1 Oral Vitamin D supplementation induces transcriptomic changes in rectal

- 2 mucosa that are consistent with anti-tumour effects.
- 3
- 4 Short title: Vitamin D and rectal gene expression

5 Authors: P.G. Vaughan-Shaw^{1,2}, G. Grimes^{1,2}, JP Blackmur^{1,2}, M. Timofeeva^{3,4}, M

- 6 Walker^{1,2}, L.Y Ooi², Victoria Svinti^{1,2}, Kevin Donnelly^{1,2}, FVN Din^{1,2}, **S.M. Farrington**^{1,2},
- 7 **M.G. Dunlop**^{1,2.}
- 8

9 Affiliations:

- ¹MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of
- 11 Edinburgh, Edinburgh, United Kingdom.
- ¹² ² Cancer Research UK Edinburgh Centre, Institute of Genetics and Molecular Medicine,
- 13 University of Edinburgh, Edinburgh, United Kingdom.
- ³ DIAS, Danish Institute for Advanced Study, Department of Public Health, University of
- 15 Southern Denmark, Odense, Denmark
- ⁴ Deanery of Molecular, Genetic & Population Health Sciences, in the College of Medicine &
- 17 Veterinary Medicine, University of Edinburgh, Edinburgh, United Kingdom
- ¹⁸ ³Department of Pathology, National University Hospital, Singapore.
- 19
- 20 Corresponding Author: Professor Malcolm G Dunlop, Institute of Genetics and Molecular
- 21 Medicine, University of Edinburgh and MRC Human Genetics Unit, Western General
- 22 Hospital Edinburgh, Crewe Road, Edinburgh, EH4 2XU, United Kingdom; M;
- 23 <u>malcolm.dunlop@igmm.ed.ac.uk</u>
- 24
- 25
- 26

27 Abstract

28 Background

Risk for several common cancers is influenced by the transcriptomic landscape of the respective tissue-of-origin. Vitamin D influences *in-vitro* gene expression and cancer cell growth. We sought to determine whether oral vitamin D induces beneficial gene expression effects in human rectal epithelium and identify biomarkers of response.

33

34 Methods

35 Blood and rectal mucosa was sampled from 191 human subjects and mucosa gene expression 36 (HT12) correlated with plasma vitamin D (25-OHD) to identify differentially expressed 37 genes. Fifty subjects were then administered 3200IU/day oral vitamin D3 and matched 38 blood/mucosa resampled after 12 weeks'. Transcriptomic changes (HT12/RNAseq) after 39 supplementation were tested against the prioritised genes for gene-set and GO-process 40 enrichment. To identify blood biomarkers of mucosal response, we derived receiver-operator 41 curves and C-statistic (AUC) and tested biomarker reproducibility in an independent 42 Supplementation Trial (BEST-D).

43

44 **Results**

45 629 genes were associated with 25-OHD level (P<0.01), highlighting 453 GO-term processes 46 (FDR<0.05). In the whole intervention cohort, vitamin D supplementation enriched the 47 prioritised mucosal gene-set (upregulated gene-set P<1.0E-07; downregulated gene-set 48 P<2.6E-05) and corresponding GO terms (P=2.90E-02), highlighting gene expression 49 patterns consistent with anti-tumour effects. However, only 9 individual participants (18%) 50 showed a significant response (NM gene-set enrichment P<0.001) to supplementation. 51 Expression changes in *HIPK2* and *PPP1CC* expression served as blood biomarkers of 52 mucosal transcriptomic response (AUC=0.84 [95%CI:0.66-1.00]), and replicated in BEST-D 53 trial subjects (HIPK2 AUC=0.83 [95%CI:0.77-0.89]; PPP1CC AUC=0.91 [95%CI:0.86-54 0.95]).

55

56 Conclusions

57 Higher plasma 25-OHD correlates with rectal mucosa gene expression patterns consistent 58 with anti-tumour effects and this beneficial signature is induced by short-term vitamin D 59 supplementation. Heterogenous gene expression responses to vitamin D may limit the ability 60 of randomised trials to identify beneficial effects of supplementation on CRC risk. However,

61	in the current study blood expression changes in HIPK2 and PPP1CC identify those
62	participants with significant anti-tumor transcriptomic responses to supplementation in the
63	rectum. These data provide compelling rationale for a trial of vitamin D and CRC prevention
64	using easily assayed blood gene expression signatures as intermediate biomarkers of
65	response.
66	
67	Keywords: Vitamin D; colorectal cancer; gene expression; biomarker
68	
69	
70	Glossary
71	25-OHD – 25-hydroxyvitamin D
72	AUC – Area under curve
73	BEST-D - Biochemical Efficacy and Safety Trial of Vitamin D trial
74	CRC – colorectal cancer
75	DNA – Deoxyribonucleic acid
76	FC – fold-change
77	GO – Gene ontology
78	NM – normal mucosa
79	PBMC – peripheral blood mononuclear cells
80	RCT – randomised control trial
81	RNA – ribonucleic acid
82	SCOVIDS - Scottish Vitamin D study
83	SNP – single nucleotide polymorphism
84	VDR- vitamin D receptor gene
85	
86	
87	
88	
89	
90	
91	
92	

93

94

95 INTRODUCTION

96 Vitamin D deficiency is associated with risk of several common cancers, the strongest 97 evidence supporting a link between vitamin D and colorectal cancer [1, 2]. However, a causal 98 association has yet to be convincingly demonstrated, because the available observational 99 evidence may be participant to several potential confounders. Environmental risk factors 100 associated with CRC also associated with vitamin D status (i.e. co-causality; e.g., physical 101 activity), while CRC or its treatment may itself lower plasma vitamin D levels (i.e. reverse 102 causation). However, a recent randomised-control trial (RCT) reported an association 103 between supplementation, vitamin D receptor genotype and risk of colorectal adenoma, 104 supporting the premise that the beneficial effect may be causal [3]. Meanwhile, vitamin D-105 related genetic variation has been shown to influence the association between 25-OHD level 106 and CRC survival [4-6], with a recent meta-analysis of RCT data strongly supporting a causal 107 effect for vitamin D supplementation on CRC mortality [7, 8].

108 Differences in gene expression have been reported in CRC and adenoma tissue relative 109 to normal colorectal tissue [9-12], with genes involved in metabolism, transcription and 110 translation and cellular processes commonly altered [13]. Recent transcriptome wide 111 association studies confirm importance of gene expression in carcinogenesis [14, 15]. 112 Vitamin D broadly influences gene expression through activation of the ligand-activated 113 transcription factor VDR, which has been shown to influence cancer cell growth in vitro [16]. 114 Therefore, investigation of gene expression in the colorectum in the context of vitamin D 115 status or supplementation may provide fresh insight into mechanisms underlying the 116 relationship between CRC and vitamin D. Recent evidence suggests one mechanism may be 117 that 1,25-dihydroxyvitamin D3 modulates immune and inflammatory pathway genes in large

bowel epithelium [17]. However, differential expression in response to high dose 1,25dihydroxyvitamin D3 may not accurately reflect the relationship between vitamin D status and gene expression at normal or low vitamin D levels, or in response to regular vitamin D3, the most commonly used vitamin D supplement.

122 We investigated whether circulating vitamin D concentration is associated with 123 differential gene expression in rectal normal mucosa using a 2-Phase approach with 124 validation of putative biomarkers in an independent study dataset. We directly assayed total 125 25-OHD, which reflects both dietary intake and skin synthesis of vitamin D [18, 19] and 126 investigated its relationship with gene expression in normal mucosa and blood, assessed by 127 microarray. In the Phase 1 correlative study we sought to identify a prioritised list of 128 differentially expressed genes associated with 25-OHD level. In Phase 2, we conducted a 129 study in human volunteers who were supplemented with oral vitamin D to determine whether 130 the corresponding transcriptomic response was induced in vivo. Using blood peripheral blood 131 mononuclear cells (PBMC) transcriptomic analysis we also identified potential blood 132 biomarkers that indirectly indicate a beneficial response in the host rectal mucosa.

- 133

134

- 135
- 136
- 137
- 138
- 139
- 140

142

143

144 **METHODS**

145 **Study Population**

Participants recruited to Phase 1 of the Scottish Vitamin D study (SCOVIDS) (n=191)
underwent sampling of blood and normal rectal mucosa by rigid sigmoidoscopic biopsy.
RNA was extracted for gene expression analysis from matched contemporaneous NM and
peripheral blood mononuclear cells (PBMCs). Plasma was collected at the same time for
vitamin D analysis and DNA was extracted from whole blood for genotyping.

151 All eligible participants from Phase 1 were invited to proceed to Phase 2 which was 152 an intervention study. Five of the 50 recruited participants to Phase 2 underwent interval 153 sampling *before* starting supplementation to assess for longitudinal changes in expression 154 before treatment. All Phase 2 participants were then administered vitamin D supplementation 155 and underwent repeat NM, PBMC and 25-OHD sampling after 12 weeks' 3200IU/ day 156 cholecalciferol (Fultium-D3) supplementation (Figure 1). Concordance with the treatment 157 protocol was assessed through a dose diary and pharmacy log of unused tablets (compliance 158 of 98% of total doses taken achieved). Demographic and clinical data were prospectively 159 collected from patient case notes.

160

161 Sample size considerations

162 There were no available published data on which to base investigation of the sample 163 size required to determine an association between vitamin D status and global gene 164 expression in normal mucosa. Thus, a formal sample size estimation was not possible.

165

166 Blood and mucosa sampling

167	Participants were sampled in outpatient clinic or during minor surgical procedures. No
168	participant received cleansing oral mechanical bowel preparation. Blood was sampled by
169	standard venepuncture of a peripheral arm vein, with plasma and PBMCs extracted
170	(Supplementary Methods). A separate blood sample was taken to allow extraction of blood
171	leukocyte DNA and genotyping of relevant vitamin D receptor and pathway SNPs. Normal
172	rectal mucosa (NM) was sampled at the same time via rigid sigmoidoscopic rectal biopsy.
173	NM and PBMC samples were immediately placed in RNAlater and kept immersed for 24-72
174	hours prior to RNA extraction or storage at -80 \Box C.
175	
176	Plasma vitamin D assay
177	All plasma samples were measured to a standard, validated and published protocol by
178	a single laboratory[20]. Total 25-OHD was measured by liquid chromatography tandem mass
179	spectrometry.

180

181 Assessment of gene expression

182 RNA was extracted and purified from NM and blood PBMC using a proprietary RNA 183 extraction kit. Gene expression profiling was undertaken using the Illumina HumanHT-184 12v4.0 Expression BeadChip Arrays and IScan NO660 scanner, providing coverage of 185 47,231 transcripts and >31,000 annotated genes. For RNA sequencing, whole-186 genome transcriptomic patterns were analysed on total RNA from selected normal mucosa 187 samples extracted as described above. RNA was sequenced on the Illumina HiSeq 2500 188 platform in "rapid mode" with 150bp paired-end reads in a single batch. Transcript indexing 189 and quantification from RNA-seq reads was performed using Salmon v1.1.0[21].

190

191 Statistical analysis

192 All statistical analysis was undertaken in R[22]. In Phase 1, linear regression modelling was 193 used to test association between 25-OHD level and NM gene expression, adjusting for age, 194 gender, CRC status and anaesthetic status (i.e. sampled under general anaesthetic). Genes 195 associated with 25-OHD level at significance level P<0.01 were termed the 'candidate gene-196 set' and taken forward for testing in the intervention dataset (Phase 2). In Phase 2, differences 197 in 25-OHD level before and after vitamin D supplementation were investigated using paired 198 Wilcoxon rank-sum test and differential gene expression analysis in response to vitamin D 199 supplementation performed using the *lmFit* and *eBayes* functions within the 'limma' package 200 [23] producing the intervention (Phase 2) dataset. Ranked lists of differentially expressed 201 genes were assessed for functional relevance using the 'GOrilla', Gene Ontology enRIchment 202 anaLysis and visuaLizAtion tool [24]. Process ontologies were investigated using gene lists 203 ranked by coefficient (Phase 1) or fold-change (Phase 2).

We tested the Phase 2 dataset (i.e. response to supplementation) for enrichment of the Phase 1 candidate gene-set and top-ranked GO terms. Directional gene-set testing was performed in R, using the gene-setTest function in the 'limma' package [25]. We performed technical replication by performing gene-set enrichment testing on differential expression data derived from RNA-seq analysis of the same NM samples from the intervention cohort.

209 To identify biomarkers of response, we performed participant-level gene-set 210 enrichment testing with a 'response' to supplementation defined as enrichment (P<0.001 211 given n=50 subjects) of the candidate gene-set after supplementation. Then, differentially 212 expressed genes in the *blood* between those with/ without rectal NM response were tested for 213 enrichment of the candidate gene-set. Logistic regression testing sought to identify potential 214 blood biomarkers of response and utility of blood biomarkers was calculated using receiver 215 operator curves and C statistic. Finally, we sought to validate putative biomarkers of response 216 in an independent blood gene expression dataset derived from the 'Biochemical Efficacy and

217 Safety Trial of Vitamin D' (BEST-D) study[26] (https://www.ebi.ac.uk/arrayexpress/files/E-

218 <u>MTAB-6246/</u>).

219 **RESULTS**

220

221 Mucosal gene expression signature associated with higher 25-OHD level

222 consistent with anti-tumor effects

- In the Phase 1, 191 participants underwent rectal mucosal biopsy and blood sampling
- (Table 1). 25-OHD was nominally associated with expression of 629 probes (P<0.01), termed
- 225 the 'candidate gene-set' ('Gene-set discovery' Figure 2).
- 226

227 Table 1 Baseline characteristics, sampling variables and vitamin D status in

228 included participants

229

	PHASE 1 Correlative dataset	PHASE 2 Intervention dataset		
Ν	191	50		
Age (median, range)	63 (24-89) years	66 (24-88) years		
Gender (male)	101 (53%)	26 (52%)		
Diagnosis				
Pre-operative; colorectal cancer	57	1		
None (healthy)	63	20		
Past medical history of CRC*	23	21		
Minor anorectal pathology, no CRC^\dagger	45	7		
Abdominal tumor (not colorectal)	3	1		
Sampled under general anaesthetic	89	3		
Median 25-OHD –Baseline	40 (IQR 37) nmol/l	36 (IQR 31) nmol/l		
Median 25-OHD –After supplementation	NA	89 (IQR 33.5) nmol/l [¥]		

230

231 * Previously treated with curative resection +/- adjuvant chemotherapy and no evidence of recurrence

232 at time of recruitment. † Full diagnosis list given in Supplementary Table 1

233

234	No individual probe was significantly associated with 25-OHD after adjustment for
235	genome-wide multiple testing (Supplementary Table 2), yet the top three hits have previous
236	reported association with colorectal tumorigenesis CNN1[27], COX7A1[28], PIP5K1C[29].
237	Gene ontology analysis demonstrated significant enrichment of 453 processes
238	(Supplementary Table 3) with many highly relevant to carcinogenesis e.g. 'regulation of cell
239	migration' (FDR=7.55E-08), 'regulation of programmed cell death' (FDR=5.38E-03), and
240	'regulation of cell differentiation' (FDR=2.55E-05). Several genes from the candidate gene-
241	set with higher expression associated with higher 25-OHD are included in enriched GO
242	ontology terms relevant to carcinogenesis, and have reported tumor suppressor activity (e.g.
243	FOXOs, CAV1, LRP1, Supplementary Tables 4 and 5). This suggests that the NM gene
244	expression signature, i.e. Phase 1 'candidate gene-set', associated with higher 25-OHD level
245	is consistent with anti-tumor effects.

246

Oral vitamin D supplementation enriches anti-tumor expression signature in normal rectal mucosa

In Phase 2, 50 participants were administered vitamin D supplementation and underwent repeat sampling after 12-weeks'. Post-hoc analysis revealed age, gender and baseline 25-OHD to be similar between Phase 1/2 participants (P>0.05). Supplementation induced an increase in plasma 25-OHD after 12-weeks (median plasma 25-OHD before/after supplementation was 36nmol/l, 89nmol/l; P= 2.5E-09, Supplementary Table 6).

No individual gene from the candidate gene-set showed significant differential expression after adjustment for multiple testing (Supplementary Table 7). However, testing of the Phase 1 candidate gene-set showed significant enrichment after supplementation

- 257 (upregulated gene-set P<1.0E-07; downregulated gene-set 2.8E-05, see 'Gene-set testing'
- 258 Figure 2, Table 2, Supplementary Figure 1), confirmed in RNA-seq data.

Table 2 Gene-set testing for enrichment of the candidate gene-set from Phase 1 after supplementation in Phase 2 and the

260 BEST-D study

261

	SCOVIDS PHAS Intervention dat All participants	SE 2 taset, NM	Intervention of HT12	dataset, blood,	BEST-D trial Blood, HT12	
PHASE 1 Correlative dataset	HT12	RNA-seq	All participants	With vs. without rectal response	All participants	
Candidate gene-set: positive association with 25-OHD level	P<1.0E-07	2.05E-07	P=3.89E-13	P=3.65E-12	P=4.72E-06	
	n=349	n=242	n=239	n=291	n=191	
Candidate gene-set: negative association with 25-OHD level	P=2.8E-05	6.87E-09	P=9.60E-05	P=3.50E-17	P=0.02	
	n=206	n=155	n=185	n=217	n=130	

262

264	NM - normal mucosa. P value given for directional gene-set enrichment test of whether Phase 1 candidate gene-set showed greater change after
265	supplementation when compared to randomly chosen genes. n=number of genes tested. To explore the role of VDR genotype in modifying the association
266	between 25-OHD and gene expression, VDR haplotypes were derived for the four genotyped VDR polymorphisms, rs1544410, rs10735810, rs7975232,
267	rs11568820, using BEAGLE software (version 3.3.2) with standard settings [30] and used as additional covariates in the linear regression model in a subset
268	of 125 participants, resulting in a smaller candidate gene-set which remained significantly enriched after supplementation (P<0.0001).

Enrichment of the candidate gene-set was not associated with 25-OHD response to supplementation, with those with the lowest 25-OHD FCs still showing gene-set enrichment (Supplementary Figure 1). Meanwhile, there was no enrichment of the candidate gene-set in interval NM samples taken *before* commencement of supplementation indicating it is a treatment effect (median interval 8 weeks, Supplementary Figure 1).

Of the 629 candidate genes associated with circulating vitamin D in the Phase 1 geneset, fifty-five had nominally significant expression change after supplementation (P<0.05). Concordance in direction of effect between the coefficient of association with plasma 25-OHD level and expression change after supplementation was observed in 49 of these (R0.93, P<2.2E-16, Supplementary Table 7), with these genes taken forward for biomarker discovery (see '**Biomarker discovery'** section Figure 2).

280

281

282 GO term enrichment indicates modulation of anti-tumor effects in normal

283 mucosa by supplementation

284 Functional annotation of the intervention dataset gene list identified 65 significantly 285 enriched pathways after supplementation, with many terms relevant to carcinogenesis 286 including 'regulation of programmed cell death (FDR=9.66E-03) and 'regulation of cell 287 migration' (FDR=7.83E-03) (Supplementary Table 8). Taken together, genes in the top 50 288 GO terms from Phase 1 were significantly enriched after supplementation (P=2.90E-02). 289 Common processes across both the Phase 1 and Phase 2 datasets included terms relevant to 290 carcinogenesis including 'regulation of programmed cell death', 'regulation of cell migration' 291 (Supplementary Figure 2), demonstrating that biologically relevant patterns of gene 292 expression changes associated with higher 25-OHD level and consistent with anti-tumor 293 effects could be imparted by oral supplementation.

295 Blood expression biomarkers identify participants with gene expression

296 response to supplementation

297 We identified 9 individual participants (18%) with a significant response (i.e. 298 candidate gene-set enrichment in NM P<0.001) to supplementation. Of the top 50 ranked GO 299 terms from Phase 1, 43 were enriched in these 9 participants after supplementation indicating 300 a biologically relevant NM gene expression response to supplementation, which was absent 301 in the remaining participants (Supplementary Table 9). NM gene expression response was 302 associated with an increased allele risk score of the four functionally relevant VDR SNPs 303 (p=0.006), but not with VDR gene expression or increased 25-OHD fold-change 304 (Supplementary Table 10).

305 Changes in PBMC gene expression after supplementation reflected those in the 306 rectum, with the Phase 1 candidate gene-set significantly enriched in blood after 307 supplementation (Table 2). Moderate correlation between blood and rectum fold-change in 308 the 49 genes taken forward for biomarker discovery was seen (R=0.64, P=5.63E-06). When 309 we compared PBMC gene expression after supplementation between those participants with 310 and without a rectal mucosal gene expression response, the differentially expressed genes in 311 PBMCs were enriched for the candidate gene-set, indicating potential blood biomarkers of 312 mucosal response (Table 2, see 'Biomarker discovery' section Figure 2). Five genes 313 identified from the Phase 1 which were both differentially expressed in NM after 314 supplementation and also differentially expressed in blood between participants with and 315 without a rectal response to supplementation (SMEK2, HIPK2, PPP1C, DDR1 and SNX21), 316 indicating biomarker potential (Table 3). When the genes were combined, a blood expression 317 signature based on the best derived cut-off showed strong utility in predicting NM response 318 (AUC=0.99, 95% CI: 0.97-1.00, Supplementary Table 10, Supplementary Figure 4).

319

320 Table 3 Genes prioritised from Phase 1 correlative dataset modified by supplementation in NM with evidence of potential

321 biomarker utility in PHASE 2 and BEST-D trial

322

	PHASE 1 PHASE 2 PHASE 2 PHASE 2 NM NM Blood With vs. withou response, Block		E 2 vs. without nse, Blood	BEST Blood	D trial	BEST-D With vs respon) trial 5. <i>without</i> se, Blood					
Gene	Coeff.	P value	FC	P value	FC	P value	FC	P value	FC	P value	FC	P value
SMEK2	-0.002	0.007	0.93	0.020	0.97	0.272	0.74	0.007	-	-	-	-
HIPK2	0.004	0.006	1.17	0.016	1.11	0.068	1.46	0.014	1.07	0.002	1.36	8.47E-12
PPP1CC	-0.002	0.007	0.93	0.020	0.91	0.006	0.82	0.047	0.96	0.14	0.61	1.98E-17
SNX21	0.004	0.003	1.07	0.047	0.99	0.750	1.26	0.039	1.01	0.27	1.04	0.02
DDR1	0.007	0.008	1.14	0.0008	1.03	0.572	0.72	0.040	1.02	0.03	1.04	0.08

323

324

325 NM- normal mucosa, FC- Fold-change. Results from Phase 1 and Phase 2 analysis given. For Phase 1, coefficient given for association with 25-OHD level.

326 For Phase 2, fold-change and P value given for gene expression response to supplementation in NM and in blood. In final column, blood expression fold-

327 change difference between participants with and without NM response to the candidate gene-set is given.

328

We then explored the value of these same genes as blood biomarkers of *blood* response to supplementation. The *HIPK2* and *PPP1CC* genes (Table 3, Figure 3) demonstrated the best predictive utility in identifying participants with response in both rectum (AUC=0.84, 95%CI: 0.66-1.00) and blood (AUC=0.87, 95%CI: 0.71-1.00, Supplementary Table 11).

335

336 *HIPK2* and *PPP1CC* are independent biomarkers of expression response to 337 supplementation in the BEST-D trial expression dataset

338 In the BEST-D trial, 48/172 (28%) participants showed significant enrichment of our 339 Phase 1 candidate gene-set. HIPK2 expression increased with supplementation in the BEST-340 D trial (FC=1.07, p=0.002), with the increase greatest in those with a blood response to 341 supplementation as defined by our candidate gene-set (P=4.6E-13, Figure 3) and a 1.36 fold-342 difference in expression change after supplementation between those with/ without a 343 response to our candidate gene-set (P=8.48E-12, Table 3). In the current intervention study 344 (SCOVIDS), we observed an average HIPK2 FC=1.11 in blood, with 21 (47%) participants 345 showing a FC>1.19, the optimum threshold determined by AUC calculations. In the BEST-D 346 study, 63 (37%) participants taking oral vitamin D supplementation showed a FC>1.19 347 response, with this signature more prevalent in those taking 4000IU per day (42%), 348 suggesting a dose-response. HIPK2 expression change showed utility in identifying those 349 with a response to our candidate gene-set, AUC=0.83 (95%CI:0.77-0.89, Figure 3) and when 350 the threshold from our study was used, HIPK2 FC>1.19 AUC=0.74 (95%CI: 0.66-0.81, 351 Supplementary Table 10).

352 PPP1CC expression showed non-significant decrease after supplementation in the
353 BEST-D cohort overall, but when the participants were stratified by blood response, *PPP1CC*354 was seen to increase in those without blood response (P=0.0015, Figure 3) and markedly

- decreased in those with a blood response (P= 5.9E-15). There was a 0.61 fold-difference in
- 356 *PPP1CC* expression change between those with and without a response to our candidate
- 357 gene-set (P=1.98E-17, Table 3). A total of 47 (27%) participants had PPP1CC FC<0.76 after
- 358 supplementation, closely reflecting the 11 (26%) participants in the current intervention
- 359 study. Crucially, *PPP1CC* expression change after supplementation showed utility as a
- 360 biomarker of blood response to our candidate gene-set, AUC=0.91 (95%CI=0.86-0.95, Figure
- 361 3) and when using the threshold from the SCOVIDs study, PPP1CC FC<0.76 AUC=0.83
- 362 (95%CI:0.76-0.89, Supplementary Table 10).

363 **DISCUSSION**

364 This study reveals demonstrable differences in gene expression patterns in normal 365 rectal mucosa correlated with plasma 25-OHD level. These differences are consistent with 366 beneficial effects on processes relevant to colorectal carcinogenesis. Furthermore, we show 367 that oral supplementation with vitamin D induces changes in the prioritised gene list. This 368 indicates that the beneficial expression "signature" is not static, but rather can be modified by 369 oral vitamin D supplementation, at least within the timescale tested here. Although we were 370 not able to directly test cancer endpoints, there is considerable published evidence supporting 371 the premise that enrichment of this favourable gene-set imparts anti-tumor effects.

372 Homeodomain-interacting protein kinase 2 (HIPK2) is a known tumor suppressor 373 gene [31] and the Protein Phosphatase 1 Catalytic Subunit Gamma gene (PPP1CC), a 374 published molecular marker of CRC [32]. We found expression changes in these genes in 375 blood to have predictive value in reflecting rectal mucosa response to supplementation. 376 Hence, these may have utility as blood biomarkers of a beneficial epithelial response to 377 supplementation. In addition, the effect on gene expression in blood PBMCs appears robust, 378 since we replicated the effect in a large, independent, expression dataset, namely the BEST-D 379 trial in which subjects were administered oral vitamin D supplementation (2000/4000IU 12 380 months).

We devised a 2-Phase *in vivo* approach, firstly to identify differentially expressed genes in the rectal epithelium associated with plasma 25-OHD and determine the GO terms and processes linked to that prioritised gene list. However, our ultimate aim was to establish whether these transcriptomic responses could be recapitulated by oral vitamin D supplementation, thereby demonstrating a modifiable transcriptomic landscape. Many of the top-ranked genes associated with higher 25-OHD level have links with CRC, for instance *CNN1* [27], *COX7A1* [28], PEG3 [33], *PIP5K1C* [29], *TAGLN* [34], *DAAM2* [35].

388 Furthermore, we highlight a number of genes within processes relevant to tumorigenesis 389 which are associated with 25-OHD level and influenced by supplementation. The directions of effect of these genes were consistent with tumor suppressor activity. Enrichment of 390 391 pathways involved in cell migration and cell death validate published in vitro data which 392 demonstrate vitamin D-induced growth arrest and apoptosis of CRC cells lines, modulation 393 of the Wnt signalling pathway, DNA repair and immunomodulation [16]. Published clinical 394 data also corroborate our current findings: Protiva et al., reported upregulation of genes 395 involved in inflammation, immune response, extracellular matrix, and cell adhesion in 396 response to 1,25(OH)2D3 [17], while Fedirko *et al.*, reported a reduction in tumor promoting 397 inflammation biomarkers, decreased oxidative DNA damage, increased cell differentiation 398 and apoptosis and modification of the APC/ β -catenin pathway in the normal human 399 colorectal epithelium [36]. Taken together, these data suggest possible mechanisms 400 underlying the widely reported link between vitamin D deficiency and increased CRC risk [1, 401 2]. It also might explain the recently reported beneficial impact of supplementation on CRC 402 survival outcomes [8].

403 Despite compelling published observational and pre-clinical data, the link between 404 vitamin D and risk of cancer and several other traits remains controversial. Indeed, several 405 large intervention trials have shown no benefit on cancer endpoints (VITAL Trial [37], 406 Vitamin D Assessment (ViDA) study [38, 39] and Baron et al. [40]). However, participants 407 in these trials were predominantly sufficient for vitamin D at the trial outset, thereby 408 potentially blunting beneficial effects [41]. We have previously rehearsed potential reasons 409 why previous study designs might have failed to detect real effects [41]. To counter potential 410 confounding effects, we conducted a Mendelian randomisation study but this also did not 411 demonstrate a beneficial effect of circulating vitamin D on CRC risk. However, available

genetic instrumental variables are weak and explain only a small portion of variance of250HD levels [42, 43].

414 In this study, 18% of participants receiving vitamin D supplementation exhibited a 415 response in the colorectal epithelium (the putative target tissue). If our hypothesis holds that 416 expression changes translate to cancer endpoints, this low response rate would adversely 417 impact on statistical power of trials conducted to date which have tested the effect of vitamin 418 D supplementation on clinical endpoints. Such trials routinely perform subgroup analyses 419 based on change in circulating 25-OHD level, yet the current study reveals poor correlation 420 between plasma level and mucosal gene expression changes, suggesting non-linear responses 421 to vitamin D may introduce further heterogeneity to clinical endpoints.

422 We have identified blood biomarkers that reliably identify participants who respond 423 to vitamin D supplementation by inducing gene expression changes in the target tissue. The 424 value of these biomarkers is replicated in a larger independent expression dataset. Further 425 work is required to assess the utility of respective blood protein assays (e.g. ELISA) and 426 reproducibility of these blood biomarkers in identifying mucosal response across a larger 427 cohort. Nevertheless, these exciting and novel findings provide rationale for a trial of vitamin 428 D and CRC prevention using easily assayed blood gene expression signatures as intermediate 429 biomarkers of response.

Whilst the 2-Phase design of an intervention study informed by our correlative dataset has many positive attributes, the study has a number of limitations. First, the low median level of 25-OHD and narrow positively skewed distribution of 25-OHD in the Phase 1 cohort may have masked some true associations between vitamin D level and gene expression. Failure to identify individual gene significance after adjustment for genome-wide multiple testing may also indicate inadequate sample size, physiological autoregulation maintaining

436 constant gene expression despite differences in circulating 25-OHD, or differences between

437 plasma and rectal mucosa concentrations of 25-OHD or 1,25-OHD [44].

438 This intervention study is larger than many published studies of gene expression and 439 vitamin D supplementation [17, 45-48], yet may still have limited power to achieve 440 individual gene significance. Phase 2 participants were recruited as a subset of the Phase 1 441 cohort, which may influence gene-set enrichment test results. However, we did not select 442 those who received supplementation based on 25-OHD level or baseline gene expression, 443 which could have led to overfitting of the data, but instead took an unselected group. If 444 anything, this approach could blunt the observed effect of supplementation, as mucosal 445 response to supplementation may be capped in those with specific 25-OHD or favorable 446 patterns of gene expression at baseline. Despite this we observed significant enrichment of 447 the candidate gene-set derived in Phase 1 in those receiving supplementation. Sampling of 448 rectal mucosa but not colonic mucosa avoided the use of cleansing bowel laxatives which 449 may influence gene expression [49], yet limits the generalisability of our findings to more 450 proximal colonic mucosa. Finally, sampling after 12-weeks of supplementation may not 451 adequately capture early or later gene expression changes yet more frequent or delayed 452 sampling would provide additional practical and ethical challenges.

453 In conclusion, we report for the first time patterns of gene expression and functional 454 pathways in the normal rectal mucosa that are associated with circulating plasma vitamin D 455 level. Oral vitamin D supplementation induces transcriptomic changes consistent with 456 beneficial anti-tumor effects. Blood leukocyte expression of HIPK2 and PPP1CC predicted 457 well those participants with the greatest expression response following supplementation. 458 Whilst further replication in a separate cohort is desirable, these data provide compelling 459 rationale for a trial of vitamin D and CRC prevention using easily assayed blood gene 460 expression signatures as intermediate biomarkers of response.

- 461
- 462
- 463
- 464
- 465
- 466
- 467

468 **FIGURES**

469 Figure 1 Summary of SCOVIDS study protocol

- 470 An unselected subset of Phase 1 participants (i.e. selection not based on 25-OHD or
- 471 baseline gene expression) proceeded to Phase 2 and were given 3200IU
- 472 cholecalciferol per day. Of these 5 participants were sampled 6 weeks after initial
- sampling and without supplementation to provide a control dataset, after which they
- 474 proceeded to 12 week's supplementation and final sampling.
- 475

476 Figure 2 Analysis flowchart and gene-set selection for biomarker assessment

477

478 Figure 3 Blood HIPK2 and PPP1CC expression before and after

- 479 supplementation in SCOVIDS and BEST-D trial with ROC of biomarker utility
- 480 Response defined as participant level gene-set enrichment to Phase 1 candidate
 481 gene-set from our SCOVIDS study after HIPK2, PPP1CC, SMEK2, DDR1, SNX21
- 482 excluded.

483

484 **ADDITIONAL FILES**

485 Supplementary Methods (docx)

486 Supplementary methods to be read as an adjunct to the main methods section in the

487 manuscript.

488 Supplementary Tables (docx)

489 Supplementary Tables to be read as an adjunct to the main methods section in the manuscript.

- 490 **Protocol (doc)**
- 491 Full protocol for SCOVIDS study

492

493

494

495

496 **DECLARATIONS**

497 Ethical approval and consent

498 All participants provided informed written consent. The research was approved by the South

499 East Scotland Research Ethics Committee 03 (11/SS/0109), South East Scotland Research

500 Ethics Committee 01 (13/SS/0248) and Lothian National Health Service Research and

501 Development office (2014/0058).

502

503 **Acknowledgements:** We acknowledge statistical advice from Dr Catalina Vallejos, 504 Biomedical Data Science research group, University of Edinburgh. We acknowledge the 505 excellent technical support from Stuart Reid. We are grateful to Donna Markie, and all those 506 who continue to contribute to recruitment, data collection, and data curation for the Scottish 507 Vitamin D Study. We acknowledge the expert support on sample preparation from the 508 Genetics Core of the Edinburgh Wellcome Trust Clinical Research Facility in addition to the 509 nursing and study facilities provided by the Clinical Research Facility.

510

511 Grant Support: The work reported in this manuscript was supported by a Cancer Research

512 UK Programme Grant (C348/A18927) and a Project Leader Grant to MGD (MRC Human

513 Genetics Unit Centre Grant - U127527198). MGD is an MRC Investigator. PVS was

514 supported by a NES SCREDS clinical lectureship, MRC Clinical Research Training

515 Fellowship (MR/M004007/1), a Research Fellowship from the Harold Bridges bequest and

516 by the Melville Trust for the Care and Cure of Cancer. JPB is supported by an ECAT-linked

517 CRUK ECRC Clinical training award (C157/A23218). FVND is supported by a CSO Senior

518 Clinical Fellowship. LYO was supported by a CRUK Research Training Fellowship

519 (C10195/A12996). The work was also supported by the infrastructure and staffing of the

520 Edinburgh CRUK Cancer Research Centre.

521								
522	Trans	cript profiling: Available at https://www.ncbi.nlm.nih.gov/geo/ GEO ID GSE157982.						
523	Full phenotypic data available from the corresponding author on reasonable request.							
524								
525	Autho	r contributions: Conceptualization, PGVS, MGD; Methodology, PGVS, LYO, GG,						
526	ET, FV	/ND, SMF, MGD; Investigation, PGVS, LYO, JPB, GG, MT, MW, VS, KD; Writing						
527	Origin	al Draft PVS, SMF, MGD; Writing - Review & Editing, JPB, MT, SMF, FVND,						
528	MGD;	Funding Acquisition, MGD.; Resources, MGD; Supervision, SMF and MGD.						
529	Autho	r names in bold designate shared co-senior authorship						
530								
531	Confli	ct of Interest: The authors declare that they have no competing interests.						
532								
533								
534								
535								
536								
537								
538								
539								
540								
540	Defer							
541	Relen							
542 543	1.	Theodoratou E, Tzoulaki I, Zgaga L, Ioannidis JP: Vitamin D and multiple health outcomes: umbrella review of systematic reviews and meta-analyses of observational studies and						
544		randomised trials. <i>BMJ</i> 2014, 348 :g2035.						
545	2.	Autier P, Boniol M, Pizot C, Mullie P: Vitamin D status and ill health: a systematic review.						
546 547	2	Lancet Diabetes Endocrinol 2014, 2 (1):76-89.						
547 548	5.	Barry EL, Peacock JL, Rees JR, Bostick RW, Robertson DJ, Bresaller RS, Baron JA: Vitamin D Recentor Genotyne Vitamin D3 Supplementation and Risk of Colorectal Adenomas: A						
549		Randomized Clinical Trial. <i>JAMA Oncol</i> 2016.						
550	4.	Hermann J, Eder P, Banasiewicz T, Matysiak K, Lykowska-Szuber L: Current management of						
551		anal fistulas in Crohn's disease. Prz Gastroenterol 2015, 10(2):83-88.						
552	5.	Vaughan-Shaw PG, O'Sullivan F, Farrington SM, Theodoratou E, Campbell H, Dunlop MG,						
553		Zgaga L: The impact of vitamin D pathway genetic variation and circulating 25-						
554 555		nydroxyvitamin D on cancer outcome: systematic review and meta-analysis. Br J Cancer						
333 556	6	2017, 110(8):1092-1110. Naughan Shaw DG, Zgaga L, Ooi LV, Theodoratov E, Timofoovo MA, Svinti V, Malkar MA.						
550	υ.	O'Sullivan F Ewing A Johnston S et al. Jow plasma vitamin D is associated with adverse						
558		colorectal cancer survival after surgical resection, independent of systemic inflammatory						
559		response. Gut 2020, 69(1):103-111.						

- Keum N, Lee DH, Greenwood DC, Manson JE, Giovannucci E: Vitamin D supplementation
 and total cancer incidence and mortality: a meta-analysis of randomized controlled trials.
 Ann Oncol 2019, 30(5):733-743.
- 5638.Vaughan-Shaw PG, Buijs LF, Blackmur JP, Theodoratou E, Zgaga L, Din FVN, Farrington SM,564Dunlop MG: The effect of vitamin D supplementation on survival in patients with565colorectal cancer: systematic review and meta-analysis of randomised controlled trials. Br566J Cancer 2020.
- 5679.Jones PA, Baylin SB: The fundamental role of epigenetic events in cancer. Nature Reviews568Genetics 2002, 3(6):415-428.
- 56910.Bertucci F, Salas S, Eysteries S, Nasser V, Finetti P, Ginestier C, Charafe-Jauffret E, Loriod B,570Bachelart L, Montfort J *et al*: Gene expression profiling of colon cancer by DNA microarrays571and correlation with histoclinical parameters. Oncogene 2004, 23(7):1377-1391.
- Alon U, Barkai N, Notterman DA, Gish K, Ybarra S, Mack D, Levine AJ: Broad patterns of gene
 expression revealed by clustering analysis of tumor and normal colon tissues probed by
 oligonucleotide arrays. Proc Natl Acad Sci U S A 1999, 96(12):6745-6750.
- 57512.Chan SK, Griffith OL, Tai IT, Jones SJM: Meta-analysis of colorectal cancer gene expression576profiling studies identifies consistently reported candidate biomarkers. Cancer577Epidemiology Biomarkers & Prevention 2008, 17(3):543-552.
- 578 13. Birkenkamp-Demtroder K, Christensen LL, Olesen SH, Frederiksen CM, Laiho P, Aaltonen LA,
 579 Laurberg S, Sorensen FB, Hagemann R, Orntoft TF: Gene expression in colorectal cancer.
 580 Cancer Research 2002, 62(15):4352-4363.
- 58114.Dong X, Su YR, Barfield R, Bien SA, He Q, Harrison TA, Huyghe JR, Keku TO, Lindor NM,582Schafmayer C et al: A general framework for functionally informed set-based analysis:583Application to a large-scale colorectal cancer study. PLoS Genet 2020, 16(8):e1008947.
- 58415.Bosse Y, Li Z, Xia J, Manem V, Carreras-Torres R, Gabriel A, Gaudreault N, Albanes D, Aldrich585MC, Andrew A et al: Transcriptome-wide association study reveals candidate causal genes586for lung cancer. Int J Cancer 2020, 146(7):1862-1878.
- 58716.Fleet JC, DeSmet M, Johnson R, Li Y: Vitamin D and cancer: a review of molecular588mechanisms. Biochemical Journal 2012, 441:61-76.
- 58917.Protiva P, Pendyala S, Nelson C, Augenlicht LH, Lipkin M, Holt PR: Calcium and 1,25-590dihydroxyvitamin D3 modulate genes of immune and inflammatory pathways in the591human colon: a human crossover trial. Am J Clin Nutr 2016, 103(5):1224-1231.
- 59218.Ross AC, Manson JE, Abrams SA, Aloia JF, Brannon PM, Clinton SK, Durazo-Arvizu RA,593Gallagher JC, Gallo RL, Jones G et al: The 2011 Report on Dietary Reference Intakes for594Calcium and Vitamin D From the Institute of Medicine: What Clinicians Need to Know595EDITORIAL COMMENT. Obstetrical & Gynecological Survey 2011, 66(6):356-357.
- 59619.Seamans KM, Cashman KD: Existing and potentially novel functional markers of vitamin D597status: a systematic review. American Journal of Clinical Nutrition 2009, 89(6):1997s-2008s.
- 59820.Knox S, Harris J, Calton L, Wallace AM: A simple automated solid-phase extraction599procedure for measurement of 25-hydroxyvitamin D3 and D2 by liquid chromatography-600tandem mass spectrometry. Ann Clin Biochem 2009, 46(Pt 3):226-230.
- 60121.Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C: Salmon provides fast and bias-aware602quantification of transcript expression. Nat Methods 2017, 14(4):417-419.
- R Development Core Team: R: A language and environment for statistical computing. In.: R
 Foundation for Statistical Computing, Vienna, Austria; 2013.
- 60523.Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: limma powers differential606expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015,60743(7):e47.
- 60824.Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z: GOrilla: a tool for discovery and609visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 2009, 10:48.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A,
Pomeroy SL, Golub TR, Lander ES *et al*: Gene set enrichment analysis: a knowledge-based
approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005,
102(43):15545-15550.

- 61426.Berlanga-Taylor AJ, Plant K, Dahl A, Lau E, Hill M, Sims D, Heger A, Emberson J, Armitage J,615Clarke R et al: Genomic Response to Vitamin D Supplementation in the Setting of a616Randomized, Placebo-Controlled Trial. EBioMedicine 2018, 31:133-142.
- 61727.Yanagisawa Y, Takeoka M, Ehara T, Itano N, Miyagawa S, Taniguchi S: Reduction of Calponin618h1 expression in human colon cancer blood vessels. Eur J Surg Oncol 2008, 34(5):531-537.
- 61928.Sandberg TP, Oosting J, van Pelt GW, Mesker WE, Tollenaar R, Morreau H: Molecular620profiling of colorectal tumors stratified by the histological tumor-stroma ratio Increased621expression of galectin-1 in tumors with high stromal content. Oncotarget 2018,6229(59):31502-31515.
- 62329.Xue J, Ge X, Zhao W, Xue L, Dai C, Lin F, Peng W: PIPKIgamma Regulates CCL2 Expression in624Colorectal Cancer by Activating AKT-STAT3 Signaling. J Immunol Res 2019, 2019:3690561.
- 62530.Browning SR, Browning BL: Rapid and accurate haplotype phasing and missing-data626inference for whole-genome association studies by use of localized haplotype clustering.627Am J Hum Genet 2007, 81(5):1084-1097.
- 62831.Puca R, Nardinocchi L, Givol D, D'Orazi G: Regulation of p53 activity by HIPK2: molecular629mechanisms and therapeutical implications in human cancer cells. Oncogene 2010,63029(31):4378-4387.
- 63132.Kou Y, Zhang S, Chen X, Hu S: Gene expression profile analysis of colorectal cancer to632investigate potential mechanisms using bioinformatics. Onco Targets Ther 2015, 8:745-752.
- 63333.Zhou T, Lin W, Zhu Q, Renaud H, Liu X, Li R, Tang C, Ma C, Rao T, Tan Z et al: The role of PEG3634in the occurrence and prognosis of colon cancer. Onco Targets Ther 2019, 12:6001-6012.
- 34. Zhou HM, Fang YY, Weinberger PM, Ding LL, Cowell JK, Hudson FZ, Ren M, Lee JR, Chen QK,
 Su H et al: Transgelin increases metastatic potential of colorectal cancer cells in vivo and
 alters expression of genes involved in cell motility. *BMC Cancer* 2016, 16:55.
- 638 35. Galamb O, Kalmar A, Peterfia B, Csabai I, Bodor A, Ribli D, Krenacs T, Patai AV, Wichmann B,
 639 Bartak BK *et al*: Aberrant DNA methylation of WNT pathway genes in the development and
 640 progression of CIMP-negative colorectal cancer. *Epigenetics* 2016, **11**(8):588-602.
- 64136.Bostick RM: Effects of supplemental vitamin D and calcium on normal colon tissue and642circulating biomarkers of risk for colorectal neoplasms. J Steroid Biochem Mol Biol 2015,643148:86-95.
- Manson JE, Cook NR, Lee IM, Christen W, Bassuk SS, Mora S, Gibson H, Gordon D, Copeland
 T, D'Agostino D *et al*: Vitamin D Supplements and Prevention of Cancer and Cardiovascular
 Disease. N Engl J Med 2018.
- 64738.Scragg R, Khaw KT, Toop L, Sluyter J, Lawes CMM, Waayer D, Giovannucci E, Camargo CA, Jr.:648Monthly High-Dose Vitamin D Supplementation and Cancer Risk: A Post Hoc Analysis of649the Vitamin D Assessment Randomized Clinical Trial. JAMA oncology 2018, 4(11):e182178.
- 65039.Chandler PD, Chen WY, Ajala ON, Hazra A, Cook N, Bubes V, Lee IM, Giovannucci EL, Willett651W, Buring JE et al: Effect of Vitamin D3 Supplements on Development of Advanced Cancer:652A Secondary Analysis of the VITAL Randomized Clinical Trial. JAMA Netw Open 2020,6533(11):e2025850.
- 65440.Baron JA, Barry EL, Mott LA, Rees JR, Sandler RS, Snover DC, Bostick RM, Ivanova A, Cole BF,655Ahnen DJ et al: A Trial of Calcium and Vitamin D for the Prevention of Colorectal656Adenomas. New England Journal of Medicine 2015, 373(16):1519-1530.
- 41. Vaughan-Shaw PG, Zgaga L, Theodoratou E, Blackmur JP, Dunlop MG: Whether vitamin D
 supplementation protects against colorectal cancer risk remains an open question. Eur J
 659 Cancer 2019, 115:1-3.

42. He Y, Timofeeva M, Farrington SM, Vaughan-Shaw P, Svinti V, Walker M, Zgaga L, Meng X, Li
K, Spiliopoulou A *et al*: Exploring causality in the association between circulating 25hydroxyvitamin D and colorectal cancer risk: a large Mendelian randomisation study. *BMC Med* 2018, 16(1):142.

- 664 43. Cornish AJ, Law PJ, Timofeeva M, Palin K, Farrington SM, Palles C, Jenkins MA, Casey G,
 665 Brenner H, Chang-Claude J *et al*: Modifiable pathways for colorectal cancer: a mendelian
 666 randomisation analysis. *Lancet Gastroenterol Hepatol* 2020, 5(1):55-62.
- Wagner D, Dias AG, Schnabl K, Van der Kwast T, Vieth R: Determination of 1,25 dihydroxyvitamin D concentrations in human colon tissues and matched serum samples.
 Anticancer Res 2012, 32(1):259-263.
- Ryynanen J, Neme A, Tuomainen TP, Virtanen JK, Voutilainen S, Nurmi T, de Mello VD,
 Uusitupa M, Carlberg C: Changes in vitamin D target gene expression in adipose tissue
 monitor the vitamin D response of human individuals. *Mol Nutr Food Res* 2014,
 58(10):2036-2045.
- 674 46. Gerke AK, Pezzulo AA, Tang F, Cavanaugh JE, Bair TB, Phillips E, Powers LS, Monick MM:
 675 Effects of vitamin D supplementation on alveolar macrophage gene expression:
 676 preliminary results of a randomized, controlled trial. *Multidiscip Respir Med* 2014, 9(1):18.
- 47. Hossein-nezhad A, Spira A, Holick MF: Influence of vitamin D status and vitamin D3
 678 supplementation on genome wide expression of white blood cells: a randomized double679 blind clinical trial. *PLoS One* 2013, 8(3):e58725.
- 680 48. Saksa N, Neme A, Ryynanen J, Uusitupa M, de Mello VD, Voutilainen S, Nurmi T, Virtanen JK,
 681 Tuomainen TP, Carlberg C: Dissecting high from low responders in a vitamin D3
 682 intervention study. J Steroid Biochem Mol Biol 2015, 148:275-282.
- 49. Buckmire M, Parquet G, Seeburger JL, Fukuchi SG, Rolandelli RH: Effect of bowel
 by preparation and a fiber-free liquid diet on expression of transforming growth factor and
 procollagen in colonic tissue preoperatively and postoperatively. *Dis Colon Rectum* 1998,
 41(10):1273-1280.





n = 191

Rectal biopsy NM – microarray ± RNA-seq

Blood Plasma - LC/MS 25-OHD assay **PBMC** - expression microarray Blood - genotyping array





n = 50

Rectal biopsy NM – microarray/ RNA-seq

Blood

Plasma - LC/MS 25-OHD assay **PBMC** - expression microarray



Ο



SCOVIDS study

Blood expression change with supplementation

HIPK2 as biomarker of gene-set enrichment

PPP1CC as biomarker of gene-set enrichment

