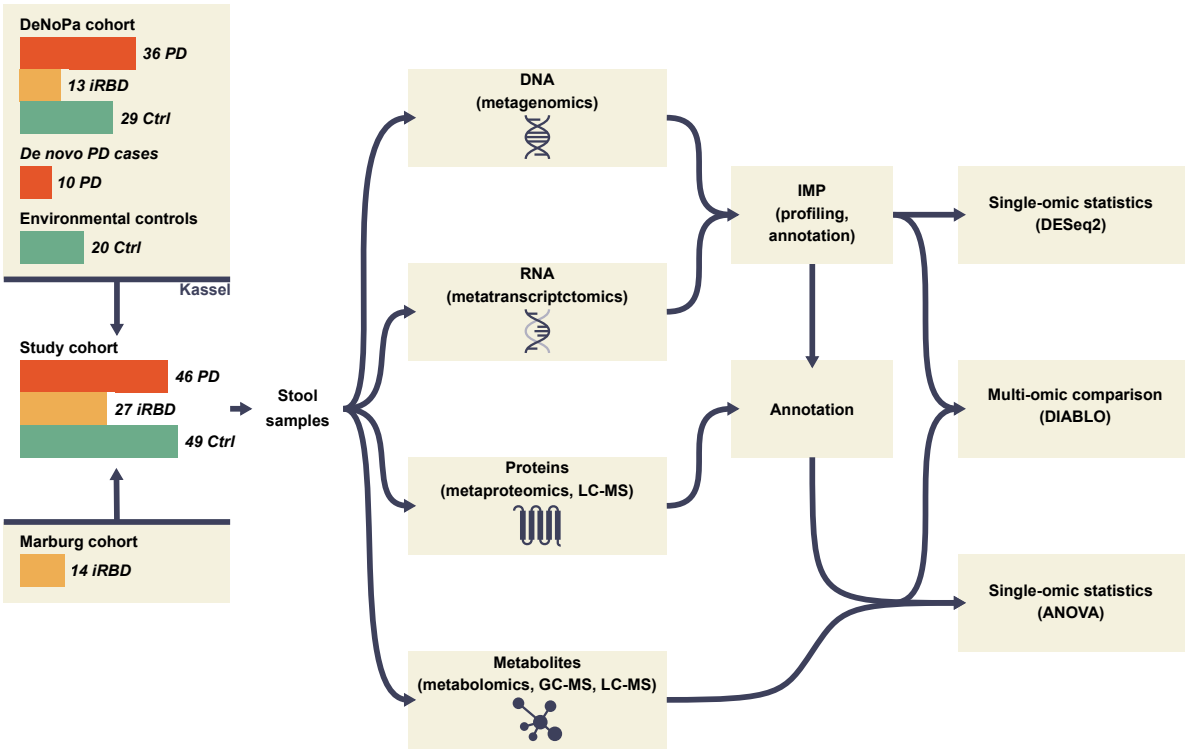
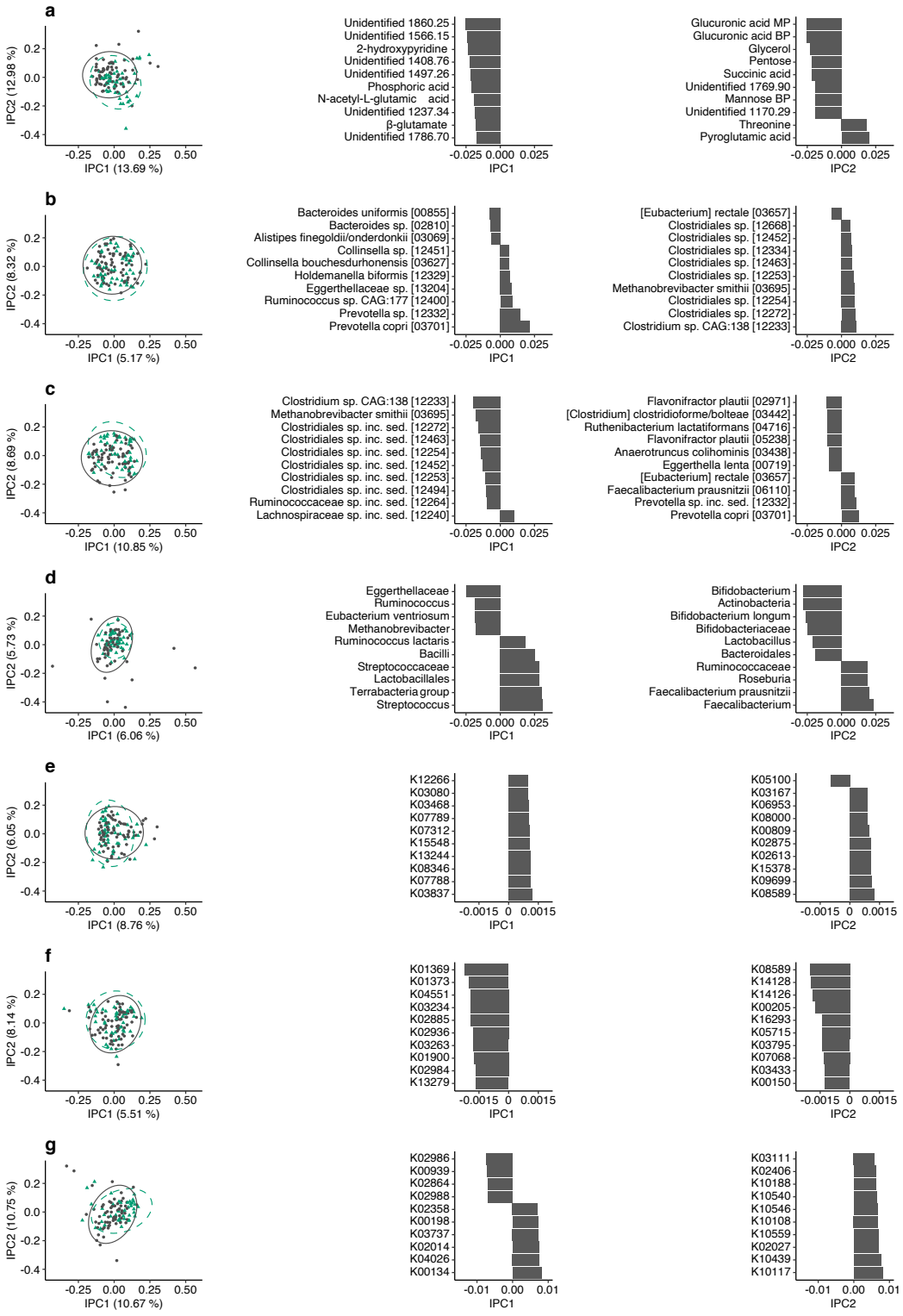


Extended Data Figure 1 | Multi-omic data analysis workflow. PD = Parkinson's disease, iRBD = idiopathic REM sleep behaviour disorder, Ctrl = control subject.

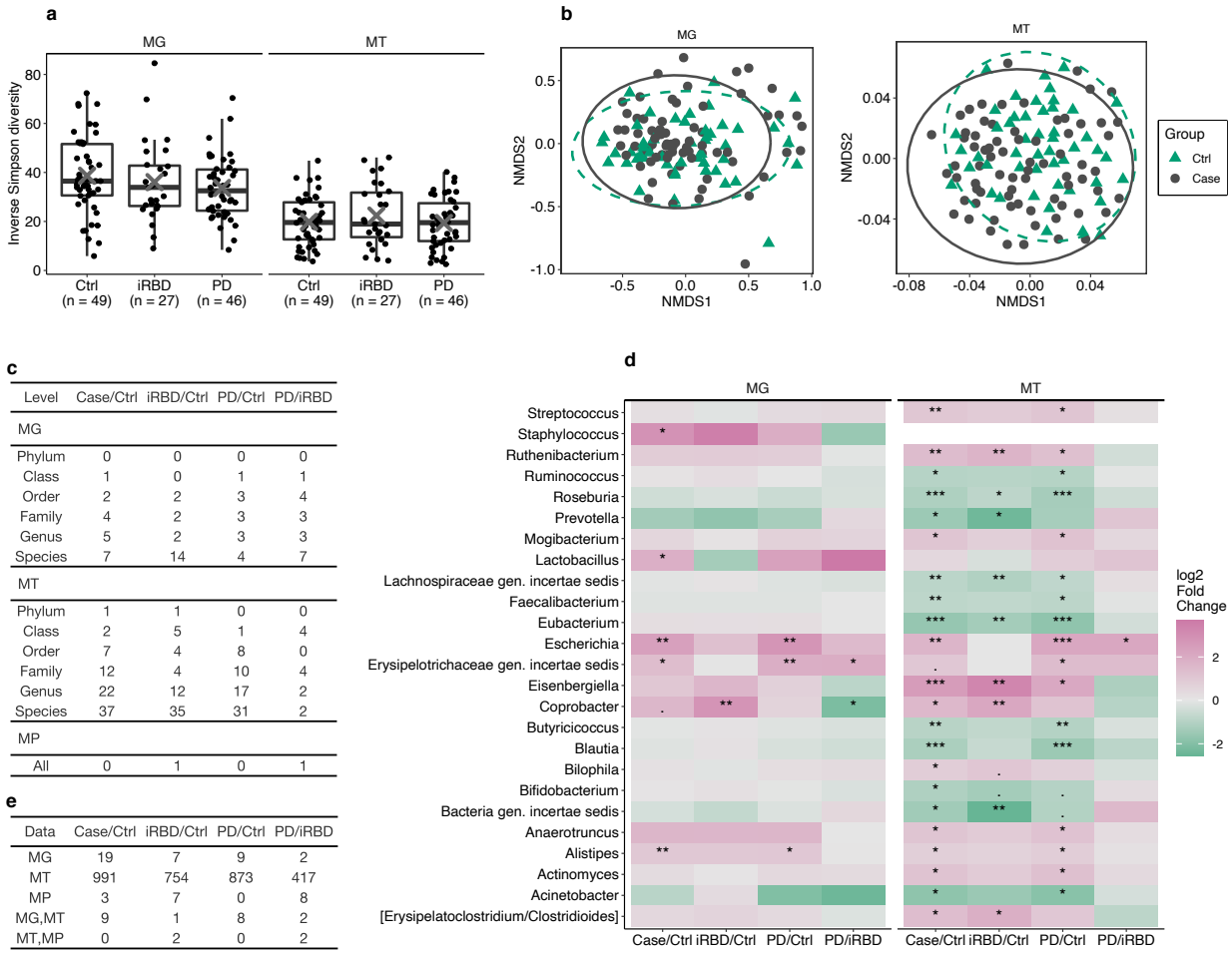
Extended Data Table 1 | Demographic and clinical data on study subjects

Variable	Ctrl	iRBD	PD
a. Main multi-omic analyses			
n	49	27	46
Sex (f/m/NA)	26/23/0	4/23/0	19/27/0
Constipation (no/yes/NA)	26/3/20	13/14/0	21/22/3
Antidepressant medication (no/yes/NA)	26/3/20	12/1/14	35/1/10
Metformin medication (no/yes/NA)	27/2/20	11/2/14	34/2/10
Statin medication (no/yes/NA)	27/2/20	9/4/14	31/5/10
PPI medication (no/yes/NA)	27/2/20	NA	29/7/10
Levodopa medication (no/yes/NA)	NA	NA	5/31/10
Agonist medication (no/yes/NA)	NA	NA	10/26/10
Entacapone medication (no/yes/NA)	29/0/20	13/0/14	31/5/10
Age sampling	68.14 ± 6.07	65.96 ± 7.83	65.8 ± 9.85
Disease duration months	NA	76.37 ± 56.3	63.46 ± 42.52
PD NMS sum	4.52 ± 3.05	8.31 ± 3.53	7.83 ± 3.65
UPDRS I III sum	3.31 ± 4.26	6.08 ± 4.94	31.47 ± 17
Scopa AUT sum	8.86 ± 6.22	10.17 ± 4.91	13.09 ± 6.21
Sniffin Sticks Identification	12.07 ± 2.07	7.46 ± 3.31	6.39 ± 4.15
Hoehn and Yahr	0 ± 0	0 ± 0	2.01 ± 0.77
b. 2-HP targeted validation			
n	30		30
Sex (f/m/NA)	12/18/0		8/22/0
Constipation (no/yes/NA)	27/2/1		21/9/0
Antidepressant medication (no/yes/NA)	28/2/0		24/6/0
Metformin medication (no/yes/NA)	30/0/0		29/1/0
Statin medication (no/yes/NA)	30/0/0		27/3/0
PPI medication (no/yes/NA)	30/0/0		23/7/0
Levodopa medication (no/yes/NA)	NA		2/28/0
Agonist medication (no/yes/NA)	NA		8/22/0
Entacapone medication (no/yes/NA)	30/0/0		25/5/0
Age sampling	68.63 ± 5.18		66 ± 8.25
Disease duration months	NA		94.43 ± 28.5
PD NMS sum	4.09 ± 2.11		8.55 ± 4.02
UPDRS I III sum	1.27 ± 2.39		38.27 ± 21.87
Scopa AUT sum	7.5 ± 4.09		14.74 ± 8.52
Sniffin Sticks Identification	12.93 ± 2.26		6.79 ± 3.76
Hoehn and Yahr	0 ± 0		2.05 ± 0.88
c. NCER-PD cohort (16S rRNA gene amplicons)			
n	343		194
Sex (f/m/NA)	145/198/0		63/131/0
Constipation (no/yes/NA)	325/18/0		116/78/0
Levodopa medication (no/yes/NA)	341/2/0		55/139/0
Other dopaminergic treatment (no/yes/NA)	340/3/0		45/149/0
Age sampling	60.15 ± 12.25		67.09 ± 9.69
Age at PD onset	NA		61.15 ± 10.99

NA = data not available, Ctrl = control subject, PD = Parkinson's disease patient, iRBD = patient with idiopathic REM sleep behaviour disorder, PPI = proton pump inhibitor, PD NMS = Non-Motor Symptoms questionnaire, UPDRS I III = Unified Parkinson's Disease Rating Scale, Scopa AUT = Scales for Outcomes in Parkinson's Disease - Autonomic Dysfunction.



Extended Data Figure 2 | Independent Principal Component Analysis (IPCA) for each data type. Testing was performed separately for each data type, with total number of components chosen based on kurtosis. The first two components are shown for each data type regardless of total number of components. In the 1st column, green triangles and dashed line correspond to control subjects, and black circles and solid line to patients with PD or iRBD. Ellipses reflect 95% confidence. 2nd and 3rd columns show 10 features with the highest loadings. **a.** Untargeted metabolomics; showing 2/2 components, **b.** Taxonomically classified metagenomic reads; 2/3 components, **c.** Taxonomically classified metatranscriptomic reads; 2/2 components, **d.** Taxonomically classified metaproteomic spectra; 2/4 components, **e.** Functionally classified (KEGG orthologs; KOs) metagenomic reads; 2/3 components, **f.** Functionally classified (KOs) metatranscriptomic reads; 2/4 components, **g.** Functionally classified (KOs) metaproteomic spectra; 2/4 components. The numbers in taxon names (**b, c**) refer to identifiers from the mOTUs workflow. For KOs (**e-g**), details are provided in Supplementary Table 3.



c

Level	Case/Ctrl	iRBD/Ctrl	PD/Ctrl	PD/iRBD
MG				
Phylum	0	0	0	0
Class	1	0	1	1
Order	2	2	3	4
Family	4	2	3	3
Genus	5	2	3	3
Species	7	14	4	7
MT				
Phylum	1	1	0	0
Class	2	5	1	4
Order	7	4	8	0
Family	12	4	10	4
Genus	22	12	17	2
Species	37	35	31	2
MP				
All	0	1	0	1

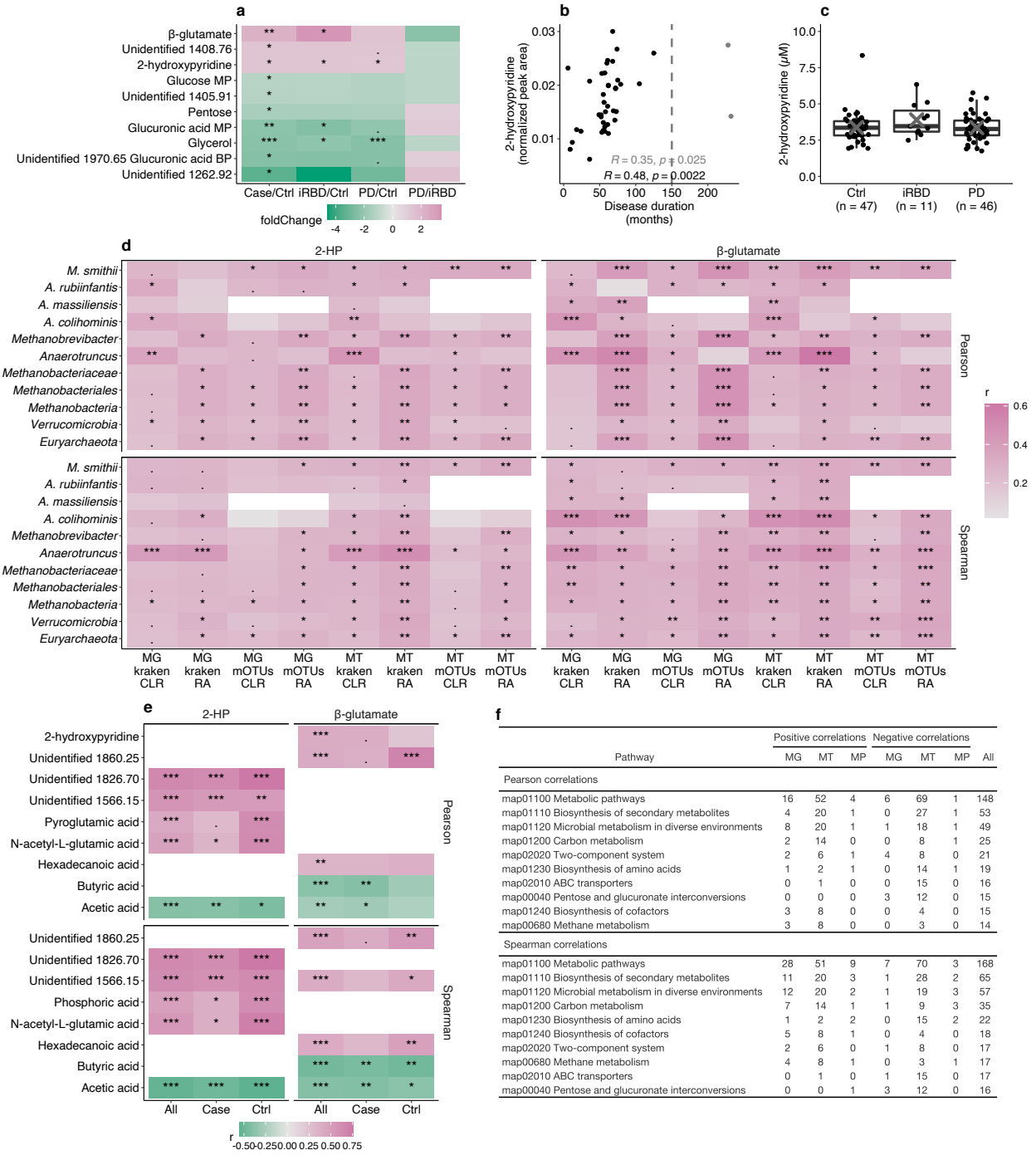
e

Data	Case/Ctrl	iRBD/Ctrl	PD/Ctrl	PD/iRBD
MG	19	7	9	2
MT	991	754	873	417
MP	3	7	0	8
MG,MT	9	1	8	2
MT,MP	0	2	0	2

f

	Decreased in cases		Increased in cases			
	MG	MT	MG	MT	MP	All
map01100 Metabolic pathways	2	156	9	173	2	342
map01110 Biosynthesis of secondary metabolites	1	61	3	61	1	127
map01120 Microbial metabolism in diverse environments	0	41	2	59	0	102
map02020 Two-component system	0	21	1	51	0	73
map01230 Biosynthesis of amino acids	0	28	1	20	0	49
map02010 ABC transporters	0	27	0	21	0	48
map01200 Carbon metabolism	0	16	1	30	0	47
map01240 Biosynthesis of cofactors	2	12	0	21	1	36
map02040 Flagellar assembly	0	25	0	1	0	26
map00620 Pyruvate metabolism	0	7	0	17	0	24

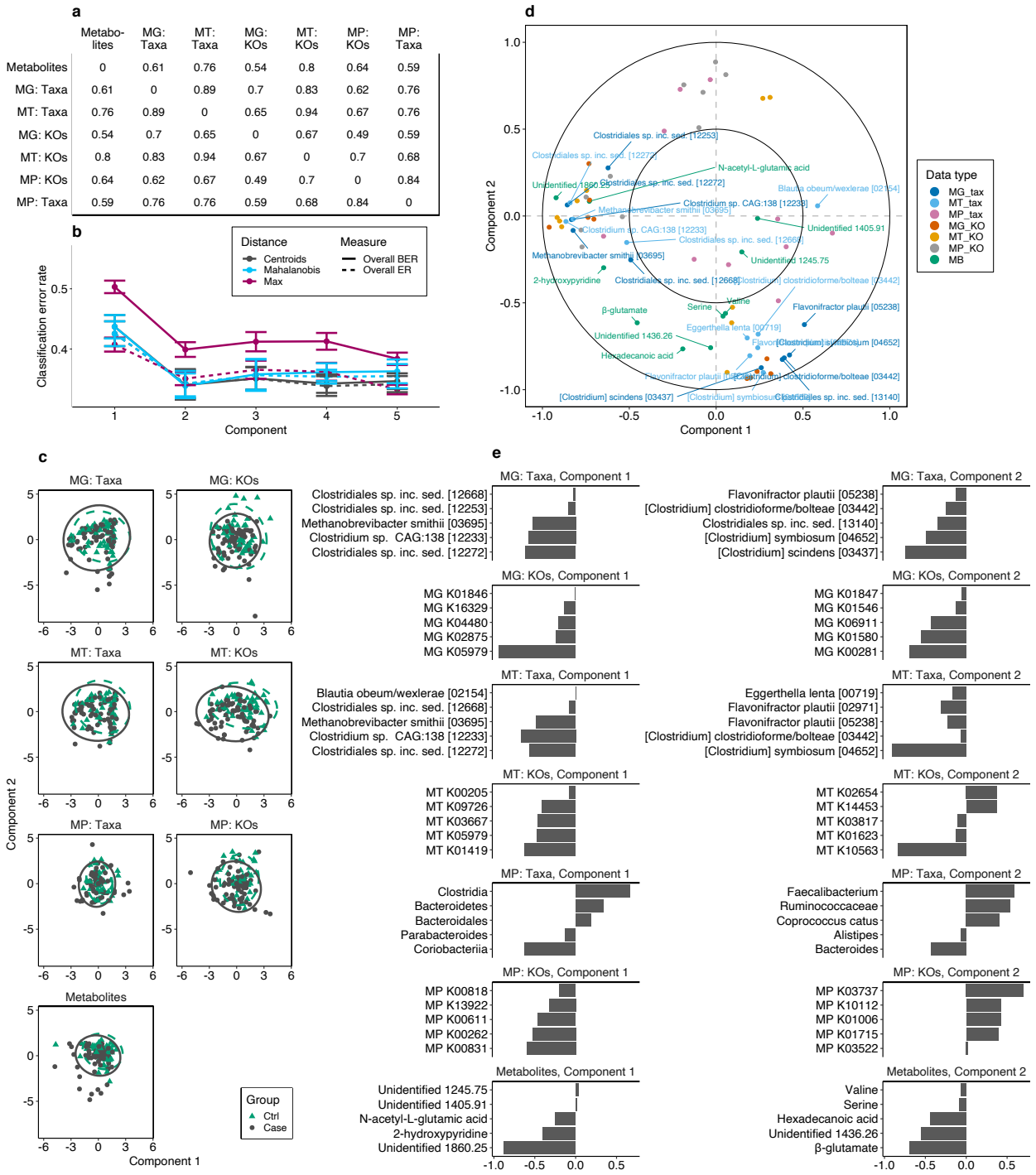
Extended Data Figure 3 | Microbial diversity and differential abundance comparisons of taxonomic and functional data. a. Alpha diversity and group in taxonomically classified metagenomic (MG) and metatranscriptomic (MT) sequence reads; box hinges: 1st and 3rd quartiles; whiskers: hinge to highest/lowest values that is within 1.5*IQR of hinge; grey cross: mean, **b.** Beta diversity and group (Non-metric Multidimensional Scaling based on Bray-Curtis dissimilarity) in MG and MT species data, with ellipses showing 95% confidence, **c.** Number of differentially abundant taxa (q (false discovery rate adjusted P) < 0.05), **d.** Genera with q < 0.05 in MG or MT data for the two-group case (either PD or iRBD) vs control contrast, **e.** Number of differentially abundant (q < 0.05) KEGG orthologs (KOs) in MG, MT, and metaproteomic (MP) data, **e.** 10 pathways most commonly represented by the differentially abundant KOs (q < 0.05) for the case vs control comparison in MG, MT and MP data. In heatmaps, . 0.1 > q > 0.05, * q < 0.05, ** q < 0.01, *** q < 0.001.



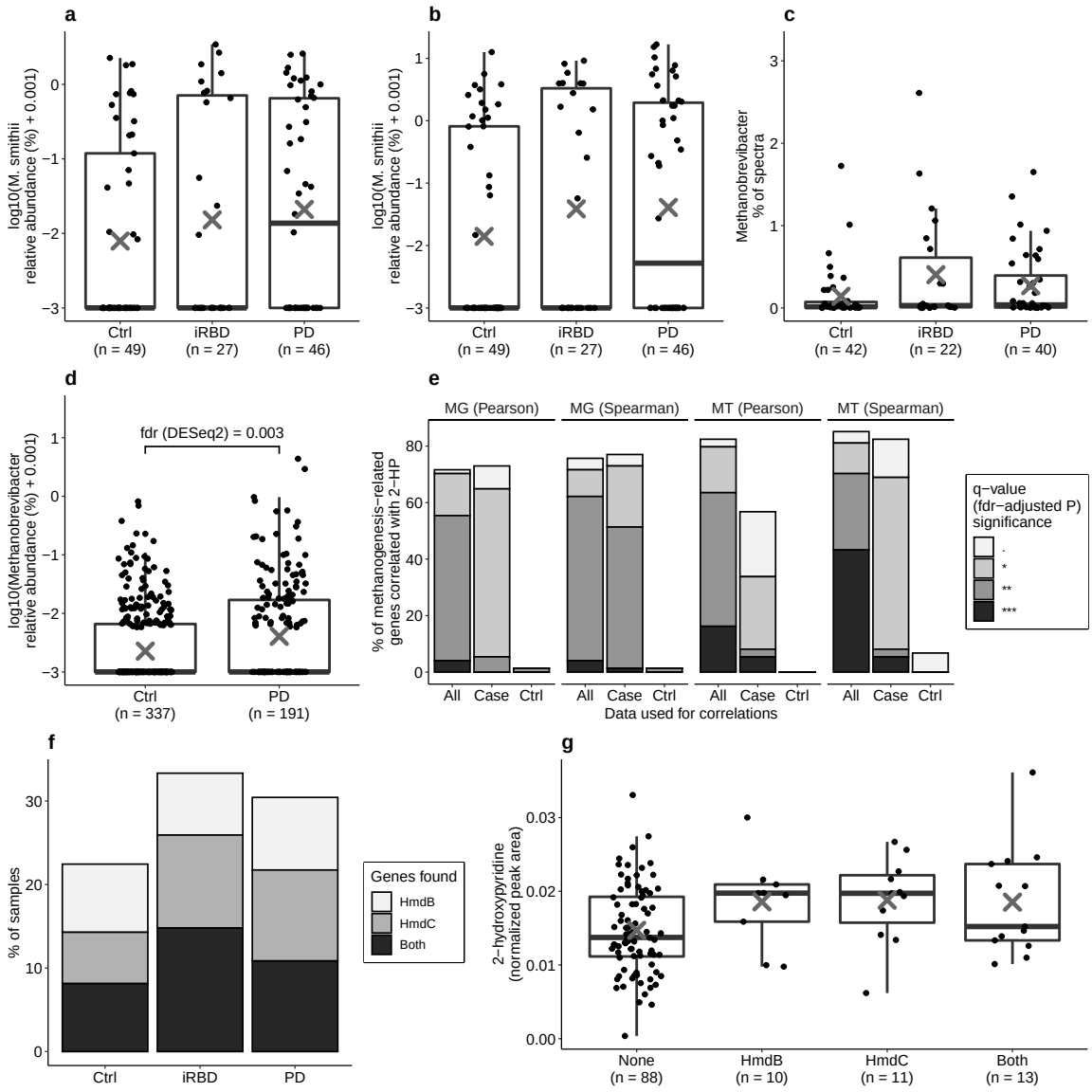
f

Pathway	Positive correlations MG	MT	MP	Negative correlations MG	MT	MP	All
Pearson correlations							
map01100 Metabolic pathways	16	52	4	6	69	1	148
map01110 Biosynthesis of secondary metabolites	4	20	1	0	27	1	53
map01120 Microbial metabolism in diverse environments	8	20	1	1	18	1	49
map01200 Carbon metabolism	2	14	0	0	8	1	25
map02020 Two-component system	2	6	1	4	8	0	21
map01230 Biosynthesis of amino acids	1	2	1	0	14	1	19
map02010 ABC transporters	0	1	0	0	15	0	16
map00040 Pentose and glucuronate interconversions	0	0	0	3	12	0	15
map01240 Biosynthesis of cofactors	3	8	0	0	4	0	15
map00680 Methane metabolism	3	8	0	0	3	0	14
Spearman correlations							
map01100 Metabolic pathways	28	51	9	7	70	3	168
map01110 Biosynthesis of secondary metabolites	11	20	3	1	28	2	65
map01120 Microbial metabolism in diverse environments	12	20	2	1	19	3	57
map01200 Carbon metabolism	7	14	1	1	9	3	35
map01230 Biosynthesis of amino acids	1	2	2	0	15	2	22
map01240 Biosynthesis of cofactors	5	8	1	0	4	0	18
map02020 Two-component system	2	6	0	1	8	0	17
map00680 Methane metabolism	4	8	1	0	3	1	17
map02010 ABC transporters	0	1	0	1	15	0	17
map00040 Pentose and glucuronate interconversions	0	0	1	3	12	0	16

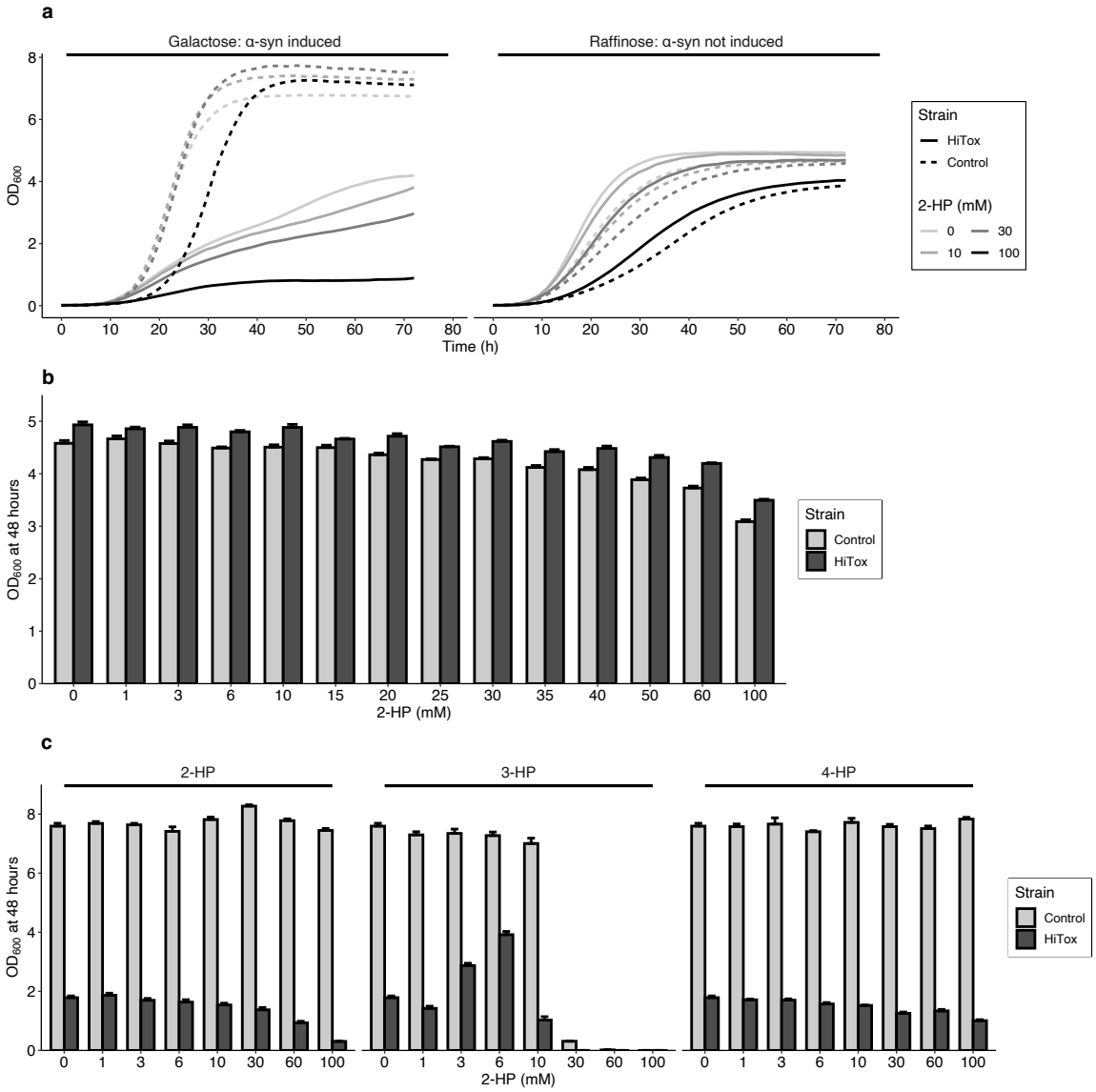
Extended Data Figure 4 | Additional results for metabolomics. a. Metabolites with *q* (false discovery rate adjusted *P*) < 0.05 for the two-group case (PD or iRBD) vs control contrast and at least one pair of three-category groups, **b.** 2-HP normalized peak area and disease duration in PD patients; grey values = full data, black values = two longest-duration subjects excluded, **c.** 2-HP in plasma (targeted metabolomics); 2 extreme outliers excluded (2-HP concentration > 30 μM, both in the iRBD group); box hinges: 1st and 3rd quartiles; whiskers: hinge to highest/lowest values that is within 1.5*IQR of hinge; grey cross: mean, **d.** Correlations of metabolites and taxa, showing taxa with *q* < 0.05 in more than 20 comparisons out of a total of 32 approaches: 2 metabolites, 2 data types (MG or MT), 2 annotation methods (Kraken or mOTUs), 2 normalizations (RA = relative abundance, CLR = centered log ratio transform) and 2 correlation coefficients (Pearson or Spearman), **e.** Correlations of 2-HP and beta-glutamate and other metabolites, showing top 5 by *q* in full data for each metabolite and correlation coefficient, **f.** 10 most common pathways represented by KEGG orthologs significantly correlated with 2-HP (*q* < 0.05, and for MG and MT data, absolute value of correlation coefficient > 0.4). In all plots, MG = metagenomic data, MT = metatranscriptomic data, MP = metaproteomic data. In plots with stars, . 0.1 > *q* > 0.05, * *q* < 0.05, ** *q* < 0.01, *** *q* < 0.001.



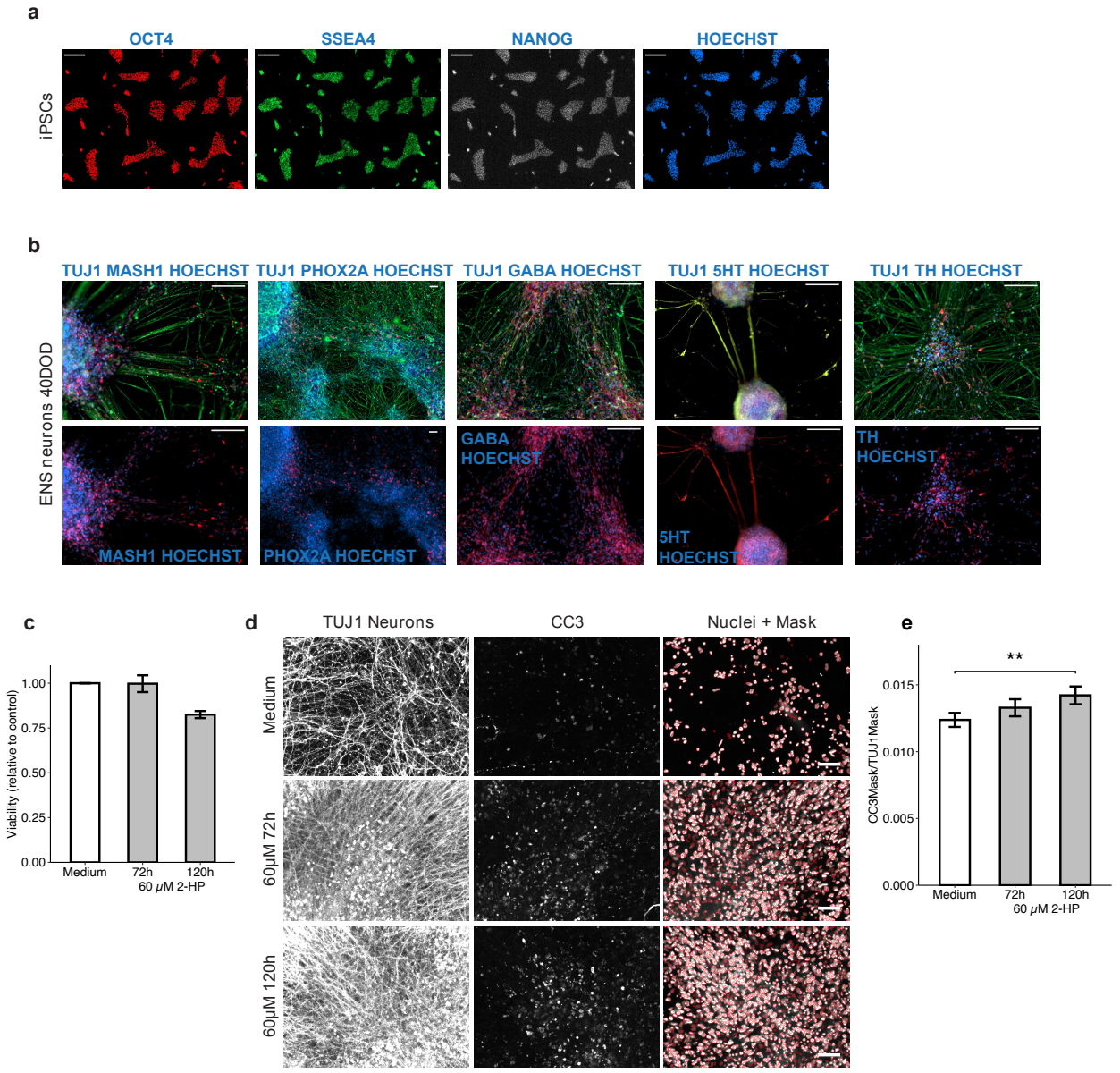
Extended Data Figure 5 | Results of integrated multi-omics with DIABLO. The classification was run for the two-category case (iRBD or PD) vs control variable, selecting the 5 best features for differentiating between these categories on 2 axes for each of the seven data types (MG = metagenomic, MT = metatranscriptomic, MP = metaproteomic; KOs: functional annotations as represented by KEGG orthologs). **a.** Design matrix used for DIABLO, calculated based on running PLS for each pair of data blocks, **b.** Component tuning, based on which 2 was selected as the number of components to use; ER = error rate, BER = balanced error rate, error bars = standard deviations, **c.** Plots of samples and sample classes per data type; green triangles with dashed lines = controls, black circles with soldi lines = PD and iRBD patients, ellipses represent 95% confidence **d.** Correlation circle plot showing the selected features together, with labels for MG and MT taxa and metabolites, **e.** Loadings for the five selected features for each data type and component. Annotations for the selected KOs are provided in Supplementary Table 10.



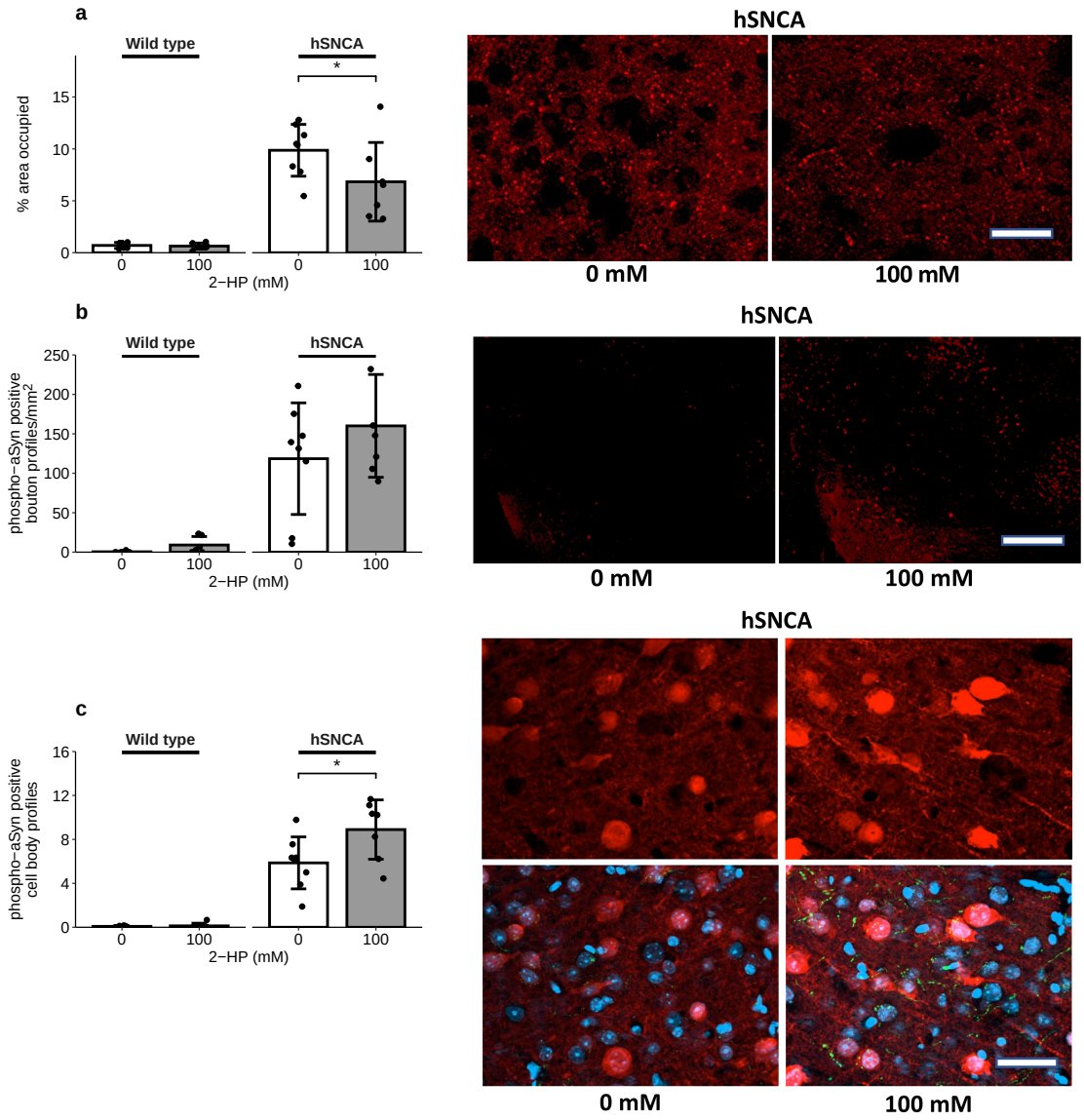
Extended Data Figure 6 | Additional results on 2-hydroxypyridine and genus *Methanobrevibacter*. **a.** log10(relative abundance) of metagenomic reads classified as *Methanobrevibacter smithii* by diagnosis group in main study cohort, **b.** log10(relative abundance) of metatranscriptomic reads classified as *Methanobrevibacter smithii* by diagnosis group in main study cohort, **c.** % of metaproteomic spectra classified as *Methanobrevibacter* by diagnosis group in the main study cohort, **d.** log10(relative abundance) of 16S rRNA gene amplicons classified as *Methanobrevibacter* in the Luxembourgish NCER-PD cohort. **e.** Summaries of *q*-values (fdr-adjusted *P*-values) of correlations for 2-hydroxypyridine and 74 methanogenesis-related genes found in the sequence data (tested either in full data or with data containing samples of a specific group, as given on the x-axis). **f.** % of samples with predicted proteins containing the HmdB or HmdC genes, which code for the biosynthesis of the cofactor for Hmd. **g.** 2-hydroxypyridine normalized peak area depending on the presence of HmdB or HmdC gene in samples. MG = metagenomic data, MT = metatranscriptomic data; in boxplots, box hinges: 1st and 3rd quartiles; whiskers: hinge to highest/lowest values that is within 1.5*IQR of hinge; grey cross: mean.



Extended Data Figure 7 | Additional results from the HiTox yeast model. a. Representative growth curves of control and HiTox strains in presence of 2-hydroxypyridine (2-HP) at the indicated concentrations. **b.** 2-HP dose-response assay in uninduced condition in HiTox and control strains. **c.** Dose-response assay in control strain and HiTox strain in aSyn-expressing condition after treatment with either 2-HP, 3-hydroxypyridine (3-HP) or 4-hydroxypyridine (4-HP). In b and c, OD_{600} was measured 48 h after inoculation and mean and SD were calculated from four biological replicates.



Extended Data Figure 8 | Additional results from the induced human pluripotent stem cell enteric neuron model. a. Induced human pluripotent stem cell characterization stainings show robust expression of OCT4, SEEA4 and NANOG pluripotency markers. Scale bars: 200 µm. **b.** Enteric neuron characterisation stainings at 40 days of differentiation detect presence of neural crest stem cell markers MASH1 and PHOX2A, and of different neuronal identities, such as GABAergic, serotoninergic and dopaminergic. Scale bars: 100 µm. **c.** Relative viability after 2-hydroxypyridine exposure. Comparison to untreated sample was not significant when evaluated with a Kruskal-Wallis test. **d.** High-content imaging of apoptosis marker cleaved-caspase 3 (CC3), TUJ1-positive neurons and Hoechst-positive nuclei. Nuclei images contain a representation of the mask applied to segment the nuclei; scale bar: 100 µm. **e.** Quantification of CC3-positive staining normalized to the amount of TUJ1-positive neurons. In c and d, data is represented as mean ± SEM of three independent neuronal differentiations, with 30 fields per well and 8 wells per condition quantified for each. ** $P < 0.01$.



Extended Data Figure 9 | Effects of intrastriatal injections of 100 mM 2-hydroxypyridine in transgenic mouse model of PD on phospho-aSyn accumulation in different brain regions. a. Striatum: 2-HP-induced decrease of phospho-aSyn signals in synaptic boutons; scale bar: 50 µm, **b.** Substantia nigra: 2-HP-induced non-significant increase of phospho-aSyn in synaptic boutons; scale bar: 350 µm, **c.** Prefrontal cortex: 2-HP-induced significant increase of phospho-aSyn signals in cell body profiles with neuronal morphology (double-stained with DAPI, lower row of images); scale bar: 50 µm. Microphotographs show examples for 0 mM and 100 mM in hSNCA mice only. Bar plots show mean and standard deviation; n of mice = 5-8 per group. *P*-values are given between group means that are different. * $P < 0.05$.