

Supplementary Table 1: Annotated differentially enriched metabolites between the newly diagnosed T2DM and IGR groups

Supplementary Table 2: Annotated differentially enriched metabolites between the newly diagnosed T2DM and NGT groups

Supplementary Table 3: Annotated differentially enriched metabolites between the IGR and NGT groups

Supplementary Figure 1: Box plot of COG function classification. Microbial sequences mapped by Greengenes database, the functional gene contents of the gut microbiota were predicted by PICRUSt and were mapped on COG.

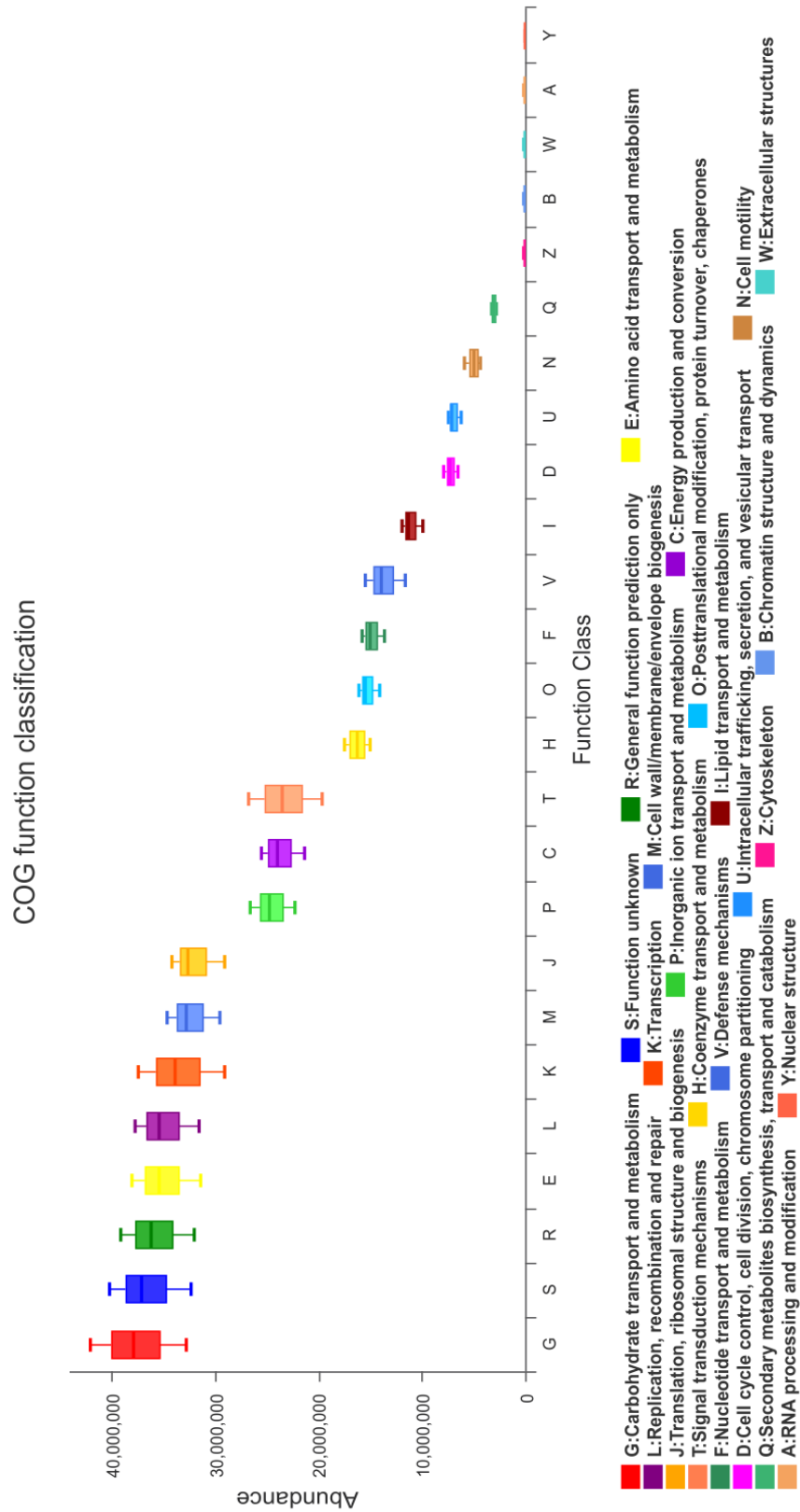
Supplementary Figure 2: Correlation heatmap of differentially enriched metabolites and gut microbiota phylum between the newly diagnosed T2DM and IGR group (a); Correlation heatmap of differentially enriched metabolites and gut microbiota phylum between the newly diagnosed IGR and NGT group (b)

Note: Each column in the graph represents a metabolite, each row represents a microbiota phylum, the colour in the graph indicates the Pearson coefficient between the bacterial phylum and the metabolite, and the red colour indicates positive correlation. The green colour is representative negative correlation. The darker colour indicate the greater the correlation. \* is represents a significant correlation ( $P < 0.05$ ).

Supplementary Figure 3: Correlation heatmap of differentially enriched metabolites and gut microbiota genus between the newly diagnosed T2DM and IGR group (a); Correlation heatmap of differentially enriched metabolites and gut microbiota genus between the newly diagnosed IGR and NGT group (b).

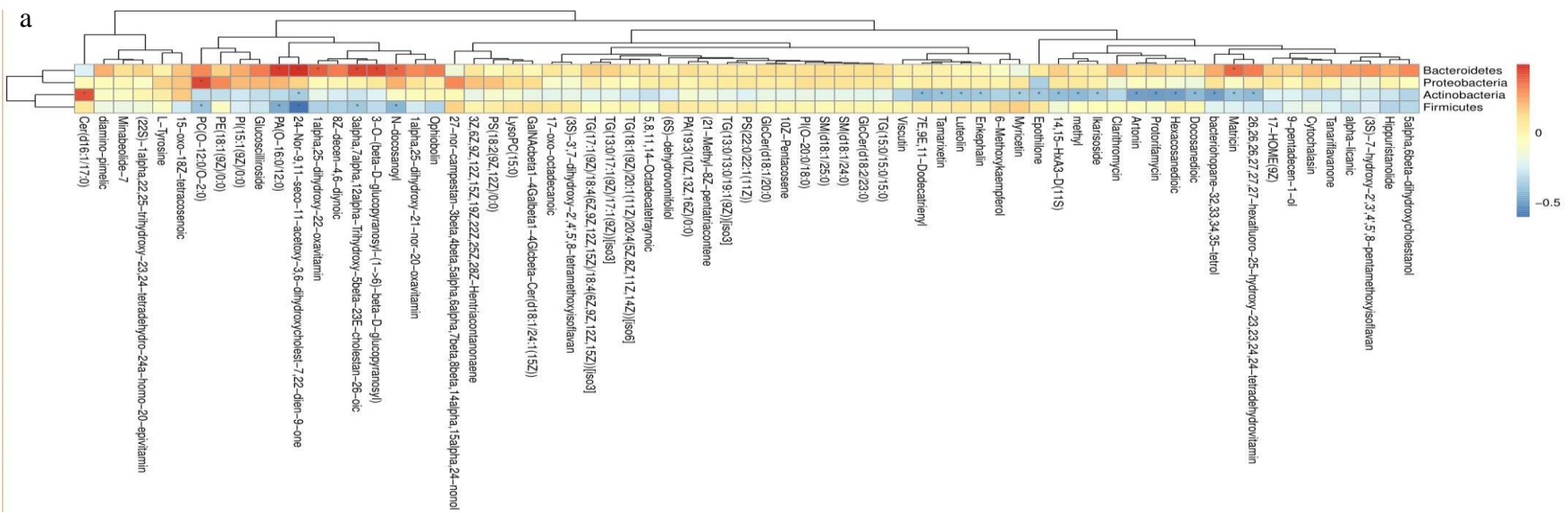
Note: Each column in the graph represents a metabolite, each row represents a microbiota genus, the colour in the graph indicates the Pearson coefficient between the bacterial genus and the metabolite, and the red colour indicates positive correlation. The green colour is representative negative correlation. The darker colour indicate the greater the correlation. \* is represents a significant correlation ( $P < 0.05$ ).

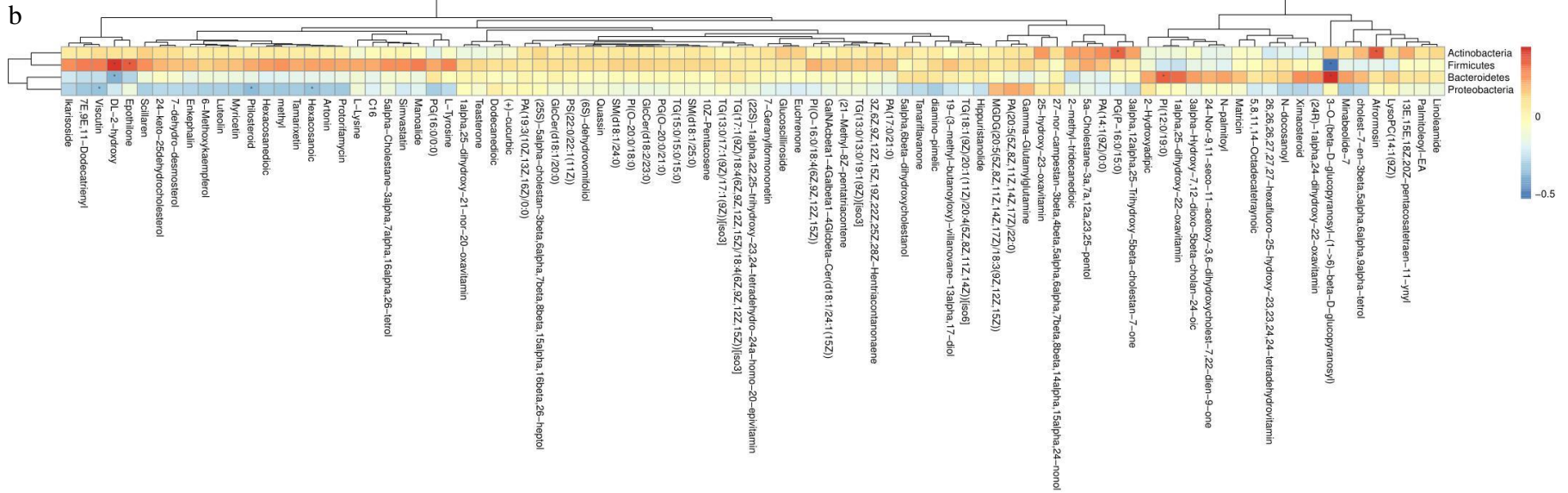
Supplementary Figure 4: Correlation heatmap of differentially enriched metabolites and gut microbiota phylum level between the newly diagnosed T2DM and NGT group (a), Correlation heatmap of differentially enriched metabolites and gut microbiota genus level between the newly diagnosed T2DM and NGT group (b).



1  
2  
3  
4  
5  
6  
7

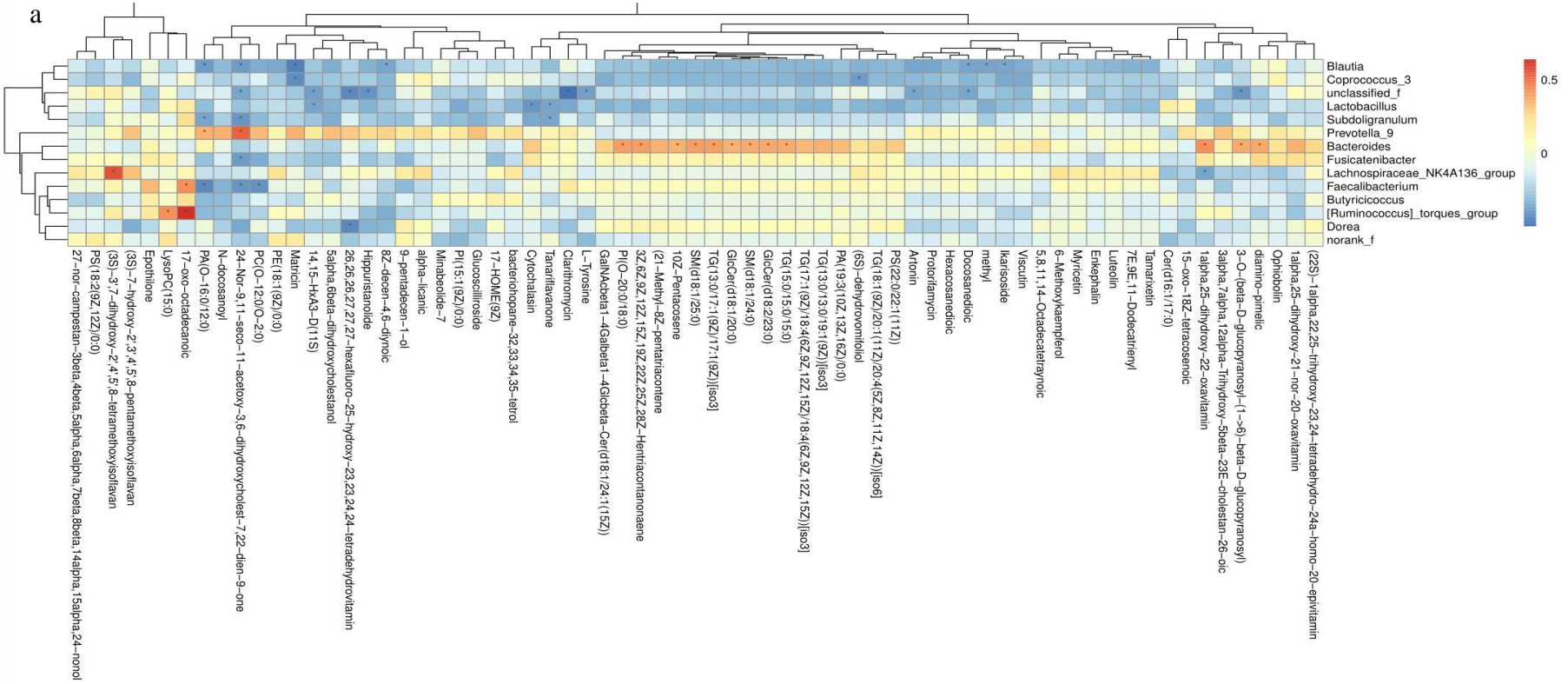
Supplementary Figure 1: Box plot of COG function classification. Microbial sequences mapped by Greengenes database, the functional gene contents of the gut microbiota were predicted by PICRUSt and were mapped on COG.

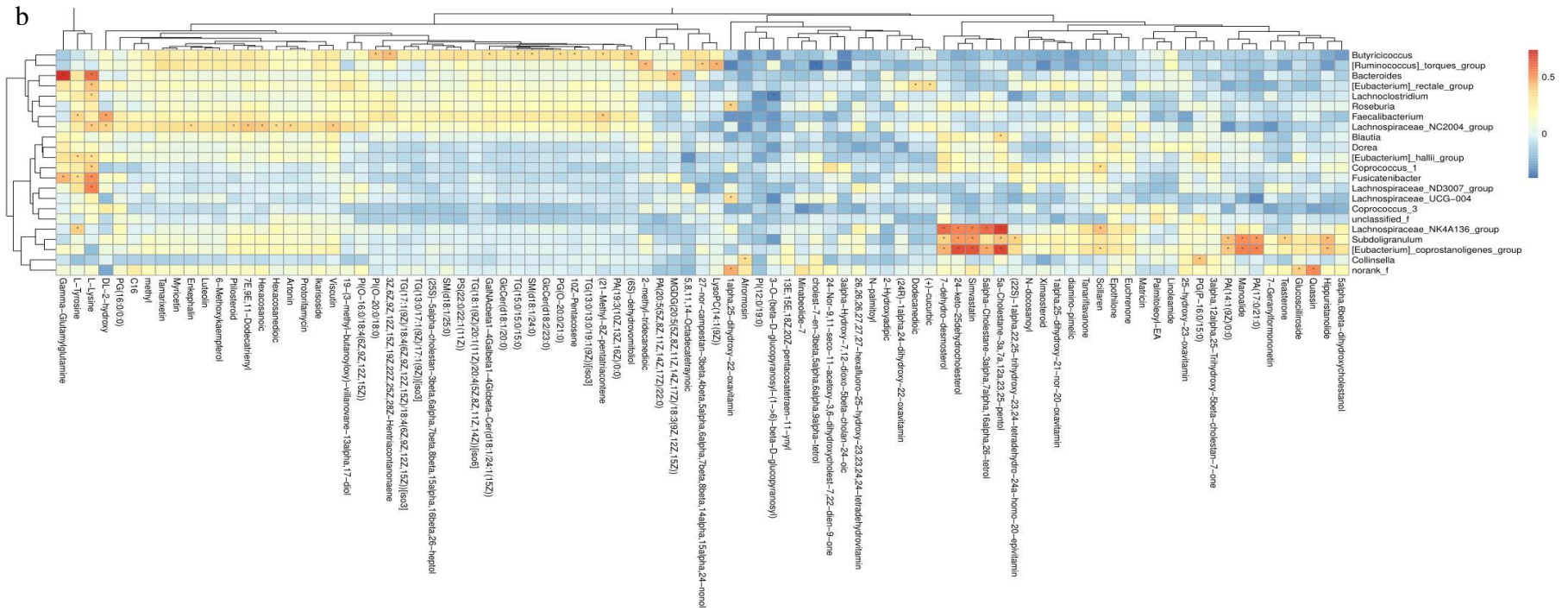




Supplementary Figure 2: Correlation heatmap of differentially enriched metabolites and gut microbiota phylum between the newly diagnosed T2DM and IGR group (a); Correlation heatmap of differentially enriched metabolites and gut microbiota phylum between the newly diagnosed IGR and NGT group (b)

Note: Each column in the graph represents a metabolite, each row represents a microbiota phylum, the colour in the graph indicates the Pearson coefficient between the bacterial phylum and the metabolite, and the red colour indicates positive correlation. The green colour is representative negative correlation. The darker colour indicate the greater the correlation. \* is represents a significant correlation ( $P < 0.05$ ).

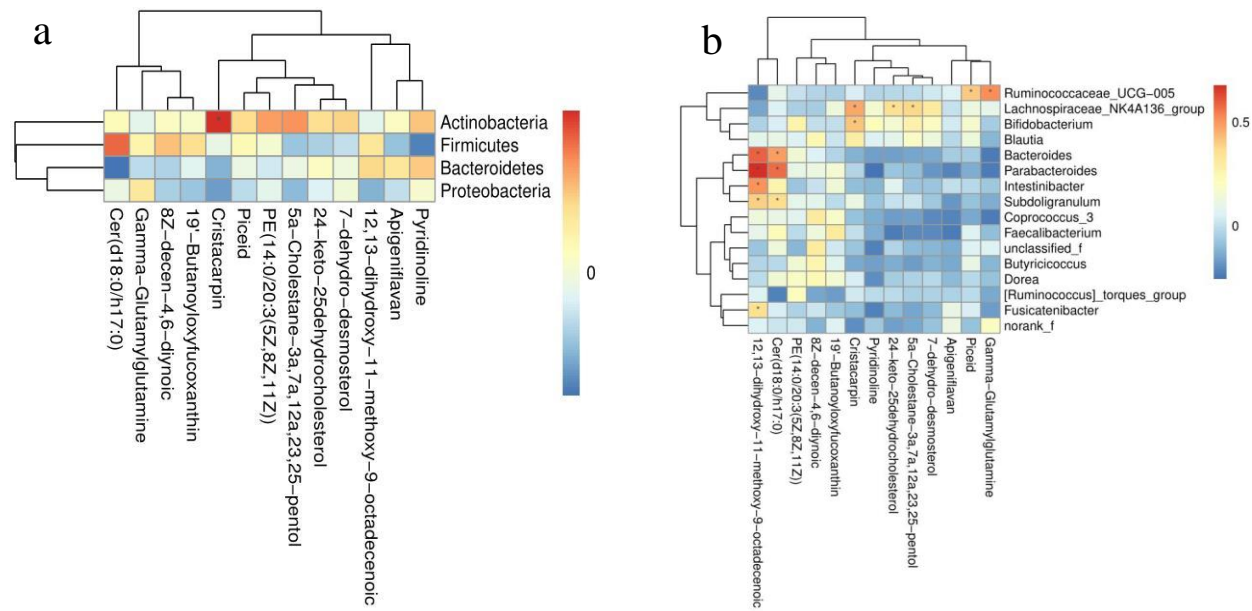




Supplementary Figure 3: Correlation heatmap of differentially enriched metabolites and gut microbiota genus between the newly diagnosed T2DM and IGR group (a); Correlation heatmap of differentially enriched metabolites and gut microbiota genus between the newly diagnosed IGR and NGT group (b).

Note: Each column in the graph represents a metabolite, each row represents a microbiota genus, the colour in the graph indicates the Pearson coefficient between the bacterial genus and the metabolite, and the red colour indicates positive correlation. The green colour is representative negative correlation. The darker colour indicate the greater the correlation. \* is represents a significant correlation ( $P < 0.05$ ).





42  
43  
44  
45  
46

Supplementary Figure 4: Correlation heatmap of differentially enriched metabolites and gut microbiota phylum level between the newly diagnosed T2DM and NGT group (a), Correlation heatmap of differentially enriched metabolites and gut microbiota genus level between the newly diagnosed T2DM and NGT group (b).

1

Supplementary Table 1: Annotated differentially enriched metabolites between the newly diagnosed T2DM and IGR groups

Metabolites	m/z	RT (min)	Compound ID	Formula	VIP value	P value	FC (T2DM/IGR)	Pathways	superclass	class	subclass
L-Tyrosine	180.07	1.14	HMDB00158	C9H11NO3	1.36	0.00	1.64	Phenylalanine, tyrosine and tryptophan biosynthesis	Organic acids and derivatives	Carboxylic acids and derivatives	Amino acids, peptides, and analogues
Epothilone A	476.25	5.84	LMPK04000040	C26H39NO6S	4.02	0.03	0.79	Type I polyketide structures	Lipids and lipid-like molecules	Polyketides	Macrolides and lactone polyketides
Anhydrorhodovibrin	565.44	6.68	LMPR01070135	C41H58O	1.86	0.01	3.27	Metabolic pathways,Carotenoid biosynthesis	Lipids and lipid-like molecules	Prenol lipids	Isoprenoids
Protorifamycin I	657.34	11.13	LMPK05000007	C35H45NO10	1.71	0.00	0.78	Biosynthesis of ansamycins,Biosynthesis of antibiotics	Phenylpropanoids and polyketides	Coumarins and derivatives	Furanocoumarins
Pimelic acid	159.07	1.16	LMFA01170051	C7H12O4	1.04	0.05	0.15	Metabolic pathways	Lipids and lipid-like molecules	Fatty acyls	Fatty acids and conjugates
LysoPC(15:0)	480.31	6.26	HMDB10381	C23H48NO7P	5.44	0.05	0.36	Glycerophospholipid metabolism,Choline metabolism in cancer	Lipids and lipid-like molecules	Glycerophospholipids	Glycerophosphocholines
Matricin	613.30	2.35	LMPR0103410003	C17H22O5	3.11	0.02	0.64	Unknown	Organoheterocyclic compounds	Lactones	Gamma butyrolactones
bacteriohopane-32,33,34,35-tetrol	564.50	7.83	LMPR04000004	C35H62O4	1.80	0.05	2.06	Unknown	Lipids and lipid-like molecules	Prenol lipids	Hopanoids
5alpha,6beta-dihydroxycholestanol	443.35	5.23	LMST01010052	C27H48O3	1.41	0.02	1.33	Unknown	Lipids and lipid-like molecules	Sterol lipids	Sterols
Cytochalasin A	460.25	2.12	LMPK11000001	C29H35NO5	1.24	0.02	1.26	Unknown	Lipids and lipid-like molecules	Polyketides	Cytochalasins
(6S)-dehydrovomifoliol	221.12	3.27	LMPR0103050009	C13H18O3	1.08	0.00	0.74	Unknown	Lipids and lipid-like molecules	Prenol lipids	Isoprenoids

2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12



13  
14  
15

Supplementary Table 2: Annotated differentially enriched metabolites between the newly diagnosed T2DM and NGT groups

Metabolites	m/z	RT (min)	Compound ID	Formula	VIP value	p value	Pathways	superclass	class	subclass
Cristacarpin	355.15	3.25	LMPK12070114	C21H22O5	3.85	0.03	Unknown	Phenylpropanoids and polyketides	Isoflavonoids	Furanoisoflavonoids
Piceid	373.13	2.35	LMPK13090012	C20H22O8	3.49	0.01	Unknown	Phenylpropanoids and polyketides	Stilbenes	Stilbene glycosides
7-dehydro-desmosterol	383.33	4.95	LMST01010121	C27H42O	2.63	0.00	Steroid biosynthesis, Metabolic pathways	Lipids and lipid-like molecules	Sterol lipids	Sterols
PE(14:0/20:3)	712.49	9.32	HMDB08836	C39H72NO8P	1.78	0.02	Unknown	Lipids and lipid-like molecules	Glycerophospholipids	Glycerophosphoethanolamines
Gamma-Glutamylglutamine	274.10	0.64	HMDB11738	C10H17N3O6	1.19	0.03	Unknown	Organic acids and derivatives	Carboxylic acids and derivatives	Amino acids, peptides, and analogues

16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37

Supplementary Table 3: Annotated differentially enriched metabolites between the IGR and NGT groups

Metabolites	m/z	RT (min)	Compound ID	Formula	VIP value	p value	FC (IGR/NGT)	Pathways	superclass	class	subclass
Dodecanedioic acid	229.14	3.27	HMDB00623	C12H22O4	4.60	0.01	2.43	Biosynthesis of secondary metabolites, alpha-Linolenic acid metabolism	Lipids and lipid-like molecules	Fatty acyls	Fatty acids and conjugates
Matricin	613.30	2.35	LMPR010341003	C17H22O5	3.05	0.01	1.68	unknown	Organoheterocyclic compounds	Lactones	Gamma butyrolactones
L-Tyrosine	180.07	1.14	HMDB00158	C9H11NO3	2.67	0.00	0.47		Organic acids and derivatives	Carboxylic acids and derivatives	Amino acids, peptides, and analogues
LysoPC(14:1)	464.28	6.06	HMDB10380	C22H44NO7P	2.59	0.01	2.92	Choline metabolism in cancer, Glycerophospholipid metabolism	Lipids and lipid-like molecules	Glycerophospholipids	Glycerophosphocholines
7-dehydro-desmosterol	383.33	4.95	LMST01010121	C27H42O	2.23	0.01	0.58	Steroid biosynthesis, metabolic pathways	Lipids and lipid-like molecules	Sterol lipids	Sterols
5alpha,6beta-dihydroxycholestanol	443.35	5.23	LMST01010052	C27H48O3	2.15	0.01	0.66	unknown	Lipids and lipid-like molecules	Sterol lipids	Sterols
Epothilone A	476.25	5.84	LMPK04000040	C26H39NO6S	1.96	0.04	1.24	Type I polyketide structures	Lipids and lipid-like molecules	Polyketides	Macrolides and lactone polyketides
Protorifamycin I	657.34	11.13	LMPK05000007	C35H45NO10	1.90	0.00	1.29	Biosynthesis of ansamycins, Biosynthesis of antibiotics	Phenylpropanoids and polyketides	Coumarins and derivatives	Furanocoumarins
(+)-cucurbitic acid	211.13	3.27	LMFA02020013	C12H20O3	1.62	0.00	1.91	unknown	Lipids and lipid-like molecules	Sterol lipids	Sterols
Teasterone	447.35	5.63	LMST01030121	C28H48O4	1.61	0.01	0.37	Biosynthesis of secondary metabolites, Brassinosteroid biosynthesis, Metabolic pathways	Lipids and lipid-like molecules	Fatty acyls	Octadecanoids
Manoalide	415.25	3.85	LMPR010502001	C25H36O5	1.59	0.01	0.40	unknown	Lipids and lipid-like molecules	Sterol lipids	Sterols
Gamma-Glutamylglutamine	274.10	0.64	HMDB11738	C10H17N3O6	1.50	0.00	0.43	unknown	Organic acids and derivatives	Carboxylic acids and derivatives	Amino acids, peptides, and analogues
(6S)-dehydrovomifoliol	221.12	3.27	LMPR010305009	C13H18O3	1.25	0.01	1.35	unknown	Lipids and lipid-like molecules	Prenol lipids	Isoprenoids
Estrone 3-glucuronide	445.19	1.40	LMST05010011	C24H30O8	1.22	0.00	0.09	Steroid hormone biosynthesis	Lipids and lipid-like molecules	Steroids and steroid derivatives	Steroidal glycosides
Anhydrosorbidol	565.44	6.68	LMPR01070135	C41H58O	1.17	0.02	0.37	Carotenoid biosynthesis, Metabolic	Lipids and lipid-like molecules	Prenol lipids	Isoprenoids

								pathways			
Quassin	389.2 0	2.68	LMPR01061100 02	C22H28O6	1.11	0.02	0.71	unknown	Lipids and lipid- like molecules	Prenol lipids	Terpene lactones
Scillaren A	715.3 3	2.30	LMST01130005	C36H52O13	1.03	0.04	0.74	unknown	Lipids and lipid- like molecules	Sterol lipids	Sterols
2-Hydroxyadipic acid	161.0 5	0.64	LMFA01170049	C6H10O5	1.01	0.04	1.97	unknown	Lipids and lipid- like molecules	Fatty acyls	Fatty acids and conjugates
L-Lysine	145.1 0	0.55	HMDB00182	C6H14N2O2	1.01	0.00	0.46		Organic acids and derivatives	Carboxylic acids and derivatives	Amino acids, peptides, and analogues