Supplementary Material

Untargeted Metabolomics Exploration of the Growth Stage-dependent Chemical Space of the Sclareol-converting Biocatalyst *Hyphozyma roseonigra*

Efficient N. Ncube[1], Lungile Sitole[1], Paul A. Steenkamp[1], Lucia H. Steenkamp[2], Ian A. Dubery[1*]

[1]*Department of Biochemistry, University of Johannesburg, P.O. Box 524, Auckland Park, 2006, South Africa; idubery@uj.ac.za*
[2]*Chemicals Cluster, Council for Scientific and Industrial Research (CSIR), P.O. Box 395, Pretoria 0001, South Africa*

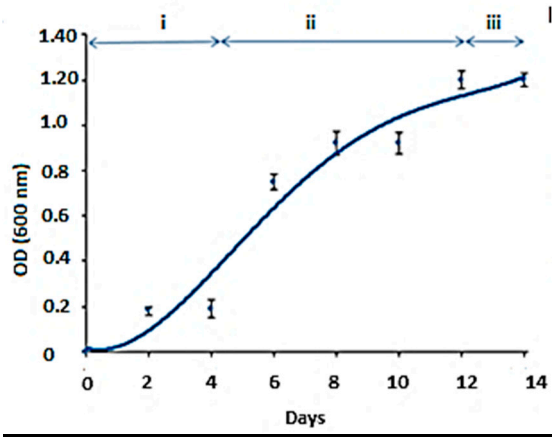

Figure S1. Assessment of the growth of the *Hyphozyma roseonigra* in batch culture. A representative growth curve in potato dextrose broth (24 g/L) over a 14 d period. Shown are the (i) early adaptation stage, (ii) logarithmic stage and (iii) stationary stage. The initial OD_{600} at inoculation was 0.015. Error bars indicate variability based on standard deviation (sample population).

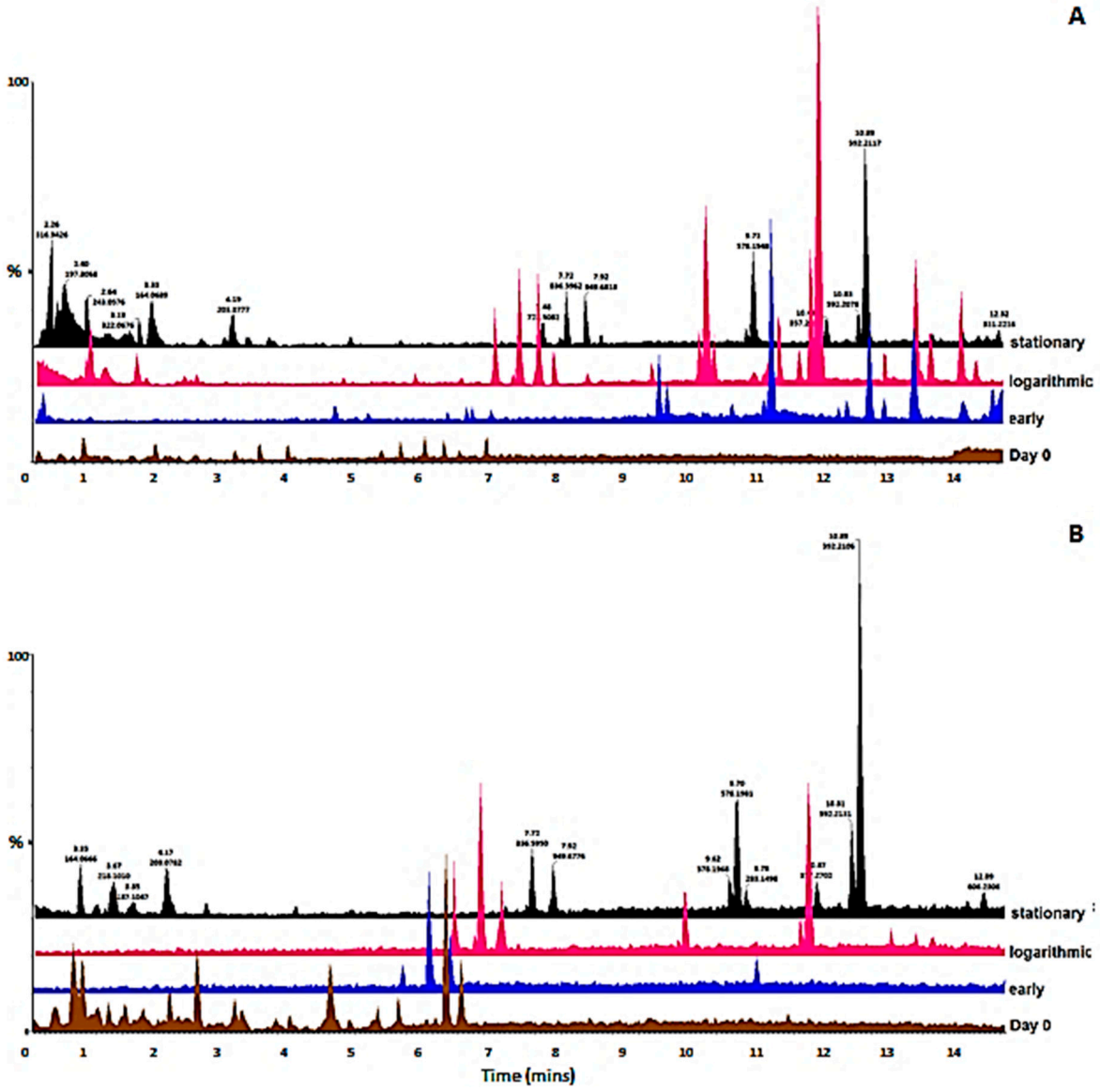

Figure S2. Ultra-high performance liquid chromatography (UHPLC) separation with high definition mass spectrometry (MS) detection of extracellular methanolic- (A) and acetonic extracts (B) of *H. roseonigra*. The cells were harvested from selected days over a 14 d (0, 2, 6, 14 d) period representing the different growth/developmental stages, *i.e.* early, log and stationary. The representative base peak intensity (BPI) chromatograms acquired in negative electrospray ionization (ESI–) mode show evident differences in peak intensities across the different days.

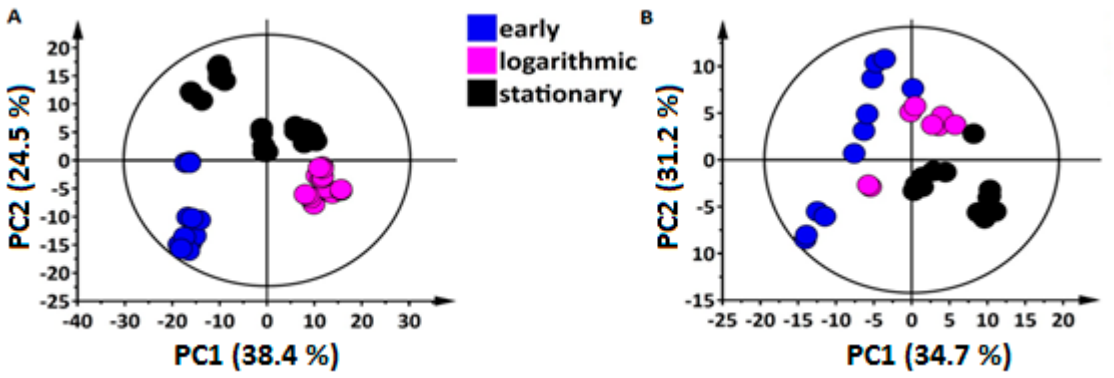

Figure S3. The PCA models of (A) methanolic- and (B) acetonic <u>extracellular extracts</u> of *H. roseonigra* analyzed using UHPLC–MS in <u>ESI(–) mode</u>. The scores plots, constructed from the first two components show clustering according to different developmental stages represented as early-, log- and stationary stage. The *Pareto*-scaled 2D scores plots (A): with PC1 and PC2 showing 38.4 % and 24.5 % respectively, explaining a total variation 62.9 %, and (B): with PC1 and PC2 showing 34.7 % and 31.2 % respectively, explaining a total variation 65.9 %, shows clear separation of the stages. The ellipse indicates Hotelling's T2 at 95% confidence interval.

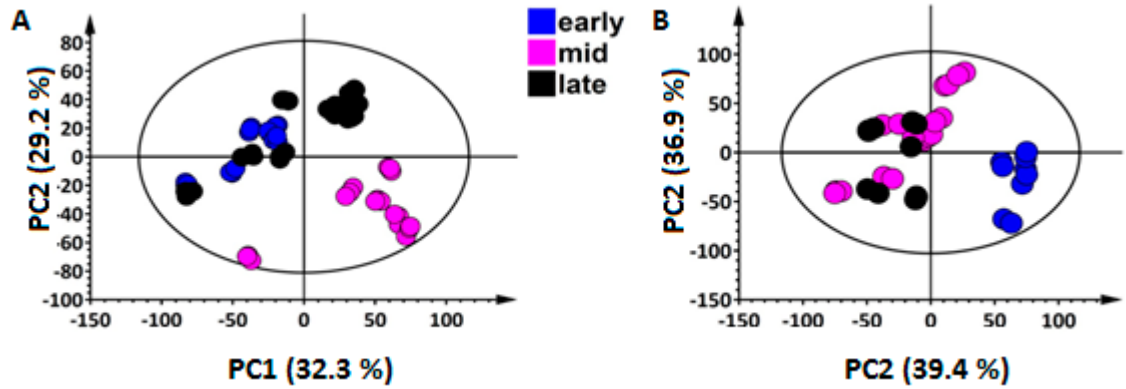

Figure S4. PCA models of (A) methanolic- and (B) acetonic <u>intracellular extracts</u> of *H. roseonigra* analyzed using LC–MS in <u>ESI(+) mode</u>. The scores plots, constructed from the first two components show clustering according to different developmental stages represented as early-, log- and stationary stage. The *Pareto*-scaled 2D scores plots (A): with PC1 and PC2 showing 32.3 % and 29.2 % respectively, explaining a total variation 61.5 %, and (B): with PC1 and PC2 showing 39.4 % and 36.9 % respectively, explaining a total variation 76.3 %, shows clear separation of the stages. The ellipse indicates Hotelling's T2 at 95% confidence interval.

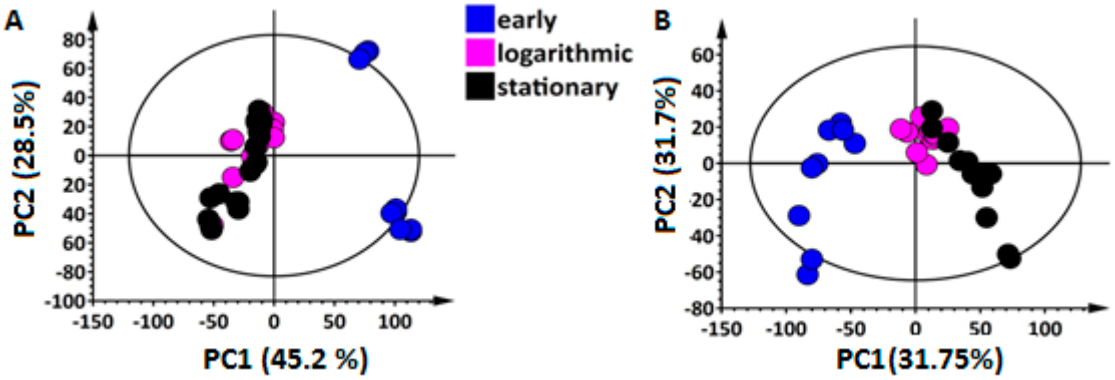

Figure S5. PCA models of (A) methanolic- and (B) acetonic <u>extracellular extracts</u> of *H. roseonigra* analyzed using LC–MS in <u>ESI(+) mode</u>. The scores plots, constructed from the first two components show clustering according to different developmental stages represented as early-, log- and stationary stage. The *Pareto*-scaled 2D scores plots (A): with PC1 and PC2 showing 45.2 % and 28.5 % respectively, explaining a total variation 73.6 %, and (B): with PC1 and PC2 showing 31.8 % and 31.7 % respectively, explaining a total variation 63.5 %, shows clear separation of the stages. The ellipse indicates Hotelling's T2 at 95% confidence interval.

Table S1. List of annotated features (metabolites) present in acetonic and methanolic extracts of *Hyphozyma roseonigra* grown in batch culture and annotated from UHPLC–MS data. (*Additional experimental detail in support of the annotations in Table 1*).

Metabolite annotation		Molecular formula	Rt (min)	Molecular weight	*m/z*	Diagnostic fragments	Adduct	Ion	Biological role	Intra-cellular	Extra-cellular
1	5-Aminoimidazole ribonucleotide [a,b]	$C_8H_{14}N_3O_7P$	2.06	295.19	378.25	216.11, 198.09, 196.69	ACN	$[M+2ACN+H]^+$	Intermediate		✓
2	Pantothenic acid [a,b]	$C_9H_{17}NO_5$	2.17	219.23	220.24	202.18	H	$[M+H]^+$	Vitamin		✓
3	Thymidine diphosphate [a,b]	$C_{10}H_{16}N_2O_{12}P_2$	2.25	402.19	401.08	383.01, 274.92, 158.92	-	$[M-H]^-$	Nucleotide	✓	
4	Glycyl-leucine[b]	$C_8H_{16}N_2O_3$	2.44	188.22	211.10	171.07	Na	$[M+Na]^+$	Dipeptide		✓
5	Xanthosine-5- [a,b] phosphate	$C_{10}H_{13}N_4O_9P$	2.59	364.21	363.21	211.00, 151.03	-	$[M-H]^-$	Purine	✓	
6	N6, N6-Dimethyladenosine [b]	$C_{12}H_{17}N_5O_4$	2.67	295.29	296.13	-	H	$[M+H]^+$	Ribonucleoside		✓
7	Maculosin[b]	$C_{14}H_{16}N_2O_3$	3.31	260.29	261.12	-	H	$[M+H]^+$	Dipeptide	✓	
8	2-Deoxyribose-1-phosphate [a,b]	$C_5H_{11}O_7P$	5.92	214.11	215.11	197.02, 179.02, 117.06	H	$[M+H]^+$	Intermediate		✓
9	S-Adenosyl-homocysteine [a,b]	$C_{14}H_{20}N_6O_5S$	7.11	384.41	385.14	242.87, 192.29, 112.12	H	$[M+H]^+$	Methyl donor	✓	
10	18-Acetoxy-1-alpha-25-dihydroxy vit. D3 [b]	$C_{29}H_{46}O_5$	8.71	474.67	473.33	-	-	$[M-H]^-$	Vitamin	✓	
11	5-Methyl-tetrahydrofolate [a,b]	$C_{20}H_{25}N_7O_6$	10.94	459.46	460.27	442.21, 414.22, 194.10	H	$[M+H]^+$	DNA synthesis	✓	
12	Reynosin [b]	$C_{15}H_{20}O_3$	14.08	248.32	247.32	184.09, 166.08	-	$[M-H]^-$	Diterpene		✓

13	Broussonin C [b]	$C_{20}H_{24}O_3$	15.61	312.40	376.15	-	ACN+Na	$[M+ACN+Na]^+$	Polyketide phenol		✓
14	Deoxyuridine [a,b]	$C_9H_{12}N_2O_5$	2.6	228.20	227.10	209.06, 184.06, 182.05	-	$[M-H]^-$	Ribonucleoside	✓	
15	Biotin [b]	$C_{10}H_{16}N_2O_3S$	5.49	244.31	245.05	227.08	H	$[M+H]^+$	Vitamin	✓	
16	UDP-glucuronate [a,b]	$C_{15}H_{22}N_2O_{18}P_2$	8.19	580.29	579.20	563.05	-	$[M-H]^-$	Sugar		✓
17	Xanthosine-5-phosphate [a,b]	$C_{10}H_{13}N_4O_9P$	8.28	364.21	363.09	211.01, 151.03	-	$[M-H]^-$	Purine	✓	

a,b denotes databases used for annotation, where a = the Yeast Metabolome Database (YMDB, http://www.ymdb.ca/ and b = Taverna workflows (in-house). ACN = acetonitrile.

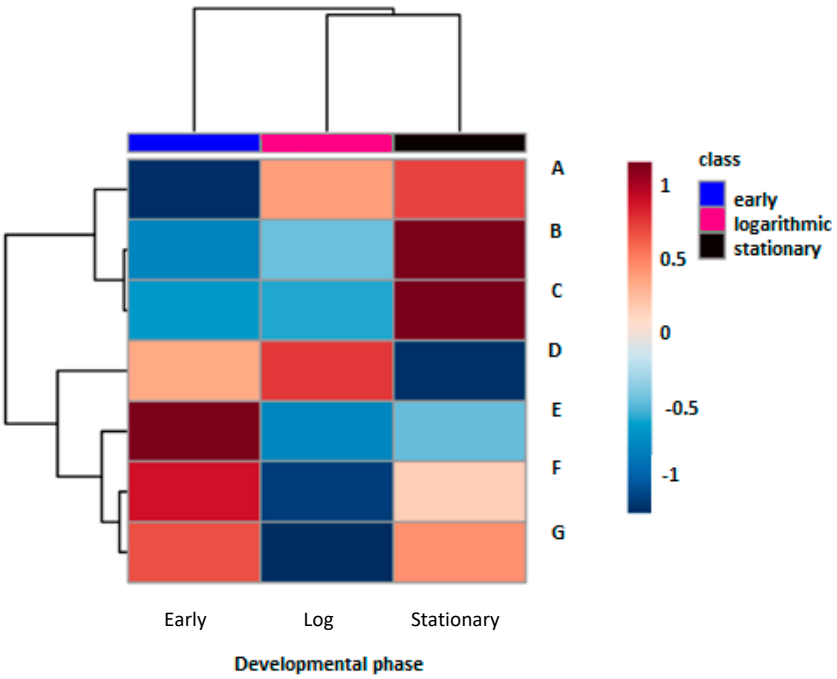

Figure S6. Heat map showing selected annotated metabolites of acetonic and methanolic extracts of *H. roseonigra* across the early adaptation-, stationary- and log stages. The selected metabolites representing each stage include 5-Aminoimidazole **(A),** Glycyl-leucine **(B),** N6,N6-Dimethyladenosine **(C),** 2-Deoxyribose-1-phosphate **(D),** Reynosin **(E),** S-Adenosyl-L-homo-cysteine **(F)** and 5-Methyl-tetrahydrofolate **(G)**. The map was generated using relative peak intensities and scored from -1 to 1, representing low to high relative concentrations, respectively. The data was normalized using *Pareto* scaling and transformed using log transformation.

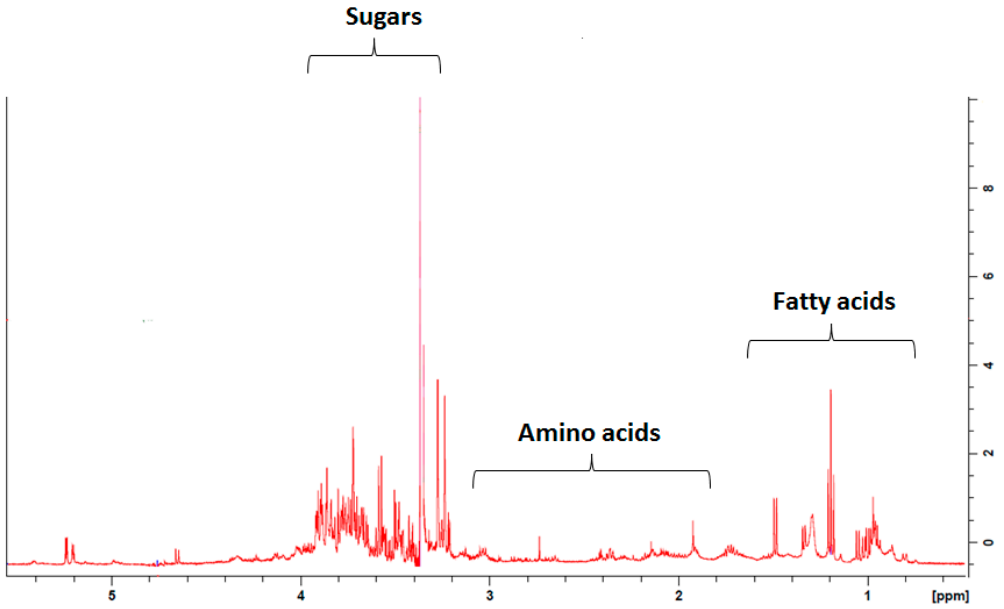

Figure S7. ^{1}H NMR spectrum and class annotation of metabolites present in methanolic extracts of *H. roseonigra*. Shown is a representative spectrum with annotated metabolites categorized to three different major classes, *i.e.* sugars, amino acids and fatty acids in that order from low to high regions, respectively. The spectra are characterized by three main regions: a low-field region between 4.0 and 3.2 ppm with intense signals due to anomeric protons of sugar units, a mid-field region between 3.1 and 1.8 with signals associated with amino acids and a high-field region between 1.6 and 0.8 ppm with strong signals due to aliphatic protons of fatty acids.

Table S2. Annotation of metabolites in methanolic extracts of *Hyphozyma roseonigra* analyzed by ^{1}H NMR. (*Additional experimental detail in support of the annotations provided in Table 2*).

Assigned Metabolite number (#)	Metabolite Annotation	Chemical shifts (ppm)	Multiplicity	Number of Hs	Database/ Reference
18	Aspartate	2.66 **2.67** 2.68	Triplet	1	YMDB
19	Glutamine	2.21 2.15 **2.11** 2.08 2.07 2.01	Multiplet	2	YMDB
20	Valine	**1.04** 1.02	Doublet	3	YMDB
21	Threonine	**1.32** 1.31	Doublet	3	YMDB
22	Alanine	1.4**6** 1.47 ; 3.76 3.77	Doublet	3 , 1	YMDB, [1]
23	Glutamate	1.99 2.00 2.04 2.05 **2.09** 2.11 2.16 2.17 2.20 2.23 2.24 2.25 2.28	Multiplet	2	YMDB
24	Isoleucine	0.94 **0.93** 0.91 0.90	Triplet	3	YMDB
25	Leucine	1.64 1.58 1.70 **1.72** 1.78	Multiplet	3	YMDB
26	Cysteine	3.01 **3.02** 3.03	Multiplet	2	YMDB
27	Ornithine	**1.80**	Singlet	2	YMDB
28	Serine	3.81 3.83 **3.84** 3.87 3.89	Quintet	1	YMDB
29	N-Acetyl aspartate	**2.44 2.49 2.51 2.55**	Doublet of doublet	1	HMDB
30	Palmitic acid, methyl ester	**0.771**	Singlet	3	[2]
31	Hexadecyl octanoate	**0.79** 0.80	Doublet	3	[2]
32	Nonadecanoic acid	0.75 **0.77**	Doublet	3	[2]
33	Linoleic acid	0.93 0.97 **0.98** 0.99	Doublet of doublet	6	[2]
34	2-Hydroxy-isobutyrate	**1.35**	-	-	[3]
35	Galactose-1-phosphate	**3.72** 3.74	Quintet	2	YMDB
36	Glucose	3.21 3.25 **3.38** 3.42 3.43 3.3.45 3.47 3.50	Multiplet	1	YMDB
37	Galactitol	3.62 3.68 **3.73** 3.75 3.78	Doublet	1	YMDB
38	Myoinositol	3.84 **3.87** 3.88 3.89 3.90	Multiplet	2	YMDB
39	Lactate	4.08 **4.10** 4.11 4.13	Quintet	1	YMDB

40	Citrate	2.51 **2.54**	Doublet	2	YMDB
41	Succinate	**2.39**	Singlet	4	YMDB
42	Betaine	3.19 3.20 3.23 **3.25** 3.28 3.29 3.31	Singlet	9	YMDB
43	Ethanol	1.16 **1.17** 1.18	Triplet	3	YMDB
44	Glutathione	3.74 **3.77** 3.83	Quintet	3	YMDB
45	Coenzyme A	**0.73 0.88**	Singlet	3	HMDB

Chemical shifts are referenced to trimethylsilylpropanoic acid (TSP). The listed numbers indicate the range of the shifts to identify the metabolite, the number in bold is the centre of the range and is of more relevance in determining the level of confidence for the particular annotation.

YMDB = Yeast Metabolome Database (http://www.ymdb.ca); HMDB = Human Metabolome Database (https://hmdb.ca).

[1] = Airoldi, C.; Tripodi, F.; Guzzi, C.; Nicastroab, R.; Coccetti, P. NMR analysis of budding yeast metabolomics: a rapid method for sample preparation. *Mol. Biosyst*. 2015, 11, 379. doi: 10.1039/c4mb00452c.

[2] = Kostidis, S.; Addie, R.D.; Morreau, H.; Mayboroda, O.A.; Giera, M. Quantitative NMR analysis of intra- and extracellular metabolism of mammalian cells: A tutorial. *Anal. Chim. Acta* **2017**, *980*, 1–24. https://doi.org/10.1016/j.aca.2017.05.011.

[3] = Marante Toledo, F.J.; Mioso, R.; Barrera, J.B.; González González, J.E.; Santana Rodríguez, J.J.; De Laguna, I.H.B. Structural characterization and metabolite profiling of the facultative marine fungus *Paecilomyces variotii*. *Ann. Microbiol*. **2012**, *6*, 1601–1607. https://doi.org/10.1007/s13213-011-0416-1

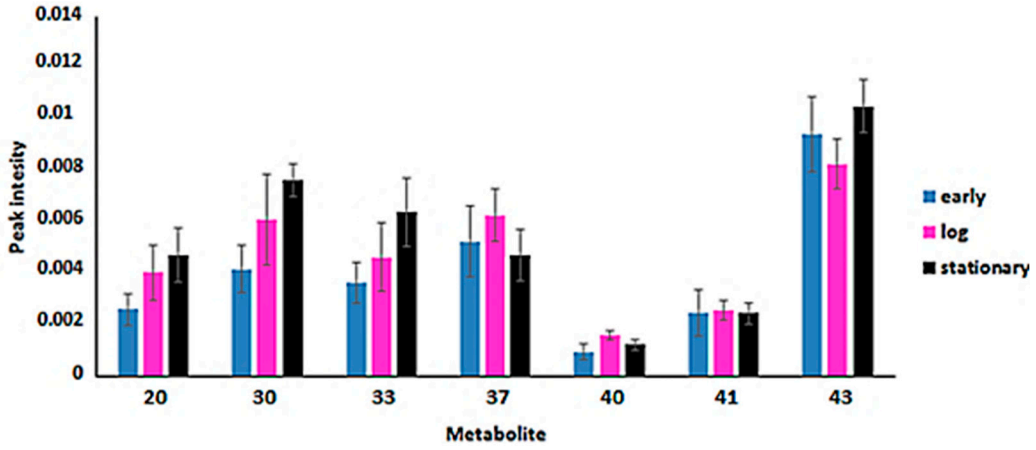

Figure S8. Semi-quantitative presentation of intracellular metabolites present in methanolic extracts of *H. roseonigra* (based on annotation from ^{1}H NMR spectra (500 Hz) data). The cells were harvested from selected days (0, 2, 6, 14 d) over a 14 d period representing the different developmental stages, *i.e.* (**E**) early-, (**L**) logarithmic- and (**S**) stationary. The representative spectra show evident differences in peak intensities across the different days, where the intensities decrease or increase for some peaks. The resonance from the internal standard (TSP) is at 0.00 ppm. (**20**) = valine, (**30**) = linoleic acid (**33**) = palmitic acid methyl ester (**37**) = galactitol, (**40**) = citrate, (**41**) = succinate and (**43**) = ethanol.

Table S3. Significant metabolic pathways determined to be active in *Hyphozyma roseonigra*, inferred from Metabolomics Pathway Analysis (MetPA) and listed according to the *p*-value. The eight pathways with an impact value > 0.12 are indicated in black.

Rank	Pathway name	*p*-value	Impact
1	Alanine, aspartate and glutamate metabolism	0.011	0.320
2	Aminoacyl-tRNA biosynthesis	0.038	0
3	Pantothenate and CoA biosynthesis	0.047	0.188
4	β-Alanine metabolism	0.074	0
5	Glyoxylate and dicarboxylate metabolism	0.090	0.015
6	Monobactam biosynthesis	0.157	0
7	Arginine biosynthesis	0.172	0
8	Valine, leucine and isoleucine degradation	0.172	0
9	Nitrogen metabolism	0.193	0
10	Valine, leucine and isoleucine biosynthesis	0.203	0
11	Citrate cycle (TCA cycle)	0.203	0.123
12	Glycolysis/Gluconeogenesis	0.266	7.90E-04
13	Cyanoamino acid metabolism	0.291	0
14	One carbon pool by folate	0.291	0
15	Glycine, serine and threonine metabolism	0.392	0.104
16	Nicotinate and nicotinamide metabolism	0.404	0
17	Pyrimidine metabolism	0.422	0.018
18	Sulfur metabolism	0.429	0.058
19	Biotin metabolism	0.429	0.127
20	Butanoate metabolism	0.453	0
21	Fructose and mannose metabolism	0.453	0.027
22	Glycerolipid metabolism	0.453	0.129
23	Tyrosine metabolism	0.476	0
24	Purine metabolism	0.490	0.121
25	Lysine biosynthesis	0.499	0
26	Galactose metabolism	0.520	0
27	Cysteine and methionine metabolism	0.522	0
28	Propanoate metabolism	0.560	0
29	Inositol phosphate metabolism	0.614	0.116
30	Pyruvate metabolism	0.631	0
31	Biosynthesis of unsaturated fatty acids	0.631	0
32	Amino sugar and nucleotide sugar metabolism	0.647	0
33	Phosphatidylinositol signaling system	0.677	0.045
34	Glutathione metabolism	0.677	0.397
35	Fatty acid degradation	0.729	0.308

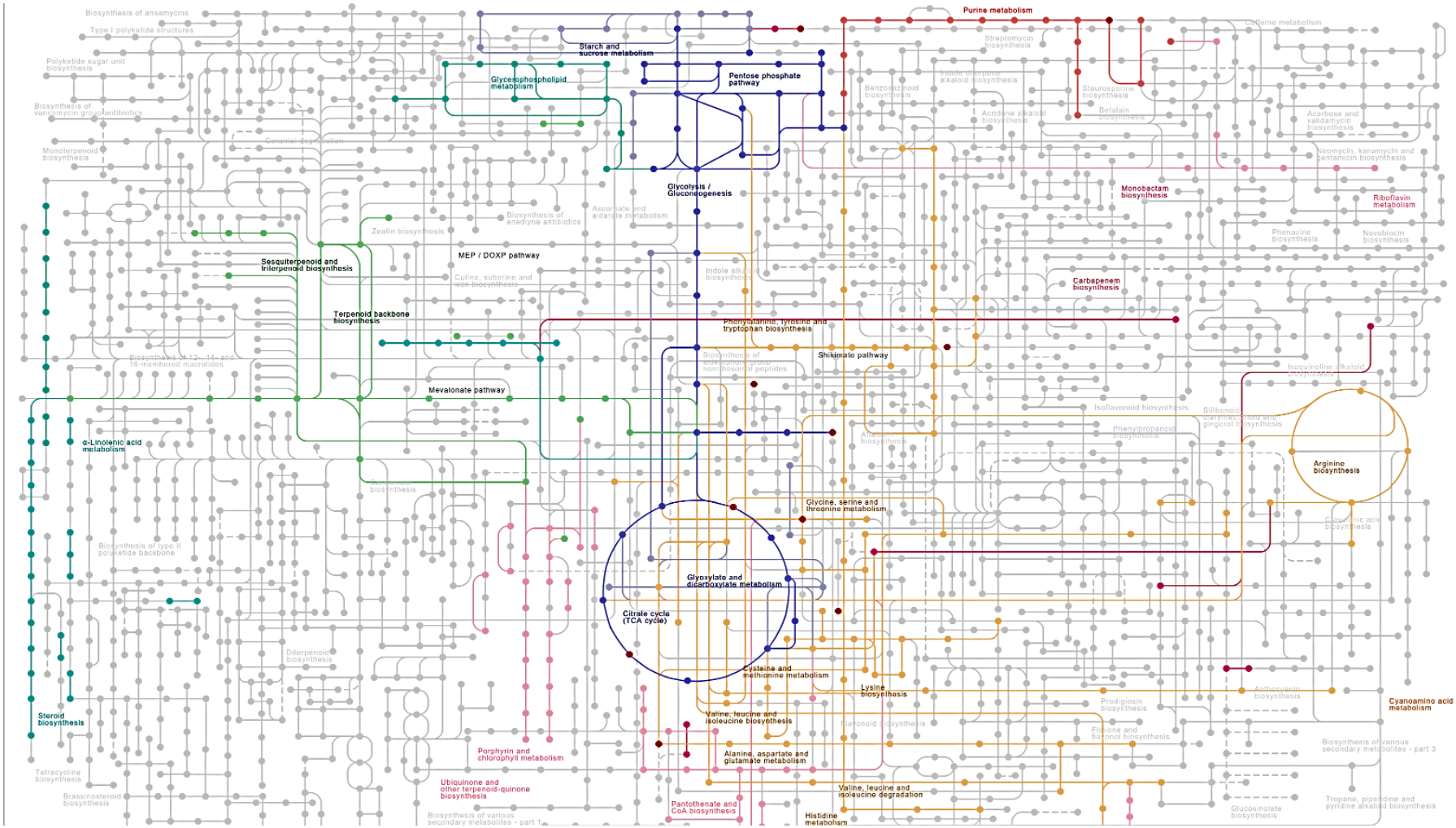

Figure S9. Biochemical roadmap: a network correlation analysis of metabolic pathways active in *Hyphozyma roseonigra* when grown in potato dextrose broth medium in batch culture. Infographic showing major metabolic pathways, *i.e.* carbohydrate metabolism, lipid metabolism, amino acid metabolism, metabolism of cofactors and vitamins, biosynthesis of terpenoids and polyketides and biosynthesis of other secondary metabolites present in *H. roseonigra* over the growth period.