bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1 ***Areca catechu*-(Betel-nut)-induced whole transcriptome changes associated with**
2 **diabetes, obesity and metabolic syndrome in a human monocyte cell line**
3
4 Short title: Betel-nut induced whole transcriptome changes
5
6
7 Shirleny Cardoso[1¶] , B. William Ogunkolade[1¶], Rob Lowe[2], Emanuel Savage[3], Charles A
8 Mein[3], Barbara J Boucher[1], Graham A Hitman[1*]
9
10
11 [1]Centre for Genomics and Child Health, Blizard Institute, Barts and the London School of
12 Medicine and Dentistry, Queen Mary University of London, United Kingdom
13
14 [2]Omnigen Biodata Ltd, Cambridge, United Kingdom
15
16 [3]Barts and The London Genome Centre, Blizard Institute, Queen Mary University of London,
17 United Kingdom
18
19 [*] Corresponding author
20 Email: g.a.hitman@qmul.ac.uk
21
22 [¶]These authors contributed equally to the work
23
24

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

Betel-nut consumption is the fourth most common addictive habit globally and there is good evidence to link it with obesity, type 2 diabetes and the metabolic syndrome. We adopted a genome-wide transcriptomic approach in a human monocyte cell line incubated with arecoline and its nitrosated products to identify gene expression changes relevant to obesity, type 2 diabetes and the metabolic syndrome. The THP1 monocyte cell line was incubated separately with arecoline and 3-methylnitrosaminopropionaldehyde (MNPA) in triplicate for 24 hours and pooled cDNA indexed paired-end libraries were sequenced (Illumina NextSeq 500). After incubation with arecoline and MNPA, 15 and 39 genes respectively had significant changes in their expression ($q<0.05$, log fold change 1.5). Eighteen of those genes have reported associations with type 2 diabetes and obesity in humans; of these genes there was strong evidence to implicate *CLEC10A*, *MAPK8IP1*, *NEGR1*, *NQ01* and *INHBE*. In summary, these pilot studies have identified a large number of genes whose expression was changed significantly in human TPH1 cells following incubation with arecoline or with 3-methylnitrosaminopropionaldehyde. These findings suggest that further investigation of these genes in betel-quid chewers with obesity and/or type 2 diabetes is warranted.

Keywords

Betel-nut, type 2 diabetes, obesity, transcriptomics, RNA-sequencing, *CLEC10A*, *MAPK8IP1*, *NEGR1*, *NQ01*, *INHBE*

Abbreviations

MNPA 3-methylnitrosaminopropionaldehyde**,** MNPN 3-methylnitrosaminopropionitrile
PMA (Phorbol-12-Myrsitate-13-Acetate)

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

48 # Introduction

49 Obesity and Type 2 diabetes are reaching epidemic proportions worldwide, but particularly
50 so in South Asian communities living in the Indian-subcontinent or who have migrated to
51 other countries[1]. In the UK there is a 3 to 4 fold increase in type 2 diabetes prevalence in
52 South Asians compared to the general population; furthermore the disease manifests at a 10-
53 15 years younger age and is strongly associated with the metabolic syndrome and
54 cardiovascular disease. Apart from lifestyle, potentially reversible environmental factors
55 driving this disease are largely unknown.
56 Betel quid consumption is the fourth most common addictive habit, used by 10% of the
57 global population and very common in South Asians. The link between cancer risks
58 (oropharyngeal, oesophageal and hepatocellular) and the 'betel-chewing' habit is well
59 established[2-4]. Evidence also implicates an association between betel consumption and
60 obesity, the metabolic syndrome and type 2 diabetes. In a meta-analysis of 17 Asian studies,
61 betel quid chewing was found to be significantly associated with obesity (relative risk (RR) =
62 1.47), metabolic syndrome (RR=1.51), diabetes (RR=1.47), hypertension (RR=1.45) and
63 cardiovascular disease (cardiovascular disease: RR=1.2)[5]. Furthermore, in the Keelung
64 Community Integrated Screening programme studies from Taiwan, paternal betel-chewing
65 was associated, dose-wise, with increases in early onset metabolic syndrome in their never-
66 chewing offspring, while betel-chewing in adults increased their risks, dose-wise, of early
67 onset type 2 diabetes and cardiovascular disease [6-8]. These data in humans support earlier
68 data reported in CD1 mice, where it was found that a proportion of betel-fed adult mice
69 developed hyperglycaemia and obesity and, more remarkably, that amongst their never-betel-
70 fed offspring hyperglycaemia was detected in up to the 4th generation and that the vertical
71 transmission of hyperglycaemia was associated with paternal, but not maternal,
72 hyperglycaemia[9].

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

73 The betel quids (also known as paan) are usually contain betel (areca) nut, slaked lime and
74 sometimes tobacco wrapped in betel leaf[10]. In Taiwan tobacco is not used. Nitrosation of
75 the major arecal alkaloid, arecoline, forms 3-methylnitrosaminopropionaldehyde (MNPA),
76 and 3-methylnitrosaminopropionitrile (MNPN) and both these compounds have been
77 identified as carcinogens [11]. Many nitroso-compounds (e.g. streptozotocin) have been
78 reported as being diabetogenic, low doses causing type 2 diabetes -like diabetes, suggesting
79 that betel-chewing might be one of the aetiological factors for the increases in type 2
80 diabetes and associated metabolic disorders in South Asians[12, 13]. Mechanisms that might
81 link betel chewing with these disorders include inflammation, increases in hepatic synthesis
82 of lipids and glucose, in adipogenesis, in adipose tissue glucose uptake, reductions in
83 lipolysis and glycolysis, neurological, hepatic or intestinal effects on appetite and adverse
84 effects on vitamin D metabolism[14, 15].
85
86 In the present proof of principle study we sought to investigate possible biological links
87 between type 2 diabetes (and related disorders) and exposure to arecoline and its nitrosated
88 products in a human monocyte derived cell line (THP1) using a whole transcriptome
89 sequencing approach. THP1 was chosen due the central role that low-grade inflammation
90 plays in the underlying causes and progression of obesity (for instance adipose tissue
91 macrophages), metabolic syndrome and related disorders including both cardiovascular
92 disease and type 2 diabetes[16, 17].
93
94

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

95 # Materials and Methods

96 The THP1 (human acute monocytic leukemia derived; ATCC number TIB-202 purchased
97 from Thermofisher) cell line[18] was regularly maintained in RPMI 1640 medium containing
98 GlutaMAX™, supplemented with 10% FCS (Gibco™ Newborn Calf Serum [heat
99 inactivated], of New Zealand origin; Thermo Fisher Scientific), 5% AA (Gibco® MEM),
100 Non-Essential Amino Acids, 100 U/mL penicillin and 100 μg/mL streptomycin (Thermo
101 Fisher Scientific). Cells were grown at 37 °C in a humidified atmosphere of 5% CO2 in air,
102 and sub-passaged with fresh complete RPMI medium every three days. The cell line was
103 regularly checked to be mycoplasma free using the VenorGeM Mycoplasma detection kit
104 (Cambio Ltd, UK) according to manufacturer's instructions.
105
106
107 For chemical treatments, arecoline, MNPA and PMA (Phorbol-12-Myrsitate-13-Acetate)
108 were diluted in 100 % methanol while MNPN was diluted in 100 % ethyl acetate. $1x10^6$
109 THP1 cells were treated with either 100 ng/ml Arecoline, 2-5 ng/ml MNPA, 50 ng/ml MNPN
110 or 200 ng/ml PMA as a positive control in 6-well plates and cells were harvested after 6 h, 24
111 h and 48 h of treatment. Methanol and ethyl acetate were used as negative controls. Three
112 independent experiments were performed for each exposure. All chemicals were purchased
113 from Sigma-Aldrich.
114

115 ## RNA extraction and RT-qPCR for gene expression of human TNFa, IL-6
116 ## and IL-8 analysis

117 Total RNA was extracted from treated cells using QIAGEN RNeasy Kit according to
118 manufacturer's instructions. cDNA was synthesized using 1 μg of the extracted RNA with an

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

119 Oligo (dT) primer using a SuperScript® IV First-Strand Synthesis System (Thermo Fisher
120 Scientific) as follows: primer annealing at 65 °C for 5 min; RNA reverse transcriptase at 50
121 °C for 1 h 10 min and at 70 °C for 15 min. The cDNA was used as a template to determine
122 the expression level of human TNFa, IL-6 ,IL-8 and 18S in (arecoline, MNPA or MNPN)
123 treated/untreated THP1 cells. The RT-PCR was performed on StepOne™ Real-Time PCR
124 System thermal cycler (Applied Biosystems™). Each PCR reaction consisted of 2 µl of
125 cDNA, 2X SYBR® Green JumpStart™ Taq ReadyMix™ (Sigma-Aldrich), 0.2 µM of each
126 forward and reverse primers (table S1). qPCR reaction conditions were: cDNA denaturation
127 at 95 °C for 5 min, cDNA amplification at 95 °C for 15 sec, primer annealing at 62 °C for 1
128 min for 45 cycles, then melt curves were obtained at 95 °C for 15 s, 60 °C for 1 min and a
129 final step at 95 °C for 15 s. All target genes were normalised to *18S RNA* using the
130 standard *ΔΔ*Ct method. Results were analysed with Thermo Fisher StepOne software v2.3.
131 Each experiment was performed in duplicate and fold change expression level was reported
132 relative to 18S level.
133
134
135 **RNA-sequencing and bioinformatics**
136 Fragmented cDNA Sequencing libraries were generated from 100ng of Total RNA using
137 NEBNext Ultra II with polyA isolation module (Illumina, San Diego, California, USA)
138 according to manufacturer's protocol. cDNA quantity and quality were evaluated using the
139 Qubit dsDNA HS assay kit (Thermo Fisher Scientific, Erembodegem-Aalst, Belgium). Size
140 distribution of our library was determined using an Agilent 2100 Bioanalyser. Pooled indexed
141 paired-end libraries were sequenced on the Illumina NextSeq 500 (Illumina, USA) using the
142 manufacturer's instructions. Sequencing was performed at Queen Mary University of
143 London Genome Centre core facility in the Blizard Institute London.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

144

145 Sequenced reads were mapped using Kallisto[19] with default settings. Mean insert sizes and
146 standard deviations were provided as input. Analysis of differential gene expression was
147 performed using sleuth, applying a generalized linear model and utilising bootstraps on reads
148 to estimate inferential variance. Genome-wide corrected p-values were calculated using the
149 Bonferroni multiple testing adjustment procedure.

150

151 Functional annotation as well as pathway enrichment analyses were performed using
152 DAVID, Reactome and Metascape ((https://metascape.org/gp/index.html#/main/step1).

153

154 **Candidacy of genes identified were assessed by look-ups in:**

155 1. GeneCards: The human genome database (https://www.genecards.org/) to check for
156 alias's, gene summary (Entrez, Genecards and UniProtKB/Swiss-Prot) and mRNA
157 expression in normal human tissues (GTEx, Illumina, BioGPS)
158 2. Type 2 diabetes knowledge portal (http://www.type2diabetesgenetics.org/). Genes
159 considered were only those with strong evidence for signal defined by either at least one
160 variant within the coding sequence ± 100kb that is associated with at least one phenotype
161 with a p value<5e-8 identified by a genome-wide association scan (GWAS), or at least one
162 known variant with a missense or protein truncating mutation in the encoded protein that is
163 associated with at least one phenotype with a p value <5e-6.
164 3. PubMed (GAH and BJB independently) by searching biology of the identified gene
165 and biological relevance to obesity, type 2 diabetes or the metabolic syndrome before
166 collation of results.

167

168

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

169 # Results

170

171 ## Baseline experiments in triplicate

172 Whilst an inflammatory response was seen after incubation with either arecoline, MNPA or
173 PMA, no response was found at 6, 24 or 48 hours with MNPN – it was therefore decided that
174 we could not proceed further with investigation of the last of these compounds.

175

176 After 24 hours incubation with PMA increased expression was found for TNFA, IL6 and IL8
177 (mean fold change 2.5, 20.6 and 23.6 units respectively); arecoline incubation increased IL6
178 expression (mean fold change 2.8) and MNPA incubation also increased IL6 expression
179 (mean fold change 2.8). PMA increased expression at 48 hours for IL6 and IL8. We therefore
180 decided to proceed with the incubations at 24 hours and whole transcriptome expression
181 experiments were run in triplicate.

182

183 ## Whole transcriptome analysis

184 ### Incubation with arecoline

185 275 gene hits were identified with a $q<0.05$ (table S2) reducing to 15 with a log-fold change
186 in either direction of 1.5 (table 1). Amongst the 15 genes, 4 genes have relevance to diabetes,
187 obesity and/or metabolic syndrome listed below

188 a. Insulin Like Growth Factor Binding Protein 3 (*IGFB3* non-logged fold change 0.08),
189 b. C-Type Lectin Domain Containing 10A (*CLEC10A* fold change 0.14),
190 c. Junction Plakoglobin (*JUP* fold change 0.21)
191 d. Mitogen-Activated Protein Kinase 8 Interacting Protein 1 (*MAPK8IP1* fold change
192 0.21)

193

194 Table 1 Genes identified after incubation with arecoline

195

Gene Name	Gene ID	Gene Name	q-value	logfc	Gene Description
H3F3AP4	ENSG00000235655	H3F3AP4	0.0328274	-6.7682546	H3 Histone Pseudogene 6
MEP1A	ENSG00000112818	MEP1A	0.0432902	-2.1619323	Meprin A Subunit Alpha
KBTBD11-OT1	ENSG00000283239	KBTBD11-OT1	0.0461988	-3.7981785	KBTBD11 Overlapping Transcript 1
IGFBP3	ENSG00000146674	IGFBP3	0.0432902	-2.2897729	Insulin Like Growth Factor Binding Protein 3
AC097372.1	ENSG00000250673	AC097372.1	0.0435549	-2.1772391	Reeler Domain Containing 1
PTCRA	ENSG00000171611	PTCRA	0.0432902	-1.9842384	Pre T Cell Antigen Receptor Alpha
IL3RA	ENSG00000185291	IL3RA	0.0420978	-1.9557171	Interleukin 3 Receptor Subunit Alpha
CLEC10A	ENSG00000132514	CLEC10A	0.0332128	-1.8776835	C-Type Lectin Domain Containing 10A
ACKR3	ENSG00000144476	ACKR3	0.0432902	-1.7457949	Atypical Chemokine Receptor 3
JUP	ENSG00000173801	JUP	0.0354724	-1.5030002	Junction Plakoglobin
MAPK8IP1	ENSG00000121653	MAPK8IP1	0.0461988	-1.5085173	Mitogen-Activated Protein Kinase 8 Interacting Protein 1*
MATK	ENSG00000007264	MATK	0.0295919	-1.5132972	Megakaryocyte-Associated Tyrosine Kinase
TREML3P	ENSG00000184106	TREML3P	0.0328274	1.93092581	Triggering Receptor Expressed On Myeloid Cells Like 3, Pseudogene
AC245036.5	ENSG00000269271	AC245036.5	0.0432902	2.98837483	RNA gene; lncRNA
AC113189.4	ENSG00000272884	AC113189.4	0.0466794	2.75155456	RNA gene; lncRNA

196

197 Listed genes satisfied the following criteria q<0.05 and a log-fold change of 1.5

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Pathway (Metascape) analysis of the 275 genes (figure 1) listed in the online table 2, revealed 5 significant pathways after statistical correction: myeloid cell activation involved in immune response, cellular response to thyroid hormone stimulus, responses to toxic substances and Hematopoietic cell lineage.

Figure 1. Pathway (Metascape) analysis arecoline incubation

Legend

Metascape Bar plot (P value (log10 scale)) showing Top 20 arecoline-induced, enriched functional ontology clusters (GO and KEGG terms) one per cluster.

Incubation with MNPA

359 gene hits were identified after incubation with MNPA with a q<0.05 (table S3) reducing to 39 with a log-fold change of +/-1.5 (table 2). Amongst the 39 genes, 14 have relevance to diabetes, obesity and/or metabolic syndrome listed below

a. Gliomedin (*GLDN* fold change 0.11)

b. Glutamate Receptor Interacting Protein 1 (*GRIP1* fold change 0.15)

c. Neuronal Growth Regulator 1 (*NEGR1* fold change 0.14)

d. Potassium Voltage-Gated Channel Subfamily Q Member 5 (*KCNQ5* fold change 0.21)

e. Cytotoxic And Regulatory T-Cell Molecule (*CRTAM* fold change 4.9)

f. NAD(P)H Quinone Dehydrogenase 1 (*NQO1* fold change 4.9)

g. Semaphorin 6B (*SEMA6B* fold change 4.9)

h. Inhibin Subunit Beta E (*INHBE* fold change 5.4)

i. Clusterin (*CLU* fold change 6.7)

j. Spectrin Alpha, Erythrocytic 1 (*SPTA1* fold change 9.3)

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

k. Heme Oxygenase 1 (*HMOX1* fold change 8.4)

l. Transmembrane Protein 140 (*TMEM140* fold change 10.0)

m. Sarcoglycan Gamma (*SGCG* fold change 24.8)

n. Triggering Receptor Expressed On Myeloid Cells Like 4 (*TREML4* fold change 26.7)

Table 2 Genes identified after incubation with MNPA

Gene Name	Gene ID	q-value	logfc	Gene Description
H3F3AP4	ENSG00000235655	0.019321694	-6.711052836	H3F3AP4; H3 Histone Pseudogene 6
PRKN	ENSG00000185345	0.049545404	-2.35379076	Parkin RBR E3 Ubiquitin Protein Ligase
ARSEP1	ENSG00000224060	0.036109343	-2.280202028	Arylsulfatase L Pseudogene 1
MYO7B	ENSG00000169994	0.027029548	-2.241484905	Myosin VIIB
SIGLEC6	ENSG00000105492	0.036109343	-1.968691623	Sialic Acid Binding Ig Like Lectin 6
WDR49	ENSG00000174776	0.012806465	-1.931428555	WD Repeat Domain 49
TENM3	ENSG00000218336	0.036109343	-1.91600906	Teneurin Transmembrane Protein 3
GLDN	ENSG00000186417	0.037309037	-1.898183997	Gliomedin
GRIP1	ENSG00000155974	0.043660886	-1.781556379	Glutamate Receptor Interacting Protein 1
NEGR1	ENSG00000172260	0.036189832	-1.734231847	Neuronal Growth Regulator 1
LRMDA	ENSG00000148655	0.027029548	-1.716827546	Leucine Rich Melanocyte Differentiation Associated
CCDC26	ENSG00000229140	0.04925699	-1.59660453	CCDC26 Long Non-Coding RNA
AL023693.1	ENSG00000224374	0.046572961	-1.587030076	lncRNA
KCNQ5	ENSG00000185760	0.041953451	-1.557854453	Potassium Voltage-Gated Channel Subfamily Q Member 5
AL109914.1	ENSG00000229646	0.047783786	-1.537281802	lncRNA
C2orf81	ENSG00000284308	0.04949737	-1.52240063	Chromosome 2 Open Reading Frame 81
CNTN4	ENSG00000144619	0.019201788	-1.516481444	Contactin 4
SLFN5	ENSG00000166750	0.027029548	1.537983467	Schlafen Family Member 5
CRTAM	ENSG00000109943	0.036109343	1.551523621	Cytotoxic And Regulatory T-Cell Molecule
NQO1	ENSG00000181019	0.016630714	1.588553651	NAD(P)H Quinone Dehydrogenase 1
SEMA6B	ENSG00000167680	0.012806465	1.593928782	Semaphorin 6B
INHBE	ENSG00000139269	0.029979007	1.602149811	Inhibin Subunit Beta E
DLGAP1-AS2	ENSG00000262001	0.03563153	1.804231675	DLGAP1 Antisense RNA 2 lncRNA
CLU	ENSG00000120885	0.027029548	1.885964664	Clusterin
EFNB2	ENSG00000125266	0.043624886	1.895780629	Ephrin B2
HTRA3	ENSG00000170801	0.04411751	1.90224253	HtrA Serine Peptidase 3

SPTA1	ENSG00000163554	0.028589527	2.020181542	Spectrin Alpha, Erythrocytic 1
HMOX1	ENSG00000100292	0.01569204	2.12103634	Heme Oxygenase 1
LUCAT1	ENSG00000248323	0.027029548	2.125101368	Lung Cancer Associated Transcript 1
OLAH	ENSG00000152463	0.029979007	2.172530721	Oleoyl-ACP Hydrolase
TMEM140	ENSG00000146859	0.033667005	2.184236911	Transmembrane Protein 140
NMRAL2P	ENSG00000171658	0.04552215	2.207970407	NmrA Like Redox Sensor 2, Pseudogene
U62317.1	ENSG00000226738	0.037309037	2.329720462	Uncharacterized LOC105373098 RNA gene
KLHDC7B	ENSG00000130487	0.041241836	2.412309763	Kelch Domain Containing 7B
AL596330.1	ENSG00000229400	0.046161935	2.590886006	Subcategory (RNA class) for ENSG00000229400 Gene
TREML3P	ENSG00000184106	0.012806465	2.6773663	Triggering Receptor Expressed On Myeloid Cells Like 3, Pseudogene
SGCG	ENSG00000102683	0.027029548	2.756561884	Sarcoglycan Gamma
NEUROD4	ENSG00000123307	0.027029548	2.78155677	Neuronal Differentiation 4
TREML4	ENSG00000188056	0.025802657	2.881204182	Triggering Receptor Expressed On Myeloid Cells Like 4

Listed genes satisfied the following criteria q<0.05 and a log-fold change of 1.

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

198 Pathway (Metascape) analysis of the 359 genes (figure 2) listed in the online table 2, revealed
199 5 significant pathways after statistical correction: regulation of cell adhesion, response to
200 inorganic substances, apoptotic signalling pathway, response to toxic substances and
201 regulation of the innate immune response.
202
203 **Figure 2. Pathway (Metascape) analysis MNPA incubation**
204
205 **Legend**
206 Metascape Bar plot (P value (log10 scale)) showing Top 20 MNPA-induced, enriched
207 functional ontology clusters (GO and KEGG terms) one per cluster.
208
209
210

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

211 # Discussion

212

213 Whole transcriptome analysis by RNASeq of the human monocyte line THP1 reveals a
214 significant number of genes that are either downregulated or upregulated in response to
215 incubation with arecoline or with MNPA. The aim of our study was to identify genes
216 associated with diabetes, obesity and metabolic syndrome whose expression was significantly
217 altered by exposure to the arecal compounds arecoline and its nitrosated metabolite, 3-
218 methylnitrosaminopropionaldehyde (MNPA). It was also hoped to determine whether the
219 strength of the evidence for any of those genes significantly affected might warrant further
220 investigation amongst betel-chewing communities with a high prevalence of metabolic
221 syndrome related disorders including obesity and type 2 diabetes.

222

223 Consistent with the established effects of betel nut ingestion in humans, a number of cellular
224 pathways and genes have been identified as being significantly affected by the arecal
225 compounds used in our approach; these genes are known to relate to immune responses, to
226 cell differentiation and lineage, to responses to toxic and inorganic substances and to the
227 development of obesity and type 2 diabetes in humans. Other genes with significantly altered
228 expression are known to be associated with carcinogenesis and immune function and a
229 smaller number of affected genes are related to neural development and could be associated
230 with addiction.

231

232 *Genes of relevance to obesity, diabetes and metabolic syndrome with good evidence that*
233 *arecoline and/or MNPA may alter its expression based on published studies and genetic*
234 *evidence include:*

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

235 **C-Type Lectin Domain Containing 10A (*CLEC10A*)** is a calcium dependent endocytic
236 receptor also known as the macrophage galactose-type lectin (MGL or CD301). It has been
237 demonstrated to have a role in regulating adaptive and innate immune responses and is
238 expressed in adipose tissue macrophages where it is associated with phenotypic switching of
239 ATM subclasses in mice that then demonstrate either a lean or an obese phenotype[20].
240 Furthermore, evidence in humans demonstrates that missense and protein truncating
241 mutations of the CLEC10A gene are strongly associated with the development of type 2
242 diabetes (http://www.type2diabetesgenetics.org/gene/geneInfo/CLEC10A). In earlier rodent
243 experiments MGL1 was found to be a novel regulator of monocyte trafficking in adipose
244 tissue in response to dietary induced obesity[20, 21].
245 **Mitogen-Activated Protein Kinase 8 Interacting Protein 1 (*MAPK8IP1*)** gene encodes a
246 regulator of pancreatic beta-cell function; it is also expressed in a large number of tissues
247 including many associated with immune function and is also a trans-activator of the glucose
248 transporter GLUT2. *MAPK8IP1* has a strong association with type 2 diabetes with a missense
249 mutation found in one family and, in *vitro,* that mutation was found to be a key down-
250 regulator of beta cell function[22].
251 **Neuronal Growth Regulator 1 (*NEGR1*)** is involved in cell adhesion and certain mutations
252 of this gene lead to Niemannn-Pick disease, a rare inherited metabolic disorder. Multiple
253 genome-wide association studies demonstrate strong genome-wide association (GWAS)
254 signals for this gene with BMI, waist circumference and type 2 diabetes[23]
255 (http://www.type2diabetesgenetics.org/gene/geneInfo/NEGR1). Furthermore, *NEGR1*
256 knockout mice develop increased adiposity including increased hepatocyte fat deposition
257 together with increases in glycaemia and in fasting serum insulin levels[24].
258 **NAD(P)H Quinone Dehydrogenase 1 (*NQO1*)** gene is a member of the NAD(P)H
259 dehydrogenase (quinone) family and encodes a cytoplasmic 2-electron reductase (Entrez

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

260 Gene) and is part of the antioxidant defence system. There is strong genetic evidence to
261 support an association between *NQO1* variants by GWAS and increased risks of type 2
262 diabetes and increased BMI (http://www.type2diabetesgenetics.org/gene/geneInfo/NQO1).
263 NQO1 is highly expressed in human adipose tissue and its expression is reduced during diet-
264 induced weight loss; furthermore its expression correlates directly with adiposity, glycaemia
265 and markers of liver dysfunction[25]. Together, these findings indicate a role for NQO1 in
266 the aetiology of obesity and type 2 diabetes.
267 **Inhibin Subunit Beta E (*INHBE*)** gene is a member of the Transforming Growth Factor
268 (TGF) beta superfamily. The transcribed peptide Activin E is ubiquitously expressed in a
269 large number of normal tissues, many being known to be especially active with cell
270 proliferation, apoptosis, immune response and hormone secretion. The highest expression is
271 found in the liver where it acts as a hepatokine with effects on energy homeostasis in both
272 brown and white adipose tissue[26]. The candidacy of the *INHBE* gene is further supported
273 by strong GWAS signals associating it with cardiometabolic traits, raised serum triglycerides
274 and with coronary heart disease
275 (http://www.type2diabetesgenetics.org/gene/geneInfo/INHBE).
276
277 *Others with suggestive evidence include:*
278 **Glutamate receptor interacting protein 1 (*GRIP1*)** mediates the trafficking and membrane
279 organisation of a number of trans-membrane proteins in various cells including neurons and
280 macrophages. Obese mice with a conditional knockout of *GRIP1* in macrophages develop
281 massive macrophage infiltration and inflammation in many metabolically active tissues
282 leading to many features that associate with the metabolic syndrome such as hepatic steatosis,
283 hyperglycaemia and insulin resistance[27]. **Clusterin (*CLU*)** is a molecular chaperone.
284 Secretory clusterin is also known as ApoJ. ApoJ has recently been identified as a novel

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

285 hepatokine, and deletion of hepatic ApoJ leads to insulin resistance and glucose
286 tolerance[28]. Furthermore, in humans, serum ApoJ levels correlate directly with increases in
287 insulin resistance and these levels decrease in response to rosiglitazone treatment[29].
288 **Insulin Like Growth Factor Binding Protein 3 (*IGFBP3*)** is the most abundant of 6 IGF-
289 binding proteins. Important interactions have been observed between IGFBP3, vitamin D
290 metabolism and obesity[30]. Furthermore, in people with type 2 diabetes IGFB3 levels may
291 inversely contribute to accelerated cerebrovascular disease[31]. **Potassium Voltage-Gated**
292 **Channel Subfamily Q Member 5 (*KCNQ5*)** is a component of potassium channels. A strong
293 GWAS association is seen between *KCNQ5* and body mass index
294 (http://www.type2diabetesgenetics.org/gene/geneInfo/KCNQ5). **Cytotoxic And Regulatory**
295 **T-Cell Molecule (*CRTAM*)** is a Protein Coding gene that has a role in the innate immune
296 system and has also been implicated as a potential determinant of insulin secretion[32]. A
297 strong GWAS association is seen between *CRTAM* and both body mass index and systolic
298 blood pressure (http://www.type2diabetesgenetics.org/gene/geneInfo/CRTAM). **Spectrin**
299 **(*SPTA1*)** is a component of the erythrocyte plasma membrane. A strong association is seen
300 between *SPTA1* and separately with HbA1c (GWAS) and type 2 diabetes adjusted for BMI
301 (mainly missense mutations) (http://www.type2diabetesgenetics.org/gene/geneInfo/SPTA1).
302
303
304 A weakness of our approach is that although we have used a human immune cell line
305 approach, we have not yet validated our findings *in vivo* in humans. Various compounds have
306 been isolated and identified from *Areca catechu* nuts including alkaloids, tannins, flavones,
307 triterpenes, steroids, and fatty acids. We have chosen to focus on arecoline and two of the
308 several nitrosated products (MNPA and MNPN), selected as being the most carcinogenic of
309 them, and in particular, because low-dose nitrosamines cause type 2 diabetes experimentally

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

310 and in humans[12, 13]. Unfortunately, for technical reasons, we did not get results with
311 MNPN. *Areca catechu* chewing quids often contain various other additives such as slaked
312 lime, spices, sweeteners, and are wrapped in leaves of the *Piper betle* vine; furthermore, in
313 many countries other than Taiwan they often contain chewing tobacco. We cannot therefore
314 exclude the possibility that major effects of chewing betel quids in humans may be due to
315 ingestion of betel quid components other than those from the *Areca catechu* nut. However,
316 obesity and hyperglycaemia were induced in CD1 mice fed *Areca catechu* nut without any
317 other betel quid component[9] and this data contributed to our focus on the findings for genes
318 associated with those particular disorders in humans.
319

320 # Conclusion

321 This pilot study has identified a large number of genes whose expression changed
322 significantly in human TPH1 cells following incubation with arecoline and MNPA and these
323 genes are known to be associated with increased risks of obesity and type 2 diabetes in
324 humans. These findings suggest that further investigation of these genes in betel-quid
325 chewers with obesity and/or type 2 diabetes is warranted.
326
327

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

328 # Acknowledgements.

329 The authors thank Nisha Patel (Borchukova laboratory, Blizard Institute, Queen Mary
330 University of London) for technical support.

331

332 # Author contributions.

333 SC, WO, CEM, ES performed the investigations and RL the bioinformatic analysis, GAH,
334 WO and BJB supervised the study, GAH conceptualized the study and acquired the funding,
335 GAH wrote the first draft of the manuscript with the help of BJB and all authors contributed
336 to the writing of the manuscript thereafter. GAH is the guarantor of this work and, as such,
337 had full access to all the data in the study and takes responsibility for the integrity of the data
338 and the accuracy of the data analysis.

339

340

341

342 This article contains supplementary material

343

344

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

345 # References

346 [1] Hills AP, Arena R, Khunti K, et al. (2018) Epidemiology and determinants of type 2
347 diabetes in south Asia. Lancet Diabetes Endocrinol 6(12): 966-978. 10.1016/S2213-
348 8587(18)30204-3

349 [2] Ko YC, Huang YL, Lee CH, Chen MJ, Lin LM, Tsai CC (1995) Betel quid chewing,
350 cigarette smoking and alcohol consumption related to oral cancer in Taiwan. J Oral Pathol
351 Med 24(10): 450-453. 10.1111/j.1600-0714.1995.tb01132.x

352 [3] Tsai JF, Chuang LY, Jeng JE, et al. (2001) Betel quid chewing as a risk factor for
353 hepatocellular carcinoma: a case-control study. Br J Cancer 84(5): 709-713.
354 10.1054/bjoc.1999.1597

355 [4] Wu MT, Lee YC, Chen CJ, et al. (2001) Risk of betel chewing for oesophageal cancer
356 in Taiwan. Br J Cancer 85(5): 658-660. 10.1054/bjoc.2001.1927

357 [5] Yamada T, Hara K, Kadowaki T (2013) Chewing betel quid and the risk of metabolic
358 disease, cardiovascular disease, and all-cause mortality: a meta-analysis. PLoS One 8(8):
359 e70679. 10.1371/journal.pone.0070679

360 [6] Tung TH, Chiu YH, Chen LS, et al. (2004) A population-based study of the association
361 between areca nut chewing and type 2 diabetes mellitus in men (Keelung Community-based
362 Integrated Screening programme No. 2). Diabetologia 47(10): 1776-1781. 10.1007/s00125-
363 004-1532-2

364 [7] Yen AM, Chiu YH, Chen LS, et al. (2006) A population-based study of the association
365 between betel-quid chewing and the metabolic syndrome in men. Am J Clin Nutr 83(5):
366 1153-1160. 10.1093/ajcn/83.5.1153

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

367 [8] Yen AM, Boucher BJ, Chiu SY, et al. (2016) Longer Duration and Earlier Age of Onset
368 of Paternal Betel Chewing and Smoking Increase Metabolic Syndrome Risk in Human
369 Offspring, Independently, in a Community-Based Screening Program in Taiwan. Circulation
370 134(5): 392-404. 10.1161/CIRCULATIONAHA.116.021511

371 [9] Boucher BJ, Ewen SW, Stowers JM (1994) Betel nut (Areca catechu) consumption
372 and the induction of glucose intolerance in adult CD1 mice and in their F1 and F2 offspring.
373 Diabetologia 37(1): 49-55. 10.1007/BF00428777

374 [10] Boucher BJ, Mannan N (2002) Metabolic effects of the consumption of Areca
375 catechu. Addict Biol 7(1): 103-110. 10.1080/13556210120091464

376 [11] IARC Working Group on the Evaluation of Carcinogenic Risks to Humans Betel-quid
377 and areca-nut chewing and some areca-nut derived nitrosamines (2004) Betel-quid and areca-
378 nut chewing and some areca-nut derived nitrosamines. IARC Monogr Eval Carcinog Risks
379 Hum 85: 1-334

380 [12] Portha B, Giroix MH, Cros JC, Picon L (1980) Diabetogenic effect of N-
381 nitrosomethylurea and N-nitrosomethylurethane in the adult rat. Ann Nutr Aliment 34(5-6):
382 1143-1151

383 [13] Karam JH, Lewitt PA, Young CW, et al. (1980) Insulinopenic diabetes after rodenticide
384 (Vacor) ingestion: a unique model of acquired diabetes in man. Diabetes 29(12): 971-978.
385 10.2337/diab.29.12.971

386 [14] Garg A, Chaturvedi P, Gupta PC (2014) A review of the systemic adverse effects of
387 areca nut or betel nut. Indian J Med Paediatr Oncol 35(1): 3-9. 10.4103/0971-5851.133702

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

388 [15] Ogunkolade WB, Boucher BJ, Bustin SA, et al. (2006) Vitamin D metabolism in
389 peripheral blood mononuclear cells is influenced by chewing "betel nut" (Areca catechu)
390 and vitamin D status. J Clin Endocrinol Metab 91(7): 2612-2617. 10.1210/jc.2005-2750

391 [16] Kasikara C, Doran AC, Cai B, Tabas I (2018) The role of non-resolving inflammation in
392 atherosclerosis. J Clin Invest 128(7): 2713-2723. 10.1172/JCI97950

393 [17] Saltiel AR, Olefsky JM (2017) Inflammatory mechanisms linking obesity and
394 metabolic disease. J Clin Invest 127(1): 1-4. 10.1172/JCI92035

395 [18] Tsuchiya S, Yamabe M, Yamaguchi Y, Kobayashi Y, Konno T, Tada K (1980)
396 Establishment and characterization of a human acute monocytic leukemia cell line (THP-1).
397 Int J Cancer 26(2): 171-176. 10.1002/ijc.2910260208

398 [19] Bray NL, Pimentel H, Melsted P, Pachter L (2016) Near-optimal probabilistic RNA-seq
399 quantification. Nat Biotechnol 34(5): 525-527. 10.1038/nbt.3519

400 [20] Lumeng CN, DelProposto JB, Westcott DJ, Saltiel AR (2008) Phenotypic switching of
401 adipose tissue macrophages with obesity is generated by spatiotemporal differences in
402 macrophage subtypes. Diabetes 57(12): 3239-3246. 10.2337/db08-0872

403 [21] Westcott DJ, Delproposto JB, Geletka LM, et al. (2009) MGL1 promotes adipose
404 tissue inflammation and insulin resistance by regulating 7/4hi monocytes in obesity. J Exp
405 Med 206(13): 3143-3156. 10.1084/jem.20091333

406 [22] Waeber G, Delplanque J, Bonny C, et al. (2000) The gene MAPK8IP1, encoding islet-
407 brain-1, is a candidate for type 2 diabetes. Nat Genet 24(3): 291-295. 10.1038/73523

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

408 [23] Willer CJ, Speliotes EK, Loos RJ, et al. (2009) Six new loci associated with body mass
409 index highlight a neuronal influence on body weight regulation. Nat Genet 41(1): 25-34.
410 10.1038/ng.287

411 [24] Joo Y, Kim H, Lee S, Lee S (2019) Neuronal growth regulator 1-deficient mice show
412 increased adiposity and decreased muscle mass. Int J Obes (Lond) 43(9): 1769-1782.
413 10.1038/s41366-019-0376-2

414 [25] Palming J, Sjoholm K, Jernas M, et al. (2007) The expression of NAD(P)H:quinone
415 oxidoreductase 1 is high in human adipose tissue, reduced by weight loss, and correlates
416 with adiposity, insulin sensitivity, and markers of liver dysfunction. J Clin Endocrinol Metab
417 92(6): 2346-2352. 10.1210/jc.2006-2476

418 [26] Hashimoto O, Funaba M, Sekiyama K, et al. (2018) Activin E Controls Energy
419 Homeostasis in Both Brown and White Adipose Tissues as a Hepatokine. Cell Rep 25(5):
420 1193-1203. 10.1016/j.celrep.2018.10.008

421 [27] Coppo M, Chinenov Y, Sacta MA, Rogatsky I (2016) The transcriptional coregulator
422 GRIP1 controls macrophage polarization and metabolic homeostasis. Nat Commun 7: 12254.
423 10.1038/ncomms12254

424 [28] Seo JA, Kang MC, Yang WM, et al. (2020) Apolipoprotein J is a hepatokine regulating
425 muscle glucose metabolism and insulin sensitivity. Nat Commun 11(1): 2024.
426 10.1038/s41467-020-15963-w

427 [29] Seo JA, Kang MC, Ciaraldi TP, et al. (2018) Circulating ApoJ is closely associated with
428 insulin resistance in human subjects. Metabolism 78: 155-166.
429 10.1016/j.metabol.2017.09.014

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

430 [30] Al-Daghri NM, Manousopoulou A, Alokail MS, et al. (2018) Sex-specific correlation of
431 IGFBP-2 and IGFBP-3 with vitamin D status in adults with obesity: a cross-sectional serum
432 proteomics study. Nutr Diabetes 8(1): 54. 10.1038/s41387-018-0063-8

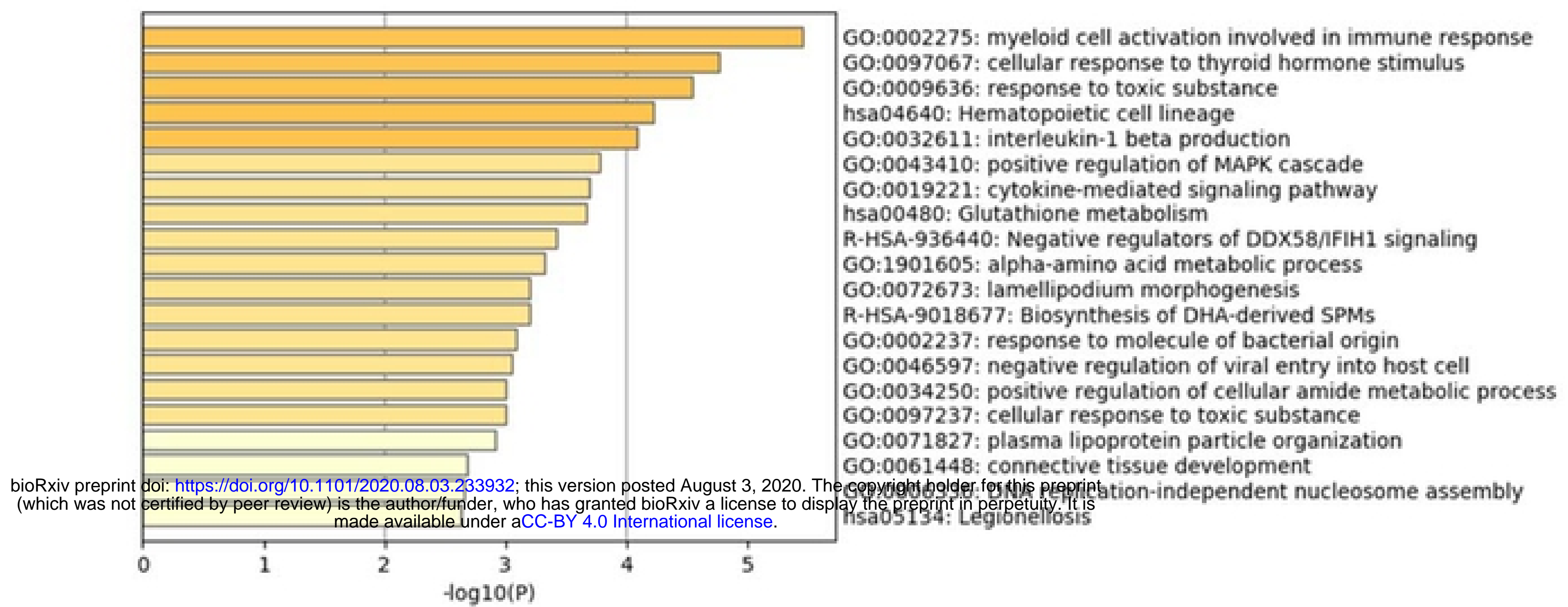
433 [31] Hjortebjerg R, Laugesen E, Hoyem P, et al. (2017) The IGF system in patients with
434 type 2 diabetes: associations with markers of cardiovascular target organ damage. Eur J
435 Endocrinol 176(5): 521-531. 10.1530/EJE-16-0940

436 [32] Curran AM, Scott-Boyer MP, Kaput J, et al. (2018) A proteomic signature that reflects
437 pancreatic beta-cell function. PLoS One 13(8): e0202727. 10.1371/journal.pone.0202727

438

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 1. Pathway (Metascape) analysis arecoline incubation



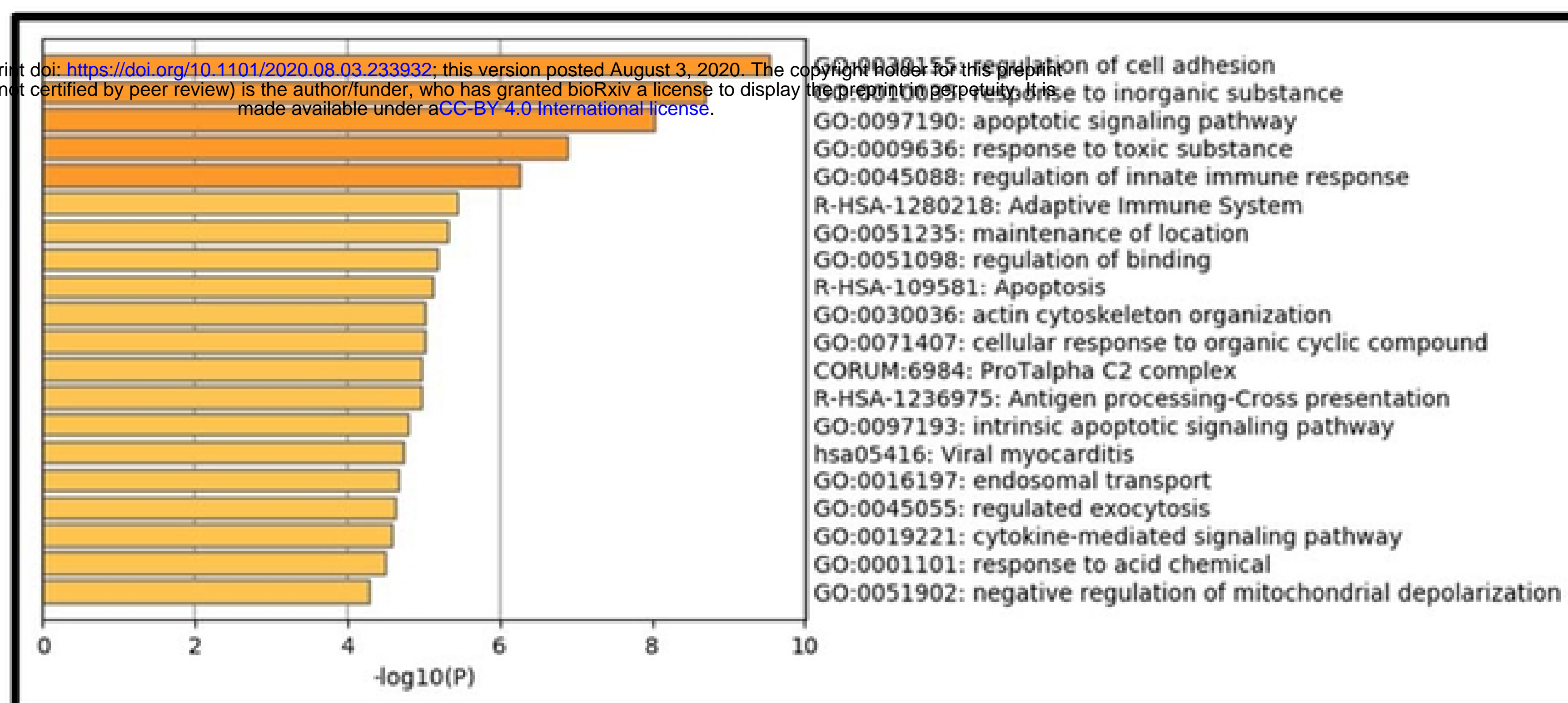
bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Metascape Bar plot (P value (log10 scale)) showing Top 20 arecoline-induced, enriched functional ontology clusters (GO and KEGG terms) one per cluster.

Figure 1

Figure 2. Pathway (Metascape) analysis MNPA incubation

bioRxiv preprint doi: https://doi.org/10.1101/2020.08.03.233932; this version posted August 3, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



Metascape Bar plot (P value (log10 scale)) showing Top 20 **MNPA-induced**, enriched **functional** ontology clusters (GO and KEGG terms) one per cluster.

figure 2