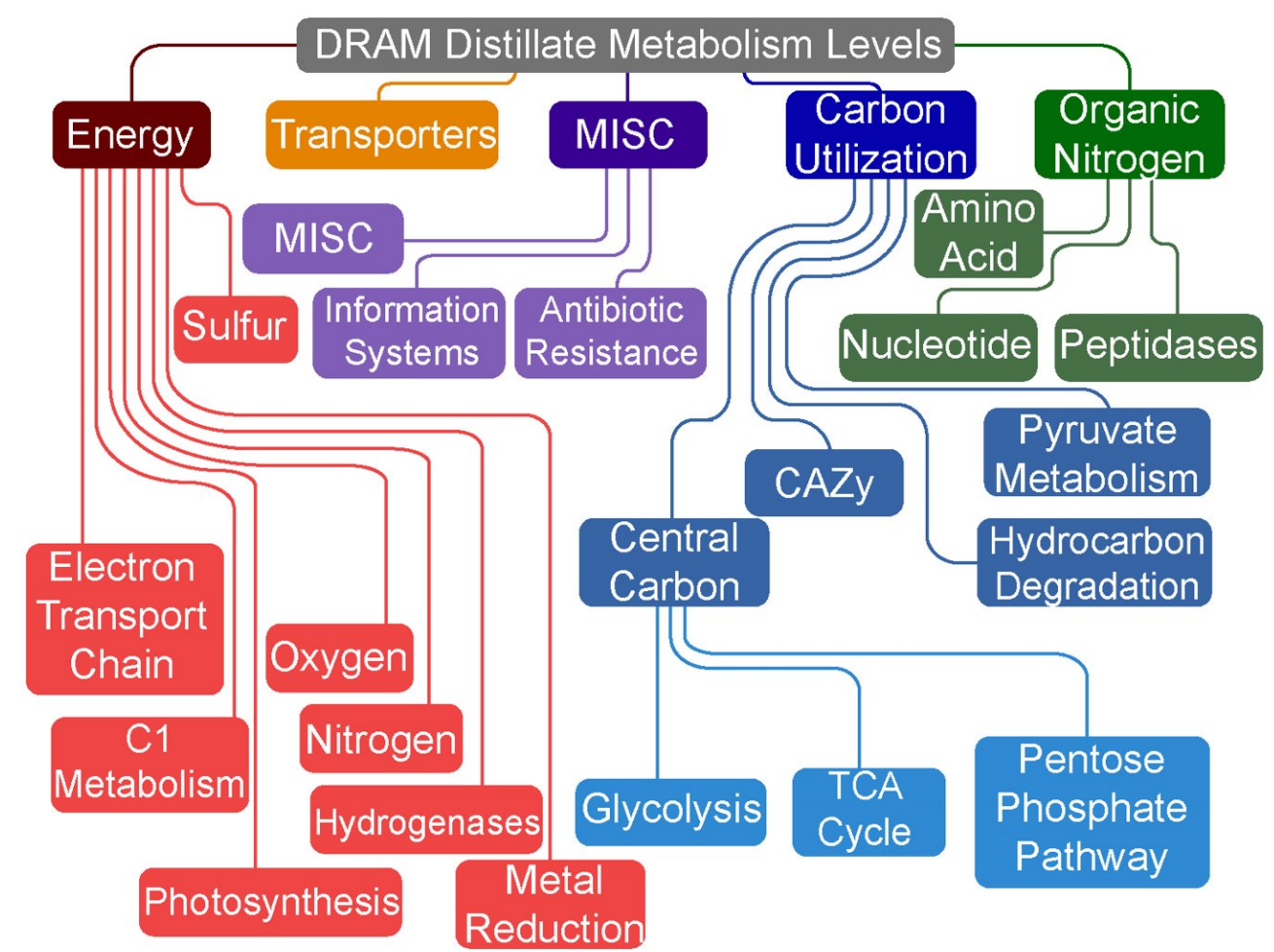
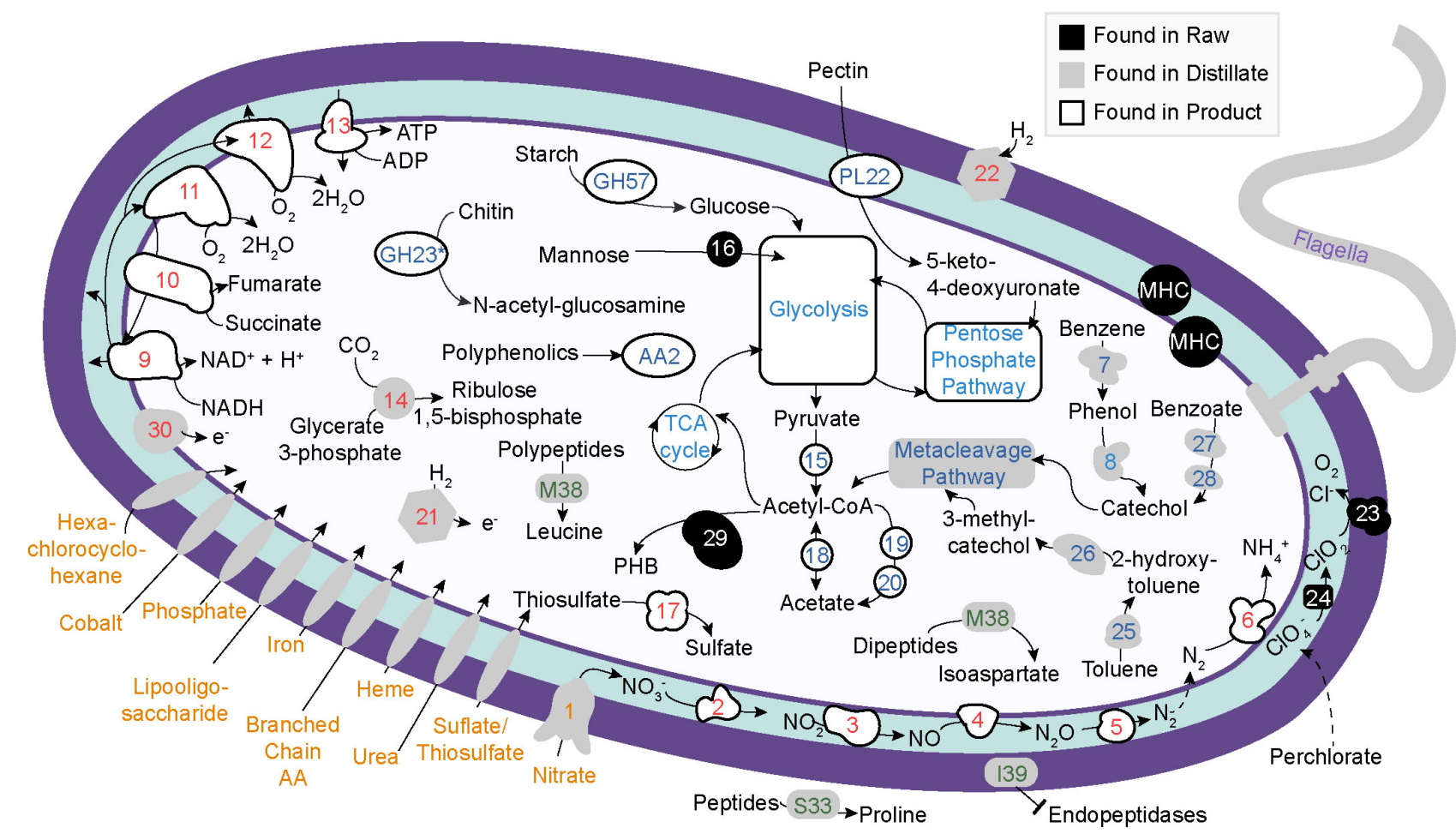


DRAM: DISTILLED AND REFINED ANNOTATION OF METABOLISM ENABLES RAPID UNDERSTANDING OF MICROBIAL AND VIRAL FUNCTIONAL POTENTIAL

 Rory M. Flynn

THE PROBLEM!

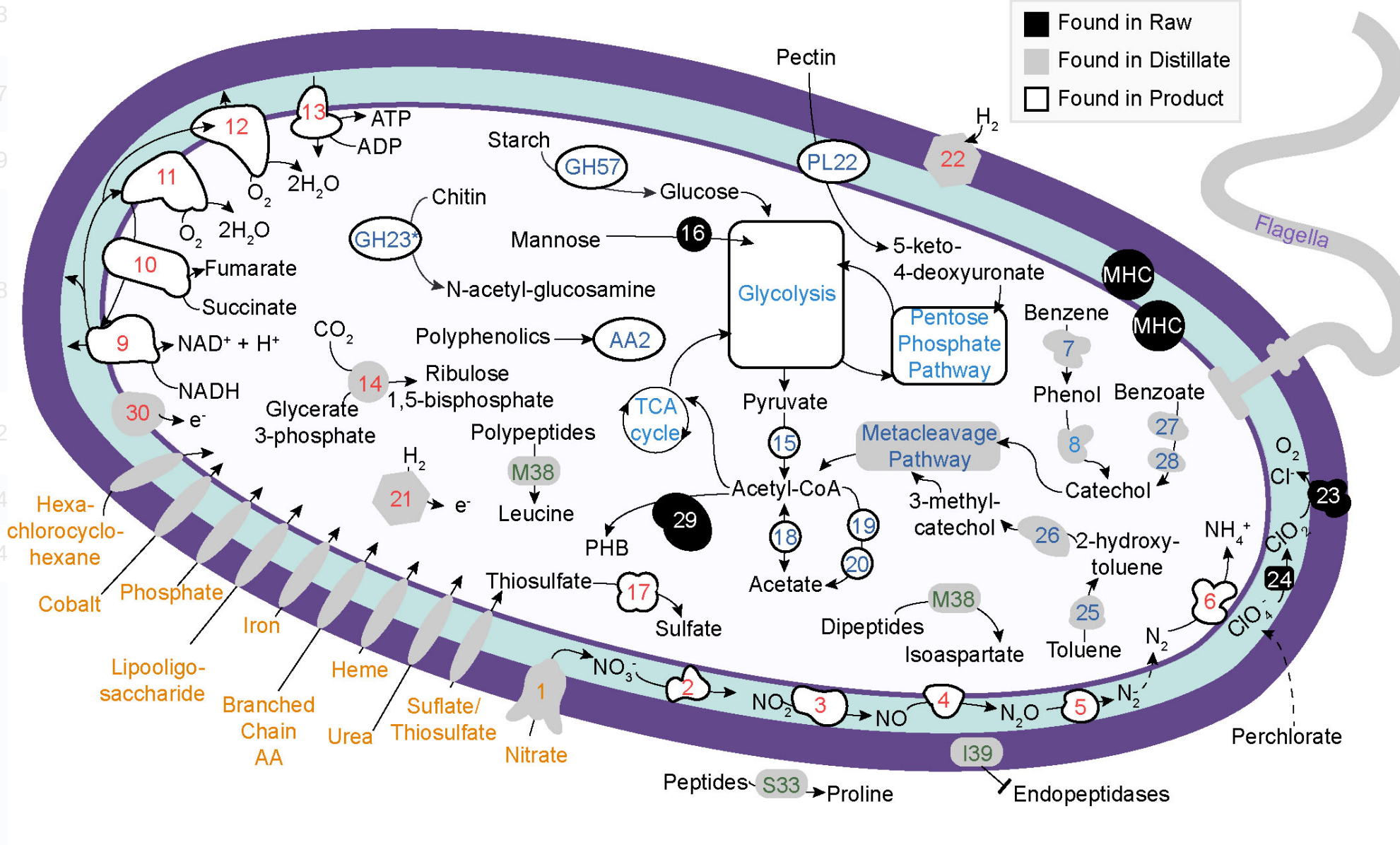


fasta	scaffold	gene_position	start_position	end_position	strandedness	rank	kegg_id	kegg_hit	kegg_RBH	kegg_identity	kegg_bitScore	kegg_eVal	peptidase_id	peptidase_family	peptidase_hit	peptidase_RBH	peptidase_identity	peptidase_bitScore	peptidase_eVal	pfam_hits
Cytophaga_hutchinsonii_ATCC_33406	NC_008255.1	1	3	110	1	E														
Cytophaga_hutchinsonii_ATCC_33406	NC_008255.1	2	107	850	1	A	K01104	chu:CHU_0001 capsular polysaccharide biosynthesis protein; K01104	True	0.966	460.0	1.712000e-144								Capsular polysaccharide synthesis, CpsB/CapB [PF19567.1]
Cytophaga_hutchinsonii_ATCC_33406	NC_008255.1	3	847	1815	1	A		chu:CHU_0002 nucleoside-diphosphate-sugar epimerase	True	0.958	650.0	4.369000e-209								3-beta hydroxysteroid dehydrogenase/isonitrate reductase family [PF01073.21]; dependent epimerase/dehydratase family [PF01370.23]; mannose 4,6 dehydratase [PF16363.7]; Male sex protein [PF07993.14]; substrate binding domain [PF04321.19]; NmrA family [PF05368.15]; NAD(P)H-binding protein [PF13460.8]; Polysaccharide biosynthesis protein [PF02719.17]; short-chain dehydrogenase [PF00271.33]
Cytophaga_hutchinsonii_ATCC_33406	NC_008255.1	4	1955	3355	1	A		chu:CHU_0003 TPR repeat containing protein	True	0.959	941.0	5.175000e-307								Tetratricopeptide repeat [PF13429.8]; MaltT-like region [PF17874.3]; repeat [PF13414.8]; Tetratricopeptide repeat [PF13424.8]; Tetratricopeptide repeat [PF13432.8]; Chaperone Arf1p-binding protein [PF09295.12]; RNPII C-terminal domain [PF18768.3]; Tetratricopeptide repeat [PF14559.8]

Cytophaga_hutchinsonii_ATCC_33406	NC_008255.1	5	3357	4127	1	A	K08097	chu:CHU_0004 phosphosulfolactate synthase; K08097	True	0.97	525.0	4.879000e-167								(2R)-phospho-3-sulfolactate synthase (ComA) [PF02679.17]
Cytophaga_hutchinsonii_ATCC_33406	NC_008255.1	6	413					chu:CHU_0005 dedA; integral												SNARE associated domain protein [PF09335.13]
Cytophaga_hutchinsonii_ATCC_33406	NC_008255.1	7	487																	Shikimate dehydrogenase substrate binding domain [PF08501.13]
Cytophaga_hutchinsonii_ATCC_33406	NC_008255.1	8	569																	Putative OmpA-OmpC porin family [PF16967.1]
Cytophaga_hutchinsonii_ATCC_33406	NC_008255.1	9	668																	Type III restriction endonuclease subunit [PF04855.1]; UvrB interaction domain [PF17757.3]; Ultra-violet resistance protein B [PF12344.10]; Helicase conserved C-terminal domain [PF00271.33]

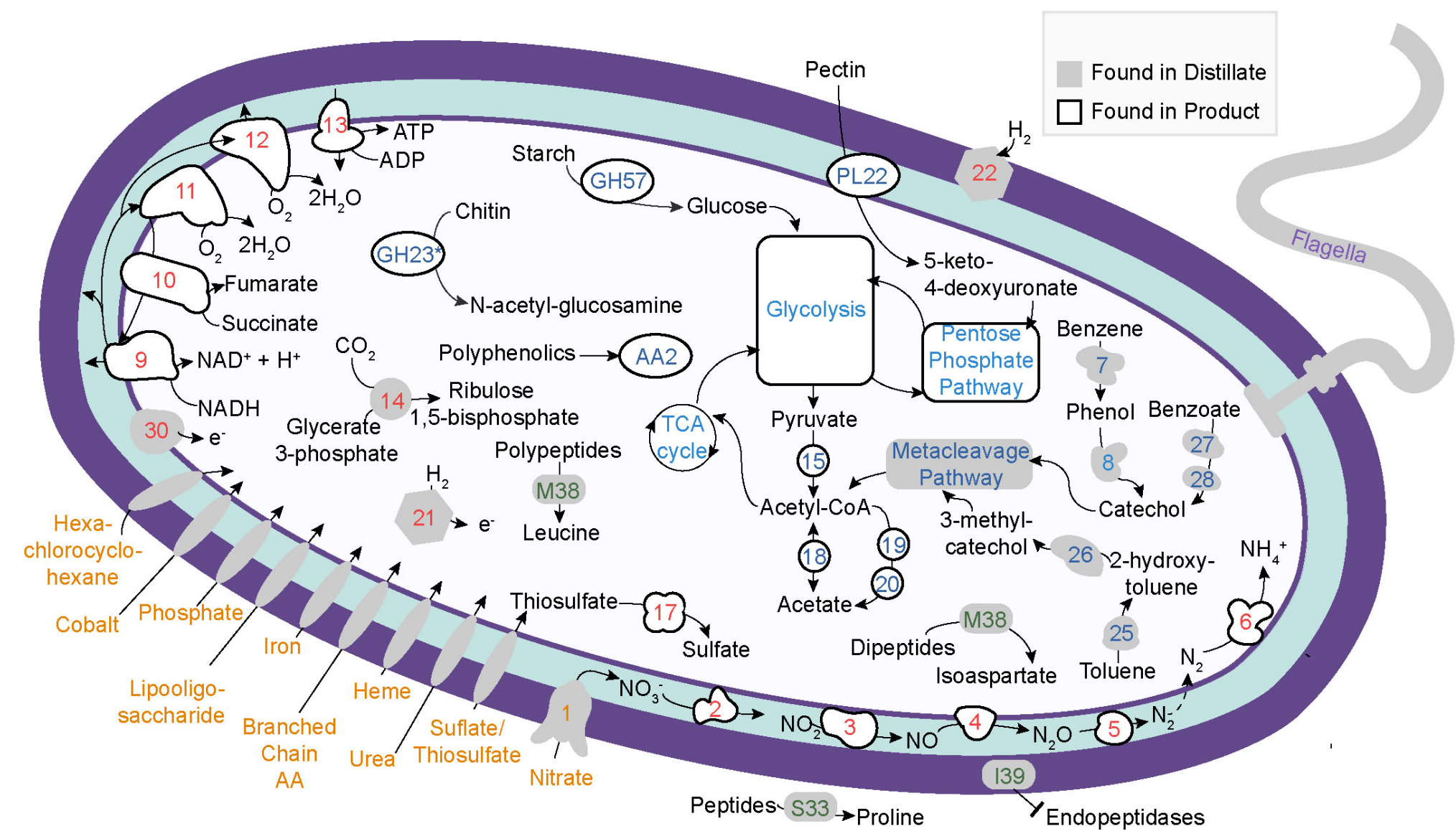
Cytophaga_hutchinsonii_ATCC_33406	NC_008255.1	10	892																	Exonuclease [PF00951.1]
Cytophaga_hutchinsonii_ATCC_33406	NC_008255.1	11	954																	
Cytophaga_hutchinsonii_ATCC_33406	NC_008255.1	12	994																	Fatty acid hydroxylase superfamily [PF04111.1]

Cytophaga_hutchinsonii_ATCC_33406	NC_008255.1	13	10631	11770	-1	A	K00754	chu:CHU_0012 a-glycosyltransferase-related protein, glycosyltransferase family 4 protein; K00754	True	0.957	764.0	1.770000e-247								Glycosyl transferase family 1 [PF00534.22]; Glycosyl transferases group 1 [PF13692.8]; Glycosyltransferase family [PF13439.8]; Glycosyl transferase 4-like [PF13477.8]
-----------------------------------	-------------	----	-------	-------	----	---	--------	--	------	-------	-------	---------------	--	--	--	--	--	--	--	--

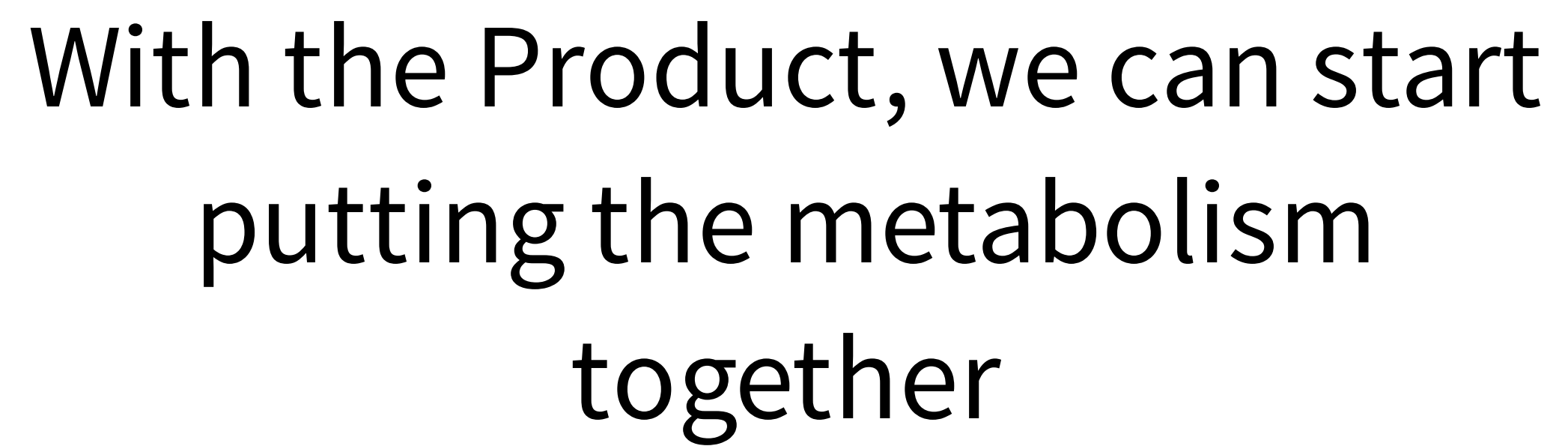


Raw contains all information

number of scaffolds	5S rRNA	16S rRNA	23S rRNA	tRNA count
1	3 present	3 present	3 present	40
1		4 present	4 present	64
1	3 present	3 present	3 present	53
1	3 present	3 present	3 present	52
1	2 present	2 present	2 present	49
1	3 present	3 present	3 present	60
1		Methylobacter_whittenburyi_NZ_JQNS01000005.1 (1, 907)		15
1	Nitrosarchaeum_koreense_MY1_NZ_AFPU01000001.1 (215335, 215447)	Nitrosarchaeum_koreense_MY1_NZ_AFPU01000001.1 (932095, 933559)	Nitrosarchaeum_koreense_MY1_NZ_AFPU01000001.1 (928963, 931928)	45
1	Nitrospira_defluvii_NC_014355.1 (1389814, 1389923)	Nitrospira_defluvii_NC_014355.1 (1384632, 1386168)	Nitrospira_defluvii_NC_014355.1 (1386680, 1389681)	48
1				20
1		Prometheoarchaeum_syntrophicum_MK-D1_CP042905.1_Anaerobic_archaeon_MK-D1_chromosome_complete_genome (724530, 726016)	Prometheoarchaeum_syntrophicum_MK-D1_CP042905.1_Anaerobic_archaeon_MK-D1_chromosome_complete_genome (656095, 659004)	45
1	8 present	7 present	7 present	76
1	6 present	6 present	6 present	71
1		2 present	2 present	41
1	2 present	2 present	2 present	52



Distillate contains more information, but less than the Raw



CONCLUSION

The future of Science is bright, and DRAM will evolve to meet that future

FUNDING



U.S. DEPARTMENT OF
ENERGY

Office of
Science



National Institutes
of Health



JOINT GENOME INSTITUTE



THE END