correspondence should be addressed to Miho Ota; ota@ncnp.go.jp

Clinical differentiation of parkinsonian syndromes such as the Parkinson variant of multiple system atrophy (MSA-P) and cerebellar subtype (MSA-C) from Parkinson's disease is difficult in the early stage of the disease. To identify the correlative pattern of brain changes for differentiating parkinsonian syndromes, we applied discriminant analysis techniques by magnetic resonance imaging (MRI). T1-weighted volume data and diffusion tensor images were obtained by MRI in eighteen patients with MSA-C, 12 patients with MSA-P, 21 patients with Parkinson's disease, and 21 healthy controls. They were evaluated using voxel-based morphometry and tract-based spatial statistics, respectively. Discriminant functions derived by step wise methods resulted in correct classification rates of 0.89. When differentiating these diseases with the use of three independent variables together, the correct classification rate was the same as that obtained with step wise methods. These findings support the view that each parkinsonian syndrome has structural deviations in multiple brain areas and that a combination of structural brain measures can help to distinguish parkinsonian syndromes.

1. Introduction

Multiple system atrophy (MSA) is an adult-onset, sporadic, progressive neurodegenerative disease characterized by varying severity of parkinsonian features, and cerebellar ataxia, autonomic failure, and corticospinal disorders [1–4]. According to the clinical presentation, a parkinsonian type (MSA-P) and a cerebellar type of MSA (MSA-C) are distinguished [2]. Parkinson's disease (PD) is a progressive neurodegenerative movement disorder characterized by rigidity, tremor, and bradykinesia. Its prevalence increases with age, and it affects 1% of the population over age 65 [5].

PD and MSA are both alpha-synucleinopathies [6, 7]. Pathologically, in Parkinson's disease a massive loss of dopaminergic neurons in pars compacta of substantia nigra and intraneuronal Lewy bodies are present [8]. In MSA, neuronal

loss and gliosis occur in the inferior olives, pons, transverse pontocerebellar fibers, cerebellum, substantia nigra, locus caeruleus, striatum, and the intermediolateral column of the spinal cord [9]. In MSA-P, the nigrostriatal system is the main site of pathology, but less severe degeneration can be widespread and usually includes the olivopontocerebellar system [9, 10]. In MSA-C, the olivopontocerebellar system is mainly affected along with loss of pontine neurons and transverse pontocerebellar fibres and atrophy of the middle cerebellar peduncles (MCPs) [9, 10]. Conventional magnetic resonance imaging (MRI) may also help distinguish the two forms of MSA. MSA-P shows "slit-like" marginal hyperintensity of the putamen [11]. Additionally, the "hot-cross bun" sign on T2-weighted and proton density images in the ventral pons has been reported to be related to MSA-C [12]. However, these MRI changes do not always occur

[3]. Clinical differential diagnosis between PD and MSA is difficult in the early stage of the disease. Relevant works on the other tools like SPECT [13], transcranial brain sonography [14–16] and on optical coherence tomography [17] showed the effectiveness for differentiating PD from healthy volunteer. Additionally, combined use of 123I-(S)-2-hydroxy-3-iodo-6-methoxy-N-((1-ethyl-2-pyrrodinyl)-methyl) benzamide (IBZM), 123I-N-v-fluoropropyl-2b-carbomethoxy-3b-(4-iodophenyl)nortropan (FP-CIT), and meta-123I-iodobenzylguanidine (MIBG) distinguishes Parkinsons disease from atypical parkinsonian disorder, such as PSP and MSA with the accuracy of about 90% [18]. However, 3 SPECT/scintigraphy tests only for diagnosis are not practical.

Over the last few years, a number of MRI studies have focused on the identification of diagnostic markers helpful in the differential diagnosis of parkinsonian syndromes such as MSA, PD, and progressive supranuclear palsy (PSP) [19–22]. However, no studies have discriminated among PD, MSA-P, MSA-C, and healthy subjects simultaneously. In the present study, we hypothesized that we would be able to distinguish the PD and healthy subjects from the MSA subjects by using the infratentorial brain images and MSA-P and PD from the MSA-C and healthy subjects by using supratentorial images. The characteristic distribution of regional brain changes revealed by the gray matter volume data using the optimized, voxel-based morphometry (VBM) method and by the diffusion tensor imaging data using tract-based spatial statistics (TBSS) would have diagnostic values for discriminating such diseases.

2. Materials and Methods

2.1. Subjects. From November 2006 to November 2010, 200 consecutive patients whose chief complaints were parkinsonism underwent brain MRI at our institution. We excluded the patients with cerebrovascular diseases cortical infarctions, multiple lacunar lesions, leukoaraiosis, and other lesions above Fazekas's Grade 2 on T2-weighted images or fluid-attenuated inversion recovery (FLAIR) MRI [23], PSP, and corticobasal degeneration (CBD). Clinical diagnosis of PD and MSA was made according to the established consensus criteria [2, 24]. A probable clinical diagnosis was determined by two neurologist with more than 20 years of experience in the diagnosis of movement disorders (MO, MM). As a consequence, 18 consecutive patients with MSA-C, 12 patients with MSA-P, and 21 patients with PD were studied. Their characteristics are shown in Table 1. 24 out of 30 MSA patients were hospitalized for the detailed diagnosis, and the diagnosis of 30 MSA patients weas not changed during follow-up clinical assessments (mean period = 2.2 years). As for PD, the follow-up clinical assessments were conducted (mean period = 4.7 years) after the MRI imaging, and no additional pathology was detected. 21 age- and sex-matched healthy persons who demonstrated no current or past history of psychiatric illness or contact with psychiatric services were enrolled as controls. Participants were excluded if they had a prior medical history of central nervous system disease or severe head injury. The study protocol was approved by the

Table 1: Characteristics of the participants.

	MSA-C	MSA-P	PD	Normal volunteers
Mean age (years)	63.6 ± 7.8	61.9 ± 7.7	62.2 ± 7.0	62.3 ± 5.6
Sex (male : female)	7 : 11	6 : 6	10 : 11	11 : 10
Duration of illness (year)	3.9 ± 2.5	3.3 ± 2.6	6.8 ± 4.1	

MSA-C: cerebellar form of multiple system atrophy; MSA-P: parkinsonism forms of multiple system atrophy; PD: Parkinson's disease.

ethics committee of the National Center of Neurology and Psychiatry, Japan.

2.2. MRI Data Acquisition and Processing. MR studies were performed on a Magnetom Symphony 1.5 Tesla (Siemens, Erlangen, Germany). First, high-spatial-resolution, 3-dimensional (3D) T1-weighted images of the brain were obtained for morphometric study. The 3D T1-weighted images were scanned in the sagittal plane (TE/TR: 2.64/1580 ms; flip angle: 15°; effective slice thickness: 1.23 mm; slab thickness: 177 mm; matrix: 208 × 256; FOV: 256 × 315 mm^2; acquisitions: 1) yielding 144 contiguous slices through the head. The raw 3D T1-weighted volume data were transferred to a workstation, and structural images were analyzed using an optimized VBM technique. Data were analyzed using Statistical Parametric Mapping 5 (SPM5) software (Welcome Department of Imaging Neuroscience, London, UK) running on MATLAB 7.0 (Math Works, Natick, MA). Images were processed using an optimized VBM script. The details of this process are described elsewhere [25]. First, each individual 3D-T1 image was normalized with the optimized VBM method. Normalized segmented images were modulated by multiplication with Jacobian determinants of the spatial normalization function to encode the deformation field for each subject, as tissue density changes in normal space. Gray matter volume and cerebrospinal fluid (CSF) volume images were smoothed using a 12 mm full width at half maximum of an isotropic Gaussian kernel. Diffusion tensor imaging (DTI) was then performed in the axial plane (TE/TR: 106/11,200 ms; FOV: 240 × 240 mm^2; matrix: 96 × 96; 75 continuous transverse slices; slice thickness 2.5 mm with no interslice gap). Diffusion was measured along 64 noncollinear directions with the use of a diffusion-weighted factor b in each direction for 1000 s/mm^2, and one image was acquired without use of a diffusion gradient. Recently, a novel processing technique has been published. In this technique, instead of trying to match each and every voxel in different subjects, DTI data is projected on a common pseudoanatomical skeleton and therefore does not need smoothing [26]. TBSS is available as part of the FSL 4.1 software package [27]. The TBSS script runs the nonlinear registration, aligning all fractional anisotropy (FA) images to the FMRIB58_FA template, which is supplied with FSL. The script then takes the target and affine-aligns it into a 1 × 1 × 1 mm MNI152 space. Once this is done, each subject's FA image has the nonlinear transform to the target and then the affine transform to the MNI152 space applied, resulting in a transformation of the original FA image into

the MNI152 space. Next, TBSS creates the mean of all aligned FA images and applies thinning of the local tract structure to create a skeletonized mean FA image. In order to exclude areas of low FA and/or high intersubject variability from the statistical analysis, TBSS thresholds a mean FA skeleton with a certain FA value, typically 0.2. The resulting binary skeleton mask is a pseudoanatomical representation of the main fiber tracks and defines the set of voxels used in all subsequent processing. Finally, TBSS projects each subject's aligned FA image onto the skeleton. This results in skeletonized FA data. It is this file that feeds into the voxelwise statistics. In addition to DTI and 3D T1-weighted images, conventional axial T2-weighted images (TE/TR: 95/3500 ms; flip angle: 150°; slice thickness: 5 mm; intersection gap: 1.75 mm; matrix: 448 × 512; field of view (FOV): 210 × 240 mm^2; acquisitions: 1) and fluid attenuation inversion recovery images in the axial plane (TE/TR: 101/8800 ms; flip angle: 150°; slice thickness: 3 mm; intersection gap: 1.75 mm; matrix: 448 × 512; FOV: 210 × 240 mm^2; acquisition: 1) were acquired to exclude cerebrovascular disease or other diseases such as tumors, and hydrocephalus. On conventional MRI, no abnormal findings were detected in the brain of any subject.

2.3. Statistical Analysis. We first evaluated the differences between the patients and healthy subjects using analysis of variance (ANOVA). These tests were performed with the SPSS software ver. 11 (SPSS Japan, Tokyo, Japan). There were no significant differences in age among patients and controls, but there were statistically significant differences in duration of illness between the patients with MSA-P and with PD ($P = 0.012$) and with MSA-C and with PD ($P = 0.005$).

The discriminant function analyses were then conducted to assess the ability of a combination of brain anatomical variables to distinguish between patients with MSA-C, MSA-P, Parkinson's disease, and controls. The independent variables were the volume data and fractional anisotropy value derived from the normalized individual image using the region of interests (ROI) method. ROIs were put on the "single_subj_T1.nii" image regarded as the anatomically standard image in SPM5, in the fourth ventricle, cerebellum hemisphere; these were derived from the WFU_PickAtlas, extension program of SPM5 [28, 29]. We also put ROIs on the "FMRIB58_FA-skeleton_1 mm.nii" image, which is the anatomically standard image in FSL, in the MCP, superior cerebellar peduncle (SCP), pons, substantia nigra, superior temporal white matter region, prefrontal white matter regions, and primary motor region where previous studies showed differences among the patients with MSA-C, MSA-P, PD, and controls (Figure 1) [20, 22, 30–37]. The value of a particular tissue was extracted using the software MarsBar [38], an extension program of SPM5.

The Box's M test confirmed the inequality of the group covariance matrices (Box-M = 76.63; $P < 0.001$). Discriminant functions were derived by step wise methods based on Mahalanobis' distance. The step wise selection criteria were decided by the overall multivariate F value of each variable to test differences between the patients and controls and to maximize the discriminant function between the groups. At

TABLE 2: Mean fractional anisotropy value in each region of interests and 4th ventricular and cerebellar volumes in patients with MSA-C, MSA-P, PD, and controls.

	MSA-C	MSA-P	PD	Normal
4th Vent_vol	0.31 ± 0.07	0.28 ± 0.07	0.19 ± 0.04	0.20 ± 0.04
Cerebellum_vol	0.28 ± 0.04	0.30 ± 0.03	0.33 ± 0.04	0.34 ± 0.04
MCP	0.32 ± 0.05	0.37 ± 0.05	0.41 ± 0.03	0.39 ± 0.03
SCP	0.44 ± 0.02	0.43 ± 0.02	0.45 ± 0.04	0.47 ± 0.03
Pons	0.12 ± 0.01	0.13 ± 0.01	0.14 ± 0.01	0.13 ± 0.01
SN	0.53 ± 0.04	0.54 ± 0.01	0.54 ± 0.05	0.56 ± 0.04
ST	0.21 ± 0.01	0.13 ± 0.01	0.13 ± 0.01	0.21 ± 0.02
PF	0.05 ± 0.00	0.06 ± 0.01	0.06 ± 0.00	0.05 ± 0.01
PM	0.20 ± 0.02	0.21 ± 0.02	0.22 ± 0.02	0.21 ± 0.02

4th Vent_vol: fourth ventricle volume; cerebellum_vol: cerebellum volume; MSA-C: cerebellar form of multiple system atrophy; MSA-P: parkinsonism forms of multiple system atrophy; MCP: middle cerebellar peduncle; SCP: superior cerebellar peduncle; PD: Parkinson's disease; SN: substantia nigra; ST: superior temporal region; PF: prefrontal region; PM: primary motor region.

TABLE 3: The coefficients of discriminant analysis.

		Factor 1	Factor 2	Factor 3
Stepwise method	4th Vent_vol	1.74	19.47	3.33
	SN	19.16	−9.81	38.32
	ST	−127.87	3.14	−32.07
	PF	196.3	80.22	−188.97
	(Constant)	0.66	−4.13	−5.99
Two independent variables analysis	Pons	−76.33	*69.41*	(—)
	ST	*97.92*	13.01	(—)
	(Constant)	−6.98	11.27	(—)
Three independent variables analysis	4th Vent_vol	1.14	*15.82*	12.23
	MCP	−24.59	−5.39	*23.04*
	ST	*100.68*	−13.21	6.19
	(Constant)	−4.01	−3.68	−13.58

4th Vent_vol: fourth ventricle volume; MCP: middle cerebellar peduncle; SN: substantia nigra; ST: superior temporal region; PM: primary motor region.

the same time, we entered the two or three independent variables together and estimated the predictive power of the discriminant function.

3. Results

We first calculated the volume and FA value from the spatially normalized images using ROIs. The mean values of these parameters are summarized in Table 2 and Figure 2. The mean FA value of the prefrontal region was too small to examine, so we did not evaluate the influence of the FA value in the prefrontal white matter region. We then conducted discriminant function analyses. The following five variables were entered in a step wise manner: fourth ventricle volume, substantia nigra, superior temporal, and prefrontal white matter region. The discriminant coefficients are shown in Table 3. The use of these variables resulted in correct

TABLE 4: Classification results.

			MSA-C	MSA-P	PD	Control	Total
				Predicted group membership			
Stepwise method (88.9% of original grouped cases correctly classified)							
Original data	Count	MSA-C	15	0	0	3	18
		MSA-P	0	9	3	0	12
		PD	0	2	19	0	21
		Control	0	0	0	21	21
	%	MSA-C	83.3	0	0	16.7	100
		MSA-P	0	75.0	25.0	0	100
		PD	0	9.5	90.5	0	100
		Control	0.0	0	0	100.0	100
Two independent variables (84.7% of original grouped cases correctly classified)							
Original data	Count	MSA-C	15	0	0	3	18
		MSA-P	0	10	2	0	12
		PD	0	4	17	0	21
		Control	1	1	0	19	21
	%	MSA-C	83.3	0.0	0	16.7	100
		MSA-P	0	83.3	16.7	0	100
		PD	0	19.0	81.0	0	100
		Control	4.8	4.8	0.0	90.5	100
Three independent variables (88.9% of original grouped cases correctly classified)							
Original data	Count	MSA-C	14	0	0	4	18
		MSA-P	0	10	2	0	12
		PD	0	1	20	0	21
		Control	0	1	0	20	21
	%	MSA-C	77.8	0.0	0	22.2	100
		MSA-P	0	83.3	16.7	0	100
		PD	0	4.8	95.2	0	100
		Control	0.0	4.8	0.0	95.2	100

MSA-C: cerebellar form of multiple system atrophy; MSA-P: parkinsonism forms of multiple system atrophy; PD: Parkinson's disease.

classification rates of 0.89 (χ^2 = 294.66; df = 12; P < 0.001; Wilks' lambda = 0.012) (Table 4).

The correct classification rates of each combination used to run the discriminant function analyses using two or three independent variables together are listed in Table 5. The highest correct classification rates were measured when we estimated the FA value of the "pons and superior temporal region" and "superior temporal region, MCP and fourth ventricle volume", respectively. Table 3 shows the discriminant coefficients, and Table 4 shows the correct classification rates derived from the analyses using two or three independent variables together, in the same way. Figure 3 shows the discriminant scores of each subject, calculated by the analysis using three independent variables together.

4. Discussion

We found that the step wise discriminant function analysis identified with fairly good accuracy the combinations of ROIs that characterized brain anatomical features distinguishing the patients with MSA-C, MSA-P, PD, and healthy subjects,

and that when discriminate analysis was conducted using the fourth ventricle volume and the FA value of MCP and superior temporal region as independent variables together, the correct classification rate was the same as that of step wise discriminant function analysis.

One study showed that patients with MSA-C and MSA-P share similar diffusion tensor imaging features in the infratentorial region [22]. Furthermore, the combination of DTI metrics can be used to distinguish between patients with MSA and with PD. However, they could not differentiate the patients with PD from healthy subjects. This may be because they were focused on the infratentorial FA value and did not investigate the focal lesions related to parkinsonism. In this study, we used the FA value of the superior temporal regions, known to be impaired in PD as an independent variable, so we could discriminate the patients with PD from healthy subjects and with MSA-C from those with MSA-P [30–32, 34, 35].

One study reported discriminating patients with MSA-P, PD, PSP, and healthy subjects [20]. They indicated that investigating the degeneration of the MCP is useful for the in vivo differential diagnosis of MSA-P and PD. These

TABLE 5: The correct classification rates of each combination of independent variables.

(a) Two independent variables

	4th Vent_vol	Cerebellum_vol	MCP	SCP	Pons	SN	ST	PM
4th Vent_vol		0.514	0.583	0.611	0.583	0.556	0.819	0.431
Cerebellum_vol			0.528	0.556	0.514	0.458	0.722	0.458
MCP				0.708	0.472	0.667	0.833	0.528
SCP					0.681	0.528	0.736	0.542
Pons						0.667	0.847	0.514
SN							0.722	0.583
ST								0.653
PM								

4th Vent_vol: fourth ventricle volume; cerebellum_vol: cerebellum volume; MCP: middle cerebellar peduncle; SCP: superior cerebellar peduncle; SN: substantia nigra; ST: superior temporal region; PM: primary motor region.

(b) Three independent variables

4th Vent_vol	Cerebellum_vol	MCP	SCP	Pons	SN	ST	PM	Accuracy
4th Vent_vol	Cerebellum_vol	MCP						0.556
4th Vent_vol	Cerebellum_vol		SCP					0.681
4th Vent_vol	Cerebellum_vol			Pons				0.569
4th Vent_vol	Cerebellum_vol				SN			0.583
4th Vent_vol	Cerebellum_vol					ST		0.861
4th Vent_vol	Cerebellum_vol						PM	0.583
4th Vent_vol		MCP	SCP					0.708
4th Vent_vol		MCP		Pons				0.542
4th Vent_vol		MCP			SN			0.653
4th Vent_vol		MCP				ST		*0.889*
4th Vent_vol		MCP					PM	0.569
4th Vent_vol			SCP	Pons				0.722
4th Vent_vol			SCP		SN			0.625
4th Vent_vol			SCP			ST		0.833
4th Vent_vol			SCP				PM	0.681
4th Vent_vol				Pons	SN			0.681
4th Vent_vol				Pons		ST		0.889
4th Vent_vol				Pons			PM	0.597
4th Vent_vol					SN	ST		0.875
4th Vent_vol					SN		PM	0.611
4th Vent_vol						ST	PM	0.861
	Cerebellum_vol	MCP	SCP					0.681
	Cerebellum_vol	MCP		Pons				0.486
	Cerebellum_vol	MCP			SN			0.639
	Cerebellum_vol	MCP				ST		0.833
	Cerebellum_vol	MCP					PM	0.528
	Cerebellum_vol		SCP	Pons				0.625
	Cerebellum_vol		SCP		SN			0.583
	Cerebellum_vol		SCP			ST		0.764
	Cerebellum_vol		SCP				PM	0.583
	Cerebellum_vol			Pons	SN			0.653
	Cerebellum_vol			Pons		ST		0.833
	Cerebellum_vol			Pons			PM	0.528
	Cerebellum_vol				SN	ST		0.792
	Cerebellum_vol				SN		PM	0.583
	Cerebellum_vol					ST	PM	0.792

(b) Continued.

						Accuracy
MCP	SCP	Pons				0.681
MCP	SCP		SN			0.694
MCP	SCP			ST		0.819
MCP	SCP				PM	0.681
MCP		Pons	SN			0.694
MCP		Pons		ST		0.806
MCP		Pons			PM	0.486
MCP			SN	ST		0.833
MCP			SN		PM	0.625
MCP				ST	PM	0.819
	SCP	Pons	SN			0.667
	SCP	Pons		ST		0.819
	SCP	Pons			PM	0.681
	SCP		SN	ST		0.792
	SCP		SN		PM	0.611
	SCP			ST	PM	0.722
		Pons	SN	ST		0.819
		Pons	SN		PM	0.694
		Pons		ST	PM	0.819
			SN	ST	PM	0.764

(a) Middle cerebellar peduncle   (b) Pons   (c) Superior cerebellar peduncle   (d) Substantia nigra

(e) Superior temporal region   (f) Prefrontal region   (g) Primary motor region

(h) 4th ventricle   (i) Cerebellar hemisphere
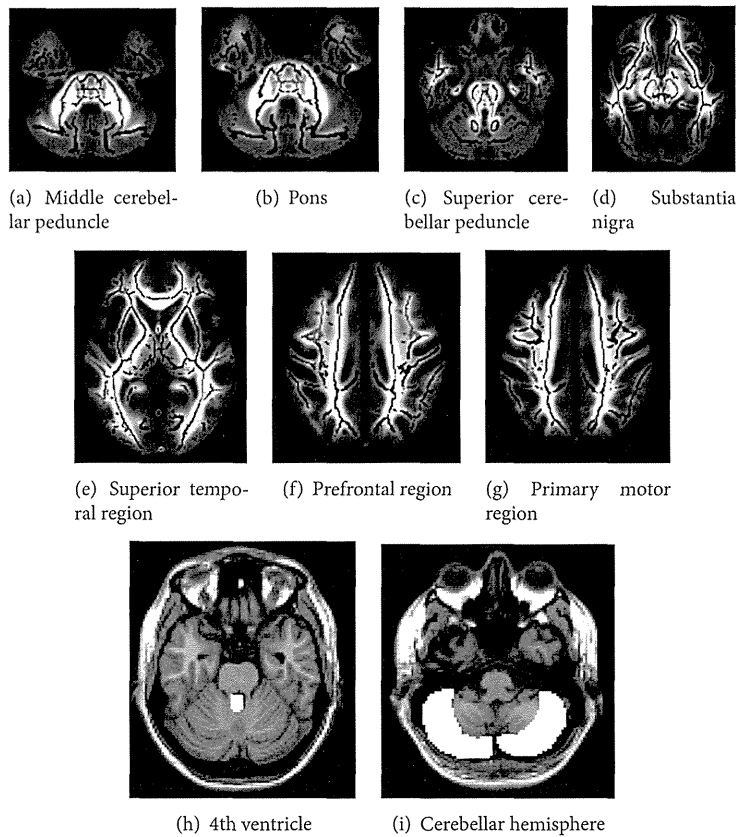
Figure 1: Locations of regions of interest. From (a) to (g) were put on the "FMRIB58_FA-skeleton_1 mm.nii" image, the anatomically standard image in FSL. Background fractional anisotropy image was the "MNI152_T1_1 mm.nii," which was also the standard image in FSL. (h) and (i) were put on a "single_subj_T1.nii" image regarded as the anatomically standard image in SPM5.
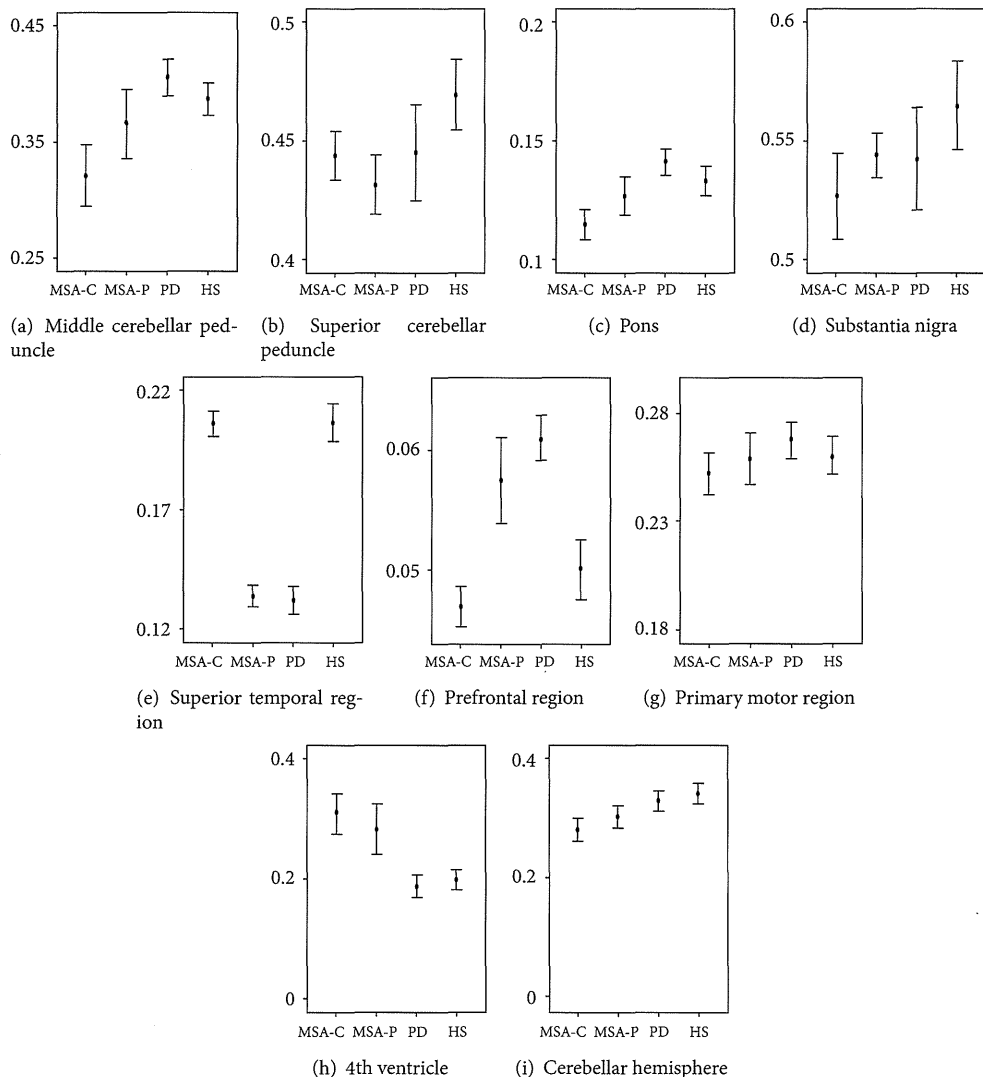
FIGURE 2: A graphic presentation of mean fractional anisotropy values in each region and mean volumes of 4th ventricular and cerebellum. MSA-C: multiple system atrophy with predominant cerebellar ataxia; MSA-P: multiple system atrophy with predominant parkinsonism; PD: Parkinson's disease; HS: healthy subjects.

results are congruent with our study. Establishing a means of differentiation using MR imaging would have potential therapeutic implications.

In this study, the participants with PD had a statistically longer duration of illness than those with MSA. It is known that MRI studies with PD show slight or no gray matter atrophy in early- to moderate-stage patients, whereas later-stage patients exhibited marked cortical atrophy [31]. We used the FA values for the independent variables to differentiate the patients with PD from others. White matter which appears normal on conventional MRI can show FA abnormalities, possibly permitting an earlier identification of the disease process which involves white matter tracts of the brain [33]. In addition, we successfully differentiated the patients with MSA-P from those with MSA-C using the same parameter.

There is significant value to distinguishing these diseases using DTI metrics.

All voxel-based analysis methods are susceptible to the effects of the spatial normalization transformation that registers images of different individuals. Regions in which this spatial transformation has relatively lower accuracy will tend to display artificially higher variability, which will adversely affect statistical significance. To date, TBSS is considered more robust and better suited for whole brain DTI data analysis. However, there are some limitations with the TBSS analysis. First, "FMRIB58_FA-skeleton_1 mm.nii" did not cover the thalamus and striatum, both of which have significant relationships with parkinsonian features [39]. A previous study demonstrated signal changes of the MR image after levodopa administration in an anatomical cluster which included the
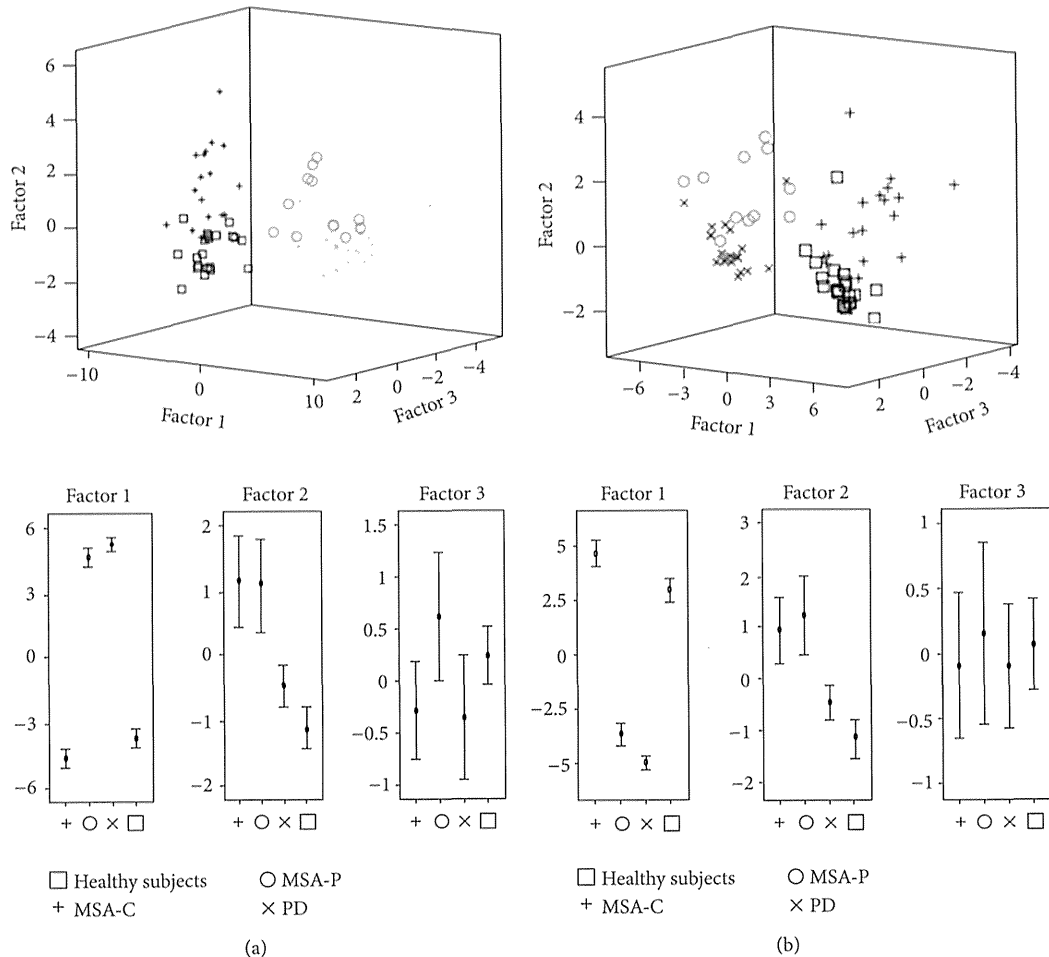
FIGURE 3: 3-dimensional scattered plots showed the discriminant scores of each subject. These were calculated (a) by the analysis using step wise method and (b) using three independent variables together. Factors 1–3 were defined in Table 3.

substantia nigra, tegmental ventral area subthalamic nucleus bilaterally, the principal origin, and first relay nuclei of projections in brain dopaminergic systems [40]. Therefore, we do not recommend using dopamine-rich regions such as the thalamus and striatum in discriminant function analysis for parkinsonian syndrome. Second, the predominant motor feature can change with time. The designation of MSA-P and MSA-C refers to the predominant feature at the time the patient is evaluated, and the predominant feature can change with time [2]. All of our MSA-C samples did not change the diagnosis to the MSA-P during follow-up clinical assessments; however, the discriminant method in this study would be fitted for the initial diagnosis. Third, our study is a small cross-sectional study, and we did not validate this discrimination method using another independent sample. MSA-P and MSA-C without cerebrovascular findings were so scarce, and we did not gather sufficient sample size. Further work with a large sample is required for the development of better discriminant capability and, if feasible, with data on another parkinsonism, PSP, would bring further clinical advantage.

5. Conclusions

Discriminant functions derived by step wise methods resulted in correct classification rates of 0.89. The present methods for automated analysis of morphometric data largely support findings from earlier studies using expert-guided ROIs or automated procedures. These findings support the view that each parkinsonian syndrome has structural deviations in multiple brain areas, and discriminant function analysis in this paper may provide objective biological information adjunct to the clinical diagnosis of parkinsonian syndromes.

References

[1] F. Geser, G. K. Wenning, K. Seppi et al., "Progression of multiple system atrophy (MSA): a prospective natural history study by the European MSA Study Group (EMSA SG)," *Movement Disorders*, vol. 21, no. 2, pp. 179–186, 2006.

[2] S. Gilman, G. K. Wenning, P. A. Low et al., "Second consensus statement on the diagnosis of multiple system atrophy," *Neurology*, vol. 71, no. 9, pp. 670–676, 2008.

[3] N. Quinn, "Multiple system atrophy—the nature of the beast," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 52, pp. 78–89, 1986.

[4] G. K. Wenning, F. Tison, Y. B. Shlomo, S. E. Daniel, and N. P. Quinn, "Multiple system atrophy: a review of 203 pathologically proven cases," *Movement Disorders*, vol. 12, no. 2, pp. 133–147, 1997.

[5] D. Aarsland, K. Andersen, J. P. Larsen, A. Lolk, H. Nielsen, and P. Kragh-Sørensen, "Risk of dementia in Parkinson's disease: a community-based, prospective study," *Neurology*, vol. 56, no. 6, pp. 730–736, 2001.

[6] M. G. Spillantini, R. A. Crowther, R. Jakes, N. J. Cairns, P. L. Lantos, and M. Goedert, "Filamentous α-synuclein inclusions link multiple system atrophy with Parkinson's disease and dementia with Lewy bodies," *Neuroscience Letters*, vol. 251, no. 3, pp. 205–208, 1998.

[7] K. Wakabayashi, M. Yoshimoto, S. Tsuji, and H. Takahashi, "α-synuclein immunoreactivity in glial cytoplasmic inclusions in multiple system atrophy," *Neuroscience Letters*, vol. 249, no. 2-3, pp. 180–182, 1998.

[8] J. M. Fearnley and A. J. Lees, "Ageing and Parkinson's disease: substantia nigra regional selectivity," *Brain*, vol. 114, part 5, pp. 2283–2301, 1991.

[9] G. K. Wenning, F. Tison, L. Elliott, N. P. Quinn, and S. E. Daniel, "Olivopontocerebellar pathology in multiple system atrophy," *Movement Disorders*, vol. 11, no. 2, pp. 157–162, 1996.

[10] A. Kume, A. Takahashi, and Y. Hashizume, "Neuronal cell loss of the striatonigral system in multiple system atrophy," *Journal of the Neurological Sciences*, vol. 117, no. 1-2, pp. 33–40, 1993.

[11] K. Bhattacharya, D. Saadia, B. Eisenkraft et al., "Brain magnetic resonance imaging in multiple-system atrophy and Parkinson disease: a diagnostic algorithm," *Archives of Neurology*, vol. 59, no. 5, pp. 835–842, 2002.

[12] M. Savoiardo, L. Strada, F. Girotti et al., "Olivopontocerebellar atrophy: MR diagnosis and relationship to multisystem atrophy," *Radiology*, vol. 174, no. 3, pp. 693–696, 1990.

[13] J. Spiegel, M. O. Möllers, W. H. Jost et al., "FP-CIT and MIBG scintigraphy in early Parkinson's disease," *Movement Disorders*, vol. 20, no. 5, pp. 552–561, 2005.

[14] G. Becker, J. Seufert, U. Bogdahn, H. Reichmann, and K. Reiners, "Degeneration of substantia nigra in chronic Parkinson's disease visualized by transcranial color-coded real-time sonography," *Neurology*, vol. 45, no. 1, pp. 182–184, 1995.

[15] K. Busse, R. Heilmann, S. Kleinschmidt et al., "Value of combined midbrain sonography, olfactory and motor function assessment in the differential diagnosis of early Parkinson's disease," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 83, no. 4, pp. 441–447, 2012.

[16] M. Izawa Okawa and H. Miwa, "Transcranial sonography findings in Parkinson's disease," *Brain Nerve*, vol. 64, no. 4, pp. 413–422, 2012.

[17] P. Albrecht, A. K. Müller, M. Südmeyer et al., "Optical coherence tomography in parkinsonian syndromes," *PLoS ONE*, vol. 7, Article ID e34891, 2012.

[18] M. Südmeyer, C. Antke, T. Zizek et al., "Diagnostic accuracy of combined FP-CIT, IBZM, and MIBG scintigraphy in the differential diagnosis of degenerative parkinsonism: a multidimensional statistical approach," *Journal of Nuclear Medicine*, vol. 52, no. 5, pp. 733–740, 2011.

[19] R. L. Gama, D. F. G. Távora, R. C. Bomfim, C. E. Silva, V. M. de Bruin, and P. F. de Bruin, "Morphometry MRI in the differential diagnosis of parkinsonian syndromes," *Arquivos de Neuro-Psiquiatria*, vol. 68, no. 3, pp. 333–338, 2010.

[20] G. Nicoletti, R. Lodi, F. Condino et al., "Apparent diffusion coefficient measurements of the middle cerebellar peduncle differentiate the Parkinson variant of MSA from Parkinson's disease and progressive supranuclear palsy," *Brain*, vol. 129, no. 10, pp. 2679–2687, 2006.

[21] A. Quattrone, G. Nicoletti, D. Messina et al., "MR imaging index for differentiation of progressive supranuclear palsy from Parkinson disease and the Parkinson variant of multiple system atrophy," *Radiology*, vol. 246, no. 1, pp. 214–221, 2008.

[22] M. F. H. Schocke, K. Seppi, R. Esterhammer et al., "Diffusion-weighted MRI differentiates the Parkinson variant of multiple system atrophy from PD," *Neurology*, vol. 58, no. 4, pp. 575–580, 2002.

[23] F. Fazekas, J. B. Chawluk, and A. Alavi, "MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging," *American Journal of Roentgenology*, vol. 149, no. 2, pp. 351–356, 1987.

[24] W. R. G. Gibb and A. J. Lees, "The relevance of the Lewy body to the pathogenesis of idiopathic Parkinson's disease," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 51, pp. 745–752, 1998.

[25] C. D. Good, I. Johnsrude, J. Ashburner, R. N. A. Henson, K. J. Friston, and R. S. J. Frackowiak, "Cerebral asymmetry and the effects of sex and handedness on brain structure: a voxel-based morphometric analysis of 465 normal adult human brains," *NeuroImage*, vol. 14, no. 3, pp. 685–700, 2001.

[26] S. M. Smith, M. Jenkinson, H. Johansen-Berg et al., "Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data," *NeuroImage*, vol. 31, no. 4, pp. 1487–1505, 2006.

[27] S. M. Smith, M. Jenkinson, M. W. Woolrich et al., "Advances in functional and structural MR image analysis and implementation as FSL," *NeuroImage*, vol. 23, no. 1, pp. S208–S219, 2004.

[28] J. A. Maldjian, P. J. Laurienti, R. A. Kraft, and J. H. Burdette, "An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets," *NeuroImage*, vol. 19, no. 3, pp. 1233–1239, 2003.

[29] J. A. Maldjian, P. J. Laurienti, and J. H. Burdette, "Precentral gyrus discrepancy in electronic versions of the Talairach atlas," *NeuroImage*, vol. 21, no. 1, pp. 450–455, 2004.

[30] M. K. Beyer, C. C. Janvin, J. P. Larsen, and D. Aarsland, "A magnetic resonance imaging study of patients with Parkinson's disease with mild cognitive impairment and dementia using voxel-based morphometry," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 78, no. 3, pp. 254–259, 2007.

[31] P. Borghammer, K. Østergaard, P. Cumming et al., "A deformation-based morphometry study of patients with early-stage Parkinson's disease," *European Journal of Neurology*, vol. 17, no. 2, pp. 314–320, 2010.

[32] A. Gerhard, N. Pavese, G. Hotton et al., "In vivo imaging of microglial activation with [11C](R)-PK11195 PET in idiopathic Parkinson's disease," *Neurobiology of Disease*, vol. 21, no. 2, pp. 404–412, 2006.

[33] N. Prakash, N. Hageman, X. Hua, A. W. Toga, S. L. Perlman, and N. Salamon, "Patterns of fractional anisotropy changes in white matter of cerebellar peduncles distinguish spinocerebellar ataxia-1 from multiple system atrophy and other ataxia syndromes," *NeuroImage*, vol. 47, supplement 2, pp. T72–T81, 2009.

[34] B. Ramírez-Ruiz, M. J. Martí, E. Tolosa et al., "Longitudinal evaluation of cerebral morphological changes in Parkinson's disease with and without dementia," *Journal of Neurology*, vol. 252, no. 11, pp. 1345–1352, 2005.

[35] K. Specht, M. Minnerop, M. Abele, J. Reul, U. Wüllner, and T. Klockgether, "In vivo voxel-based morphometry in multiple system atrophy of the cerebellar type," *Archives of Neurology*, vol. 60, no. 10, pp. 1431–1435, 2003.

[36] K. Specht, M. Minnerop, J. Müller-Hübenthal, and T. Klockgether, "Voxel-based analysis of multiple-system atrophy of cerebellar type: complementary results by combining voxel-based morphometry and voxel-based relaxometry," *NeuroImage*, vol. 25, no. 1, pp. 287–293, 2005.

[37] L. C. Tzarouchi, L. G. Astrakas, S. Konitsiotis et al., "Voxel-based morphometry and voxel-based relaxometry in parkinsonian variant of multiple system atrophy," *Journal of Neuroimaging*, vol. 20, no. 3, pp. 260–266, 2010.

[38] M. Brett, J.-L. Anton, R. Valabregue, and J.-B. Poline, "Region of interest analysis using an SPM toolbox," in *Proceedings of the 8th International Conference on Functional Mapping of the Human Brain*, Sendai, Japan, June 2002, Available on CD-ROM in *NeuroImage*, vol. 16, no. 2, 2002.

[39] P. Remy, M. Doder, A. Lees, N. Turjanski, and D. Brooks, "Depression in Parkinson's disease: loss of dopamine and noradrenaline innervation in the limbic system," *Brain*, vol. 128, no. 6, pp. 1314–1322, 2005.

[40] P. Salgado-Pineda, P. Delaveau, C. Falcon, and O. Blin, "Brain T1 intensity changes after levodopa administration in healthy subjects: a voxel-based morphometry study," *British Journal of Clinical Pharmacology*, vol. 62, no. 5, pp. 546–551, 2006.

Robustness of Gut Microbiota of Healthy Adults in Response to Probiotic Intervention Revealed by High-Throughput Pyrosequencing

Seok-Won Kim[1], Wataru Suda[1], Sangwan Kim[1], Kenshiro Oshima[1], Shinji Fukuda[2,3,4], Hiroshi Ohno[3,4], Hidetoshi Morita[5], and Masahira Hattori[1,*]

Center for Omics and Bioinformatics, The Department of Computational Biology, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwanoha 5-1-5, Kashiwa, Chiba 277-8561, Japan[1]; Institute for Advanced Biosciences, Keio University, Mizukami 246-2, Kakuganji, Tsuruoka City, Yamagata 997-0052, Japan[2]; Laboratory for Epithelial Immunobiology, RIKEN Research Center for Allergy and Immunology, Yokohama, Japan[3]; Graduate School of Nanobioscience, Yokohama City University, Yokohama 230-0045, Japan[4] and School of Veterinary Medicine, Azabu University, Fuchinobe 1-17-71, Chuo-ku, Sagamihara, Kanagawa 252-5201, Japan[5]

*To whom correspondence should be addressed. Tel. +81 4-7136-4070. Fax. +81 4-7136-4084. Email: hattori@k.u-tokyo.ac.jp

Abstract

Probiotics are live microorganisms that potentially confer beneficial outcomes to host by modulating gut microbiota in the intestine. The aim of this study was to comprehensively investigate effects of probiotics on human intestinal microbiota using 454 pyrosequencing of bacterial 16S ribosomal RNA genes with an improved quantitative accuracy for evaluation of the bacterial composition. We obtained 158 faecal samples from 18 healthy adult Japanese who were subjected to intervention with 6 commercially available probiotics containing either *Bifidobacterium* or *Lactobacillus* strains. We then analysed and compared bacterial composition of the faecal samples collected before, during, and after probiotic intervention by Operational taxonomic units (OTUs) and UniFrac distances. The results showed no significant changes in the overall structure of gut microbiota in the samples with and without probiotic administration regardless of groups and types of the probiotics used. We noticed that 32 OTUs (2.7% of all analysed OTUs) assigned to the indigenous species showed a significant increase or decrease of ≥ 10-fold or a quantity difference in >150 reads on probiotic administration. Such OTUs were found to be individual specific and tend to be unevenly distributed in the subjects. These data, thus, suggest robustness of the gut microbiota composition in healthy adults on probiotic administration.

Key words: probiotics; gut microbiota; 16S ribosomal RNA gene; pyrosequencing

1. Introduction

Probiotics are defined as live bacterial strains conferring various benefits to the consumer by modulating the intestinal ecosystem, thereby potentially promoting host health and improving host disease risk.[1–11] Various probiotic strains have been industrially developed and marketed as a variety of products and applications such as fermented foods and supplements, including yogurt[12–15] Most probiotics taxonomically belong to two genera, *Bifidobacterium* and *Lactobacillus*, that originate from various environments, including the human intestine, and both species are generally regarded as safe.[16–18]

The interaction between administrated probiotics and indigenous microbiota is one of the most attractive and important research areas, particularly because gut microbiota have been shown to be profoundly

associated with various host physiology states, including disease, diet, and age through the shift of bacterial composition, as well as metabolic and nutritional processes.[19−23] The ability of probiotics to survive through the intestine and to modulate gut microbiota is a critical factor in determining their potential for health-related outcomes.

There have been a large number of probiotic intervention studies to assess the impact of probiotics on gut microbiota in healthy adults,[24−34] infants, and children,[35,36] and in clinical trials on patients with a variety of diseases.[37,38] Most probiotic intervention studies were carried out by comparison between probiotic-treated groups and placebo controls and examined only one or two samples from periods before and during intervention or post-intervention for each subject. These experimental designs make evaluation of results obscure from a statistical viewpoint due to the high inter-individual variability of gut microbiota.[4] In addition, most of the analyses focussed on the composition of specific bacterial species or groups by conventional methods such as culturing, quantitative polymerase chain reaction (qPCR), fluorescence in situ hybridization, denaturing gradient gel electrophoresis, or terminal-restriction fragment length polymorphism based on the bacterial 16S ribosomal RNA gene (16S). These conventional methodologies may also overlook subtle changes in bacterial community structure and change of species other than targeted species. Thus, the effect of probiotic administration on the overall structure of gut microbiota is largely unknown.

Recently, a high-throughput sequencing-based analysis has been conducted for gut microbiota fed with a probiotic yogurt that provided new insights into probiotics research by utilizing a large-scale dataset.[39] However, much more data are required to understand the impact of probiotics on gut microbiota. Recent advances in sequencing technology have enabled us to elucidate complex bacterial communities, including human gut microbiota.[40,41] Particularly, 454 pyrosequencing of bacterial 16S gene tags coupled with bioinformatics provides a high-throughput and cost-effective approach for the comprehensive analysis of bacterial communities at the species level.[42−48]

In this study, we developed an analysis pipeline for bacterial communities based on barcoded 454 pyrosequencing of 16S gene tags using modified PCR primers that improved the quantitative accuracy of inferred species composition in human gut microbiota. Using this pipeline, we analysed faecal samples longitudinally collected from individuals with and without probiotic administration to evaluate the effect of probiotics on gut microbiota with respect to species richness and diversity. The results revealed the robustness and stability of gut microbiota of healthy adults in response to probiotic administration.

2. Materials and methods

2.1. Subjects, faecal sample collection, and probiotic intervention

Eighteen healthy volunteers (age: 22 ± 3.16 yrs, 6 male, 12 female) were recruited through Azabu University, Kanagawa, Japan (Supplementary Table S1). All subjects were informed of the purpose of this study. This study was approved by the ethical committee of Azabu University, and written consent was obtained from all subjects. No subjects were treated with antibiotics during faecal sample collection. The subjects were divided into six groups (three subjects per group), and each group consumed six different commercially available probiotics supplied from Yakult Honsha Co., Ltd, Kagome Co., Ltd, Morinaga Milk Industry Co., Ltd, Takanashi Milk Products Co., Ltd, Meiji Co., Ltd, and Danone Japan Co., Ltd, respectively (Supplementary Table S1). The number of each bacterial strain contained in the probiotic products was estimated as the genome equivalent by qPCR of 16S ribosomal RNA genes using 27Fmod-338R, followed by pyrosequencing of the 16S amplicons (see below). The genome equivalent per gram or millilitre and the total genome equivalent of each bacterial strain in one probiotic product are summarized in Supplementary Table S1. Three subjects in each group consumed the same probiotics daily for 8 weeks according to the schedule of sampling and probiotic intervention (Supplementary Fig. S1). Faecal samples from 4 weeks before (S00) and 8 weeks during probiotic intervention (S01–S04), and 8 weeks after cessation of probiotic intervention (S05–S08), were collected every 2 weeks from each subject. In total, we collected 158 faecal samples from the 18 subjects because we could not collect 1 sample each from 4 of the subjects.

2.2. Recovery of bacteria from faecal samples

Freshly collected faeces (1.0 g) were suspended in 20% glycerol (Wako Pure Chemical Industries, Ltd) and phosphate buffered saline solution (Life Technologies Japan, Ltd, Tokyo, Japan), frozen in liquid nitrogen, and stored at −80°C until ready for use. Bacterial pellets were prepared from frozen faecal samples as described previously.[49]

2.3. DNA isolation from bacteria

Faecal DNA was isolated and purified according to the literature, with minor modifications.[49] The bacterial pellet was suspended and incubated with 15 mg/ml lysozyme (Sigma-Aldrich Co., LCC) at

37°C for 1 h in TE10. Purified achromopeptidase (Wako Pure Chemical Industries, Ltd) was added at a final concentration of 2000 units/ml and then incubated at 37°C for 30 min. The suspension was treated with 1% (wt/vol) sodium dodecyl sulphate and 1 mg/ml proteinase K (Merck Japan) and incubated at 55°C for 1 h. The lysate was treated with phenol/chloroform/isoamyl alcohol (Life Technologies Japan, Ltd). DNA was precipitated by adding ethanol and pelleted by centrifugation at 3,300 *g* at 4°C for 15 min. The DNA pellet was rinsed with 75% ethanol, dried, and dissolved in 10 mM Tris-HCl/1 mM EDTA (TE). DNA samples were purified by treating with 1 mg/ml RNase A (Wako Pure Chemical Industries, Ltd) at 37°C for 30 min and precipitated by adding equal volumes of 20% polyethylene glycol solution (PEG6000-2.5M NaCl). DNA was pelleted by centrifugation at 8,060 *g* at 4°C, rinsed with 75% ethanol, and dissolved in TE.

2.4. 454 barcoded pyrosequencing of 16S rRNA gene

The V1–V2 region of the 16S rRNA gene was amplified using forward primer (5′-CCATCTCATCCCTGCG TGTCTCCGACTCAGNNNNNNNNNNagrgtttgatymtggc tcag-3′) containing the 454 primer A, a unique 10-bp barcode sequence for each sample (indicated in N), and 27Fmod (5′-agrgtttgatymtggctcag) in which the third base A in the original primer 27F was changed to R, and reverse primer (5′-CCTATCCCCTGTGTGC CTTGGCAGTCTCAGtgctgcctcccgtaggagt-3′) containing the 454 primer B and reverse primer 338R (5′-tgctgcctcccgtaggagt). PCR was performed in 1 × Ex Taq PCR buffer (50 μl), deoxynucleoside triphosphate (2.5 mM), Ex Taq polymerase (Takara Bio, Inc., Shiga), each primer (10 μM), and 40 ng of extracted DNA under conditions of 2 min at 96°C, 20 cycles of 96°C for 30 s, 55°C for 45 s, and 72°C for 1 min, and a final extension of 72°C for 10 min on a 9700 PCR system (Life Technologies Japan, Ltd, Tokyo, Japan). PCR products of approximately 370 bp were confirmed by agarose gel electrophoresis, purified by AMPure XP magnetic purification beads (Beckman Coulter, Inc., Brea, CA, USA), and quantified using the Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies Japan, Ltd, Tokyo, Japan). Mixed samples were prepared by pooling approximately equal amounts of PCR amplicons from each sample and subjected to 454 GS FLX Titanium or 454 GS JUNIOR (Roche Applied Science) sequencing according to the manufacturer's instructions.

2.5. Analysis pipeline for 454 barcoded pyrosequencing of 16S PCR amplicons

We developed an analysis pipeline for 454 barcoded pyrosequencing of PCR amplicons of the V1-2 region amplified by 27Fmod-338R primers. First, 16S reads were assigned to each sample based on the barcode sequence information. Second, 16S reads that did not have PCR primer sequences at both sequence termini and those with an average quality value < 25 were filtered out. Third, 16S reads containing possible chimaeric sequences that had BLAST match lengths of < 90% with reference sequences in the database were removed. Reads removed in these processes accounted for about 35% of all reads, most of which represented reads lacking PCR primer sequences (Supplementary Table S2). Finally, filter-passed reads were obtained for further analysis by trimming off both primer sequences.

All 3000 filter-passed reads of the 16S V1-2 sequences obtained from each subject were deposited in DDBJ/GenBank/EMBL with accession numbers DRA000869–DRA000886.

2.6. Assessment of the quantitative accuracy of 16S data using artificial bacterial communities

Two artificial bacterial communities (designated 'mock01' and 'mock02') were constructed by mixing genomic DNA from 10 and 11 different human gut-associated bacterial strains with an appropriate ratio, respectively (Supplementary Table S3). Genome sequences of these microbes were completely sequenced and are publicly available. From these communities, we amplified the V1-2 region by PCR using 27F-338R and 27Fmod-338R primers, the V5-6 region by 787F-1061R primers, and the V1-9 region by 27F-1492R primers. V1-2 and V5-6 amplicons were subjected to 454 pyrosequencing, and V1-9 amplicons were cloned in *Escherichia coli*, and 3000 clones were sequenced by the Sanger method, and the products were analysed with the ABI3730xl (Life Technologies Japan, Ltd, Tokyo). We also performed duplicate qPCR experiments targeting a specific genomic region of the bacterial strains in the two mock communities. All filter-passed 16S *de novo* sequences and qPCR data were then analysed by principal component analysis (PCA) to compare and assess the quantitative accuracy (Fig. 1).

The error rate of the filter-passed sequences using 27Fmod-338R primers obtained from the two mock communities was estimated by aligning the 16S V1-2 *de novo* sequences with the reference 16S sequences in the two mock communities (Supplementary Table S4).

2.7. Data analysis

2.7.1. Database Two databases were constructed for the analysis of 16S sequences. One is the 16S rRNA gene sequence database constructed
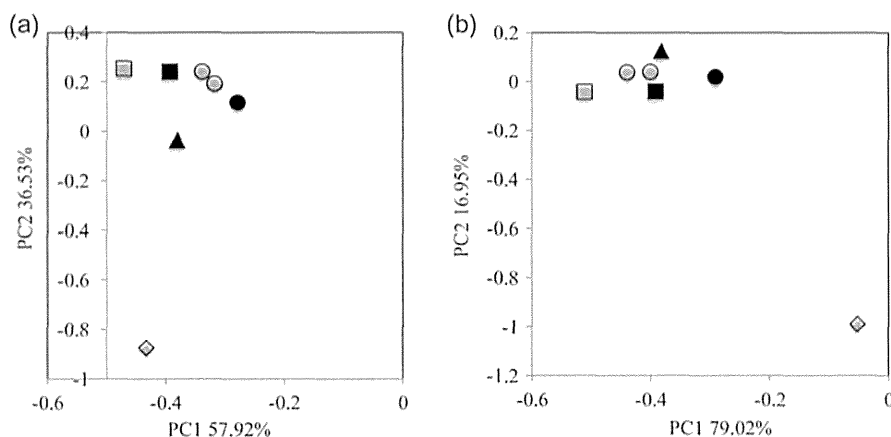
Figure 1. Assessment of the quantitative accuracy of the analysis of the bacterial composition of two mock communities by various methods. PCA analysis of the data was obtained from various methods using mock01 (a) and mock02 (b). Closed circle: expected, open circle: duplicate qPCR, closed square: pyrosequencing of 16S V1-2 region using 27Fmod, open square: pyrosequencing of 16S V1-2 region using 27F, closed triangle: pyrosequencing of 16S V5-6 region, open diamond: Sanger sequencing of nearly full-length 16S clone.

by collecting 16S sequences of ≥1200 bp of bacterial isolates from the Ribosomal Database Project v. 10.27. Another database is the reference genome database constructed by collecting genome sequences from the NCBI FTP site (ftp://ftp.ncbi.nih.gov/genbank/, Dec 2011) that includes 1482 complete and 605 draft bacterial genomes.

2.7.2. Operational taxonomic unit (OTU) and UniFrac distance analysis
We used 3000 filter-passed reads of 16S sequences for operational taxonomic unit (OTU) and UniFrac distance analysis for each sample. For OTU analysis, clustering of 16S reads was done using a 96% pairwise-identity cutoff with the UCLUST program (www.drive5.com). Representative sequences for each OTU were assigned to bacterial species by BLAST search with a 96% pairwise-identity cutoff against the two databases mentioned above. UniFrac distance analysis was used to determine the dissimilarity (distance) between two communities based on the fraction of branch length shared between two communities within a phylogenetic tree constructed from 16S sequence datasets.[44]

2.7.3. Other
Estimation of OTU numbers by extrapolation (Chao1 and ACE) was calculated with the vegan package (v2.0-5) for R (v2.15.2).

3. Results and discussion

3.1. Quantitative accuracy of 16S data produced by 454 pyrosequencing
Pyrosequencing of PCR amplicons of bacterial 16S short variable regions is the most popular and a

high-throughput approach to infer and characterize the species composition in bacterial communities.[42,45,46,48] The 454 pyrosequencing platform, which can produce over 400 bases per read, is also superior to shorter read-length sequencers with respect to sequence accuracy for single-end sequencing.[50,51] However, this PCR-based method has a problem particularly in quantification of the composition of the genus *Bifidobacterium*, a dominant species in human gut microbiota because the 16S sequence of *Bifidobacterium* has a few base mismatches with the commonly used PCR primer 27F (or 8F), underestimating this genus in the community.[52-55] To improve this, we modified primer 27F to 27Fmod by changing the third base G to R (G or A) in 27F-YM[53] that perfectly matched with the annealing site of the *Bifidobacterium* 16S gene (see Materials and methods).

To assess the 16S data using 27Fmod, we compared various 16S sequence and qPCR data obtained from two mock communities (Supplementary Table S3) that are useful to evaluate the quantitative accuracy of 16S-based data and the sequencing error rate.[56,57] Quantitative accuracy of the overall bacterial composition was evaluated by comparing the similarity of each data to the expected ('Expected') using PCA (Fig. 1). From the PCA data, Euclidean distance was calculated for evaluation of the similarity of each data with the 'Expected'. The results revealed that the order of their similarities with the 'Expected' was the qPCR data ≥ the V1-2 data using 27Fmod > the V5-6 data > the V1-2 data using 27F >> the data of Sanger sequencing-based full-length V1-9, indicating that the use of 27Fmod greatly improved the quantitative accuracy for evaluation of the overall

bacterial composition (Supplementary Table S5). This improvement was largely dependent on the improved estimation of the *Bifidobacterium* content by the use of 27Fmod. The average relative *Bifidobacterium* content in the two mock communities estimated from the data of V1-2 using 27F was only 1.5% of the 'Expected' (100%), whereas the use of 27Fmod increased the relative content to 61% that was also better than that estimated from the data of V5-6 and Sanger full-length analyses (Supplementary Fig. S2). Because qPCR can be used only when genomes of all bacteria in a given community are known, or only for a limited number of specific known species, we concluded that 454 pyrosequencing of the V1-2 region using 27Fmod-338R provided more quantitatively accurate data for bacterial composition in human gut microbiota than that using the conventional 27F primer.

We estimated the average error rate of filter-passed V1-2 data using 27Fmod-338R by aligning the V1-2 and reference 16S sequences of bacterial strains used in the two mock communities. The error rate was estimated to be 0.58 and 0.66% for mock01 and mock02 by local alignment, respectively (Supplementary Table S4). These error rates are similar to the previously published data,[43,45,50] but lower than in another study.[58] The latter may be due to differences in the examined alignment length and between local and global alignments. Errors in 454 pyrosequencing data can be the primary cause for overestimation of the OTU number that is an issue which needs to be improved for accurate estimation of species richness in bacterial communities.[59,60] We compared OTU numbers generated from clustering of various qualities of 16S reads with a 96% and a 97% pair-wise identity cutoff. For this comparison, we made and used three datasets: only primer check-passed reads having the highest error rates, filter-passed reads, and selected filter-passed reads having the lowest error rates. The results indicated that a 96% cutoff clustering of error-rich reads and a 97% cutoff clustering of filter-passed reads gave the worse results than a 96% cutoff clustering of filter-passed and selected filter-passed reads (Supplementary Fig. S3). A 97% cutoff was defined for clustering of highly accurate Sanger full-length 16S sequences.[61] Therefore, in clustering of pyrosequencing data having higher error rate than Sanger data, the use of a cutoff identity lower than 97% and a lower number of reads are reasonable to reduce overestimation of the OTU number. A 96% cutoff clustering of filter-passed reads gave similar OTU numbers up to 30−50 reads to those of filter-passed reads having the lowest error rates. These read numbers are approximately three to five times the number of input strains. After several trials

testing the mock communities, we decided to use 3000−5000 reads per sample for clustering with a 96% cutoff for the analysis of human gut microbiota. Indeed, OTU numbers using a 96% cutoff clustering of 3000 reads decreased about 15% when compared with those using a 97% cutoff clustering.

3.2. Species richness and diversity in human faecal microbiota with probiotic intervention

We randomly selected 3000 reads of 16S V1-2 sequences from all filter-passed reads for each sample (Supplementary Table S2) and used 474 000 reads in total from 158 faecal DNA samples of 18 subjects for the analysis of species richness and composition in human gut microbiota. Clustering of all reads with a 96% pairwise-identity cutoff gave a total of 2758 OTUs. After removing the minority OTUs having <0.1% abundance in any samples, 1175 OTUs having ≥0.1% abundance in at least 1 sample, accounting for 99.1% of all 16S reads, were used for further analysis.

3.2.1. Detection of administrated probiotic strains in faecal samples
We investigated whether administrated strains contained in the probiotic products can be detected in faecal DNA. We sequenced the 16S V1-2 region of all bacterial strains contained in probiotic products used in this study. The BLAST search to the databases indicated that except for the *Bifidobacterium longum* strain used in Group III, the 16S sequences of all strains in the probiotic products significantly differed from those of the indigenous species phylogenetically closest to the probiotic strains. The 16S sequence of the *B. longum* strain used in Group III was almost identical to that of an indigenous *Bifidobacterium* species, so that we used a distinguishable additive *Lactococcus lactis* strain in this product for the detection of administrated bacteria in Group III samples. The 16S sequences of these probiotic strains were included in the databases constructed in this study, and the 16S reads assigned to administrated strains had the average similarity between 99.4 and 99.9% identities with the reference sequences (data not shown). The 16S reads assigned to probiotic or additive strains were detected in samples (S01−S04) during probiotic intervention [designated 'Pro(+)'] at various frequencies, but almost none were detected in samples (S00 and S05−S08) without probiotic administration [designated 'Pro(−)'] (Supplementary Table S6). The administrated probiotic strains were shown to be more frequently detected in samples during the intervention than in the pre- and post-intervention periods using different detection methods such as culturing, targeted PCR, and hybridization.[24,26−28,30,32,33,62] In

the present study, two probiotic *Lactobacillus* and one additive *Lactococcus* strains were detected in post-intervention samples in three subjects with a minimum count, respectively. The similarity of three 16S sequences was 99.4, 99.7, and 100% identity with those of administrated *Lactobacillus* and *Lactococcus* strains, indicating that these are administrated strains. The survival of some probiotic strains in the post-intervention period was also reported previously.[28,30] Our data suggested that some probiotic strains seem to be able to persistently colonize the intestine and their survivability may be related to metabolic activity in the intestine.[63,64] Probiotic *Bifidobacterium* strains were not detected in any Pro(−) samples. However, we found two distinct 16S sequences both assigned to *Bifidobacterium animalis* in two subjects APr37 and APr39. One showed a high similarity of >98% identity with the 16S sequence of the administrated *B. animalis* and was detected with high frequency only in the Pro(+) samples, whereas another showed a low similarity of 96.5−97.4% identity (a mean of 97.2%) with low frequency in both the Pro(−) and Pro(+) samples. These data suggest the presence of unknown indigenous species phylogenetically close to, but distinct from, probiotic *B. animalis* in human gut microbiota. The total number of bacteria contained in each probiotic product was varied between 10^9 and 10^{10}, showing no large difference in quantity among them (Supplementary Table S1). No clear correlation was

also observed between the number of bacteria in the products and the frequency in detection of the administrated strains in the Pro(+) samples. From these observations, the frequency of administrated bacteria detected in faeces may not be largely affected by their amounts in the products. Therefore, detection of *Lactobacillus brevis* and *Lactobacillus delbrueck* at relatively low level in faeces cannot be simply explained by the difference in a dose, but could be considered the association with several factors such as their survivability in the intestine, diet, or physiological conditions of subjects.

3.2.2. Change of species richness in samples with and without probiotics

We analysed species richness (OTU number) in the Pro(+) and Pro(−) samples. Supplementary Figure S4 shows the change in OTU numbers for every sample in each subject, indicating that OTU numbers vary dramatically for every sample. Most of the variation can be attributed to single OTUs representing the minority species. We averaged the OTU numbers of the Pro(−) and Pro(+) samples and compared them for subject, group, type of probotics (*Lactobacillus* and *Bifidobacterium*), and all combined samples, respectively (Fig. 2). The average OTU numbers in 6 out of 18 subjects were decreased in the range of the ratio of 0.83−0.95 in the Pro(+) samples when compared with the Pro(−) samples, whereas those in other 12 subjects were increased in the range of the ratio of
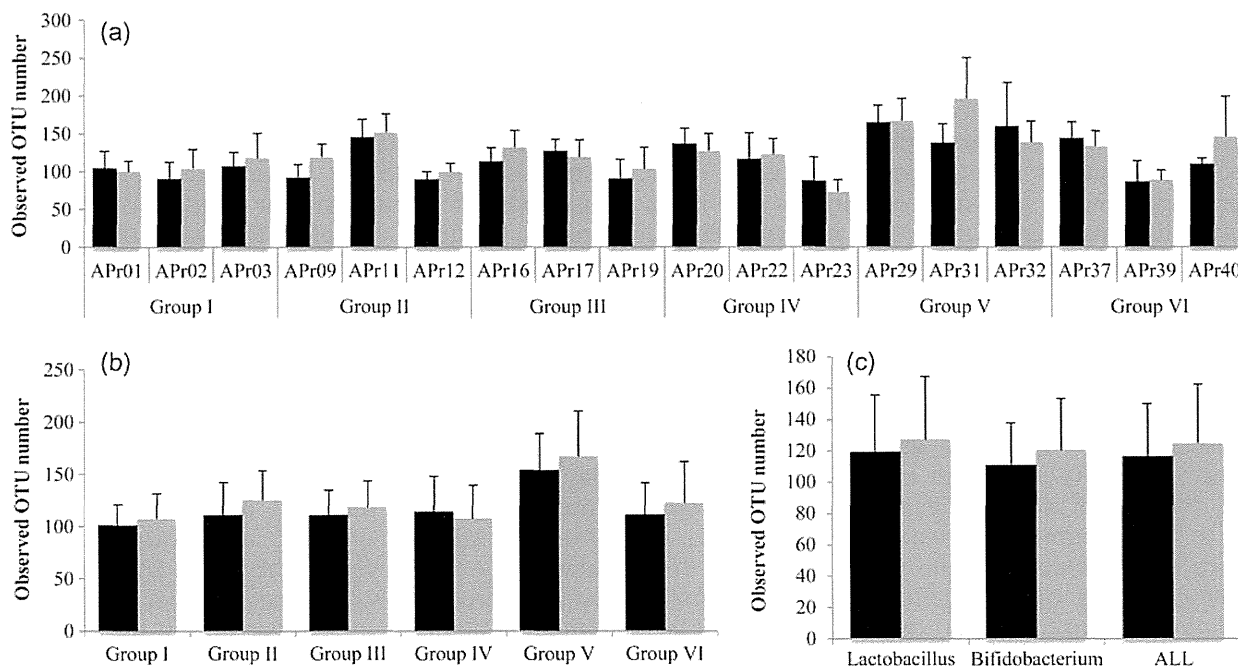

Figure 2. Change in OTU number in faecal microbiota with and without probiotic administration. (a) Individual, (b) group, (c) type of probiotics. Black bar indicates Pro(−) samples. Grey bar indicates Pro(+) samples. The error bars represent standard deviation.

1.01–1.43. For group, only Group IV showed a decrease in the average OTU number in the Pro(+) samples with the ratio of 0.94. For type of probiotics and all samples, the average OTU numbers in the Pro(+) samples were slightly more abundant (approximately 1.07-fold) than those in the Pro(−) samples, but no statistical significance was observed in any dataset. The increase in OTU number in the Pro(+) samples was largely due to the minority species (Supplementary Fig. S4), whereas the abundance of the majority species (OTUs containing ≥10 reads) was almost constant over time. We performed the same analysis using different sets of 3000 reads for each subject. The analysis reproducibly showed the similar pattern and the degree of the change in OTU numbers to which the minority species is largely attributed (data not shown). These data indicate that administration of probiotics tends to increase species richness in faecal microbiota that may

be beneficial for the consumer because the species richness in faecal microbiota of subjects afflicted with disease such as inflammatory bowel disease is significantly reduced when compared with that of healthy subjects.[65]

3.2.3. Change of species composition in samples with and without probiotics

We obtained the average weighted and unweighted UniFrac distances within Pro(−), within Pro(+), and between Pro(−) and Pro(+) samples for every group, probiotic types, and all subjects, respectively (Fig. 3). High UniFrac distance implies high variability of microbiota structure within and between samples. If the difference between any pair of the three distances is statistically significant, it can be considered that probiotic administration significantly affected the overall microbiota composition. We found the largest difference between weighted UniFac distances of the Pro(+)
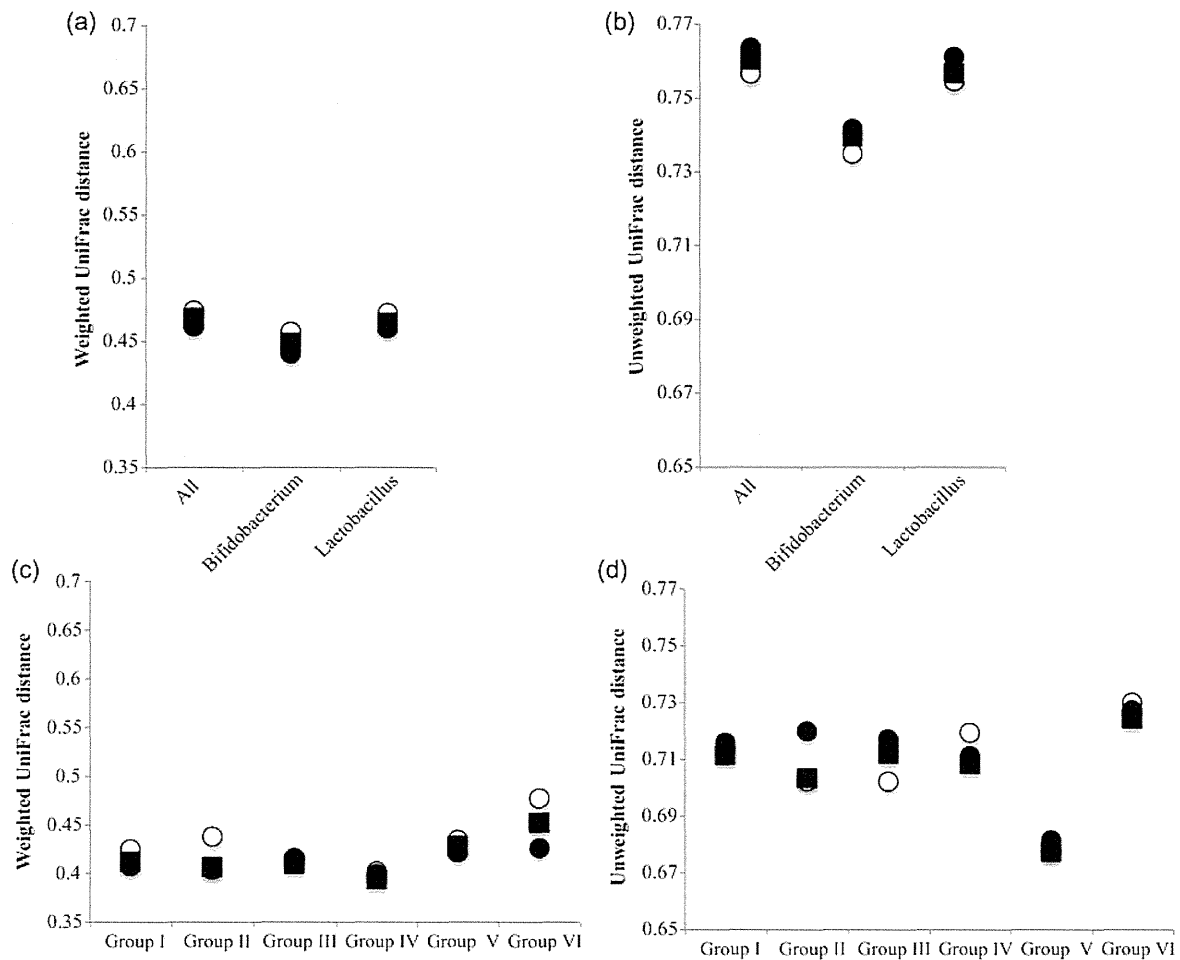

Figure 3. Average UniFrac distance within Pro (−) and Pro(+) and between Pro(−) and Pro(+) for each group, type of probitotics, and all subjects. Average UniFrac distance between any pair of the three distances for type of probiotics and all subjects (a and b), and each group (c and d). Open circle, closed circle, and closed square indicate average UniFrac distance within Pro (−), within Pro (+), and between Pro(−) and Pro(+) samples, respectively.

and Pro(−) samples in Group VI. However, statistical evaluation of this difference by the Student's t-test showed no significance (P-value > 0.05) for 781 out of 1000 times (Supplementary Table S7). These data imply high stability of gut microbiota to probiotic administration for all subjects examined. We also analysed UniFrac distances of intra-subject gut microbiota (Fig. 4). Although 5 subjects (APr02, 12, 16, 37, and 39) showed a significant difference in the UniFrac distances between Pro(−) and Pro(+) samples, the results showed that both weighted and unweighted distances between Pro(−) and Pro(+) of all intra-subjects were significantly lower than the average distance of the 18 unrelated subjects. The Welch's t-test for these differences showed statistical significance (Supplementary Table S8). We also performed the UniFrac distance analysis using different 16S datasets of 5000 reads for group, type of probiotics, all subjects, and intra-subject. The results similarly showed no statistical significance in differences between any pair of the 3 UniFrac distances and the significantly lower UniFrac distance of each intra-subject than that of the 18 unrelated subjects (data not shown). Thus, these data suggested that the perturbation of microbiota elicited by probiotics in an intra-subject did not overcome the inter-subject variations of gut microbiota, supporting high intra-specificity and stability of gut microbiota.[66,67] This robustness of gut microbiota of adults is in contrast with the profound effect of antibiotic administration on adult gut microbiota[68] and the observed response of gut microbiota of infants fed with probiotics, in which the infant gut microbiota composition was considerably affected by probiotics.[36] A short-term dietary intervention study showed that in controlled feeding of the same diet to subjects over 10 days, a marked change was observed within 1 day after the intervention initiation.[69] In the present study, no significant difference was observed between samples before (S00) and first samples (S01) after the intervention initiation (data not shown). It would be valuable to analyse faecal samples collected within a few days after administration of probiotics for evaluation of the short-term effect of probiotics.

4. Identification of bacterial species showing significant increase or decrease by probiotic administration

Although our results suggested that administration of probiotics had almost no effect on the overall structure of gut microbiota, it is possible to identify bacterial species largely responding to the administrated probiotics at the OTU/species level. We surveyed OTUs showing an increase or a decrease between the Pro(+) and Pro(−) samples by comparing the number of 16S reads for each OTU. We first enumerated the OTUs showing ≥2-fold change between the
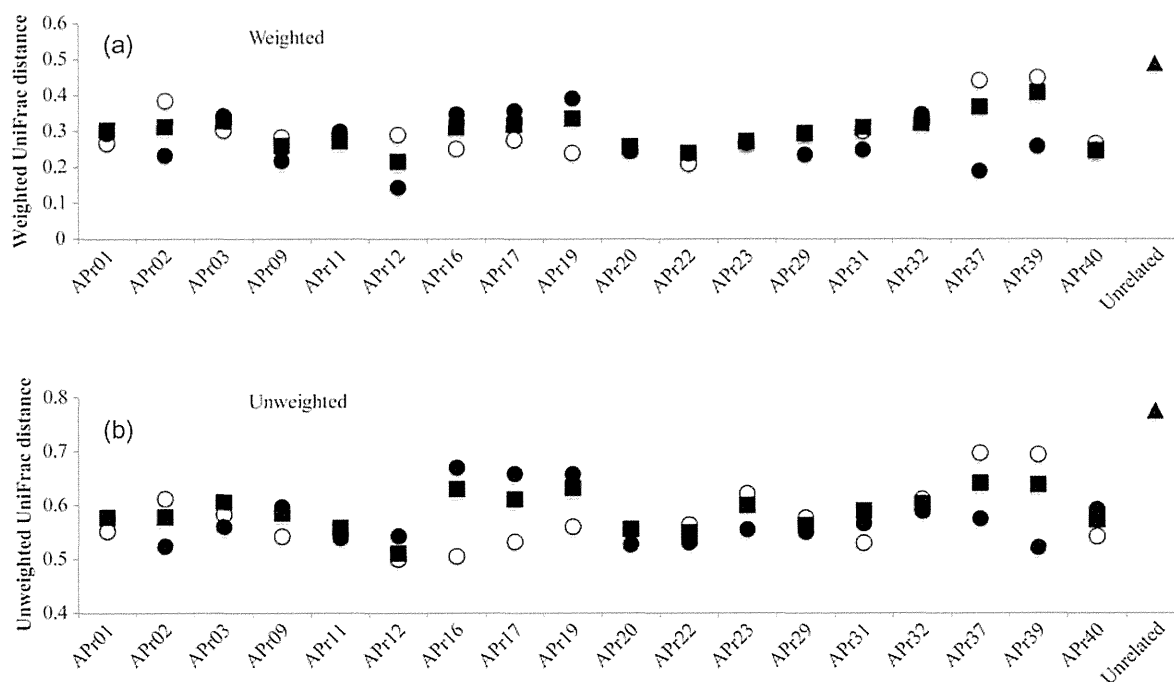

Figure 4. Average UniFrac distance within Pro(−) and Pro(+) and between Pro(−) and Pro(+) for each subject. Open circles, closed circles, and closed squares indicate average UniFrac distance within Pro(−), within Pro(+), and between Pro(−) and Pro(+) samples, respectively. Closed triangles indicate average UniFrac distance between samples (S00) of 18 unrelated individuals.