

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/111324/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Bakhsh, Ameen D., Ladas, Ioannis, Hamshere, Marian L. , Bullock, Martyn, Kirov, George , Zhang, Lei , Taylor, Peter N. , Gregory, John W. , Scott-Coombes, David, Völzke, Henry, Teumer, Alexander, Mantripragada, Kiran , Williams, E. Dillwyn , Clifton-Bligh, Roderick J., Williams, Nigel M. and Ludgate, Marian E. 2018. An InDel in Phospholipase-C-B-1 is linked with euthyroid multinodular goiter. *Thyroid* 28 (7) , pp. 891-901. 10.1089/thy.2017.0312

Publishers page: <https://doi.org/10.1089/thy.2017.0312>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



1 **An InDel in Phospholipase-C-B-1 is linked with euthyroid multinodular goiter**

2

3 *Ameen D Bakhsh,¹ *Ioannis Ladas,¹ Marian L Hamshere,² Martyn Bullock,³ George Kirov,² Lei Zhang,¹
4 Peter N Taylor,¹ John W. Gregory,¹ David Scott-Coombes,⁴ Henry Völzke,⁶ Alexander Teumer,⁶ Kiran
5 Mantripragada,² E Dillwyn Williams,⁵ Roderick J Clifton-Bligh,³ Nigel M Williams,² Marian E Ludgate,¹.

6 * Authors contributed equally ¹**Division of Infection & Immunity, School of Medicine, Cardiff**
7 **University, Cardiff CF14 4XN, UK. ²Institute of Psychological Medicine & Clinical Neurosciences,**
8 **School of Medicine, Cardiff University, Cardiff CF24 4HQ. UK. ³Kolling Institute of Medical**
9 **Research and Dept Endocrinology, University of Sydney, Royal North Shore Hospital, Sydney, NSW**
10 **2065, Australia. ⁴Department of Endocrine Surgery, University Hospital of Wales, Cardiff CF14**
11 **4XW, UK. ⁵Strangeways Research Laboratory, Cambridge CB1 8RN, UK. ⁶Institute for Community**
12 **Medicine, Study of Health in Pomerania, Ernst Moritz Arndt University, D-17475 Greifswald,**
13 **Germany.**

14

15 **Corresponding Author & address for reprints:**

16 **Dr. Ioannis Ladas**

17 **Harvard Institutes of Medicine,**

18 **Department of Radiation Oncology,**

19 **Dana Farber Cancer Institute,**

20 **4 Blackfan Circle, Lab 342 - Boston, MA**

21 **Phone +1 9786895751; email Ioannis_Ladas@dfci.harvard.edu**

22 **Abbreviated title:** PLCB1 Intronic deletion linked with MNG

23 **Key Words:** Multinodular goiter; genome-wide linkage analysis; copy-number variation; next generation
24 sequencing

25 **Word Count:** 3731 [excluding abstract, references & legends]

26

27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

Abstract

Euthyroid multinodular goiter (MNG) is common but little is known about the genetic variation conferring predisposition. Previously we reported a family with MNG of adolescent onset in which some family members developed papillary thyroid carcinomas (PTC). We conducted a genome-wide linkage analysis and next generation sequencing to identify genetic variants that may confer disease predisposition. A multipoint nonparametric LOD score of 3.01 was obtained covering 19 cM on chromosome 20p. Haplotype analysis reduced the region of interest to 10 cM; analysis of copy number variation identified an intronic InDel (~1000 bp) in the *PLCBI* gene in all 8 affected family members and carriers (an unaffected person who has inherited the genetic trait); this InDel is present in ~1% of ‘healthy’ Caucasians. Next generation sequencing of the region identified no additional disease-associated variant, suggesting a possible role of the InDel. Since *PLCBI* contributes to thyrocyte growth regulation, we investigated the InDel in relevant Caucasian cohorts. It was detected in 0/70 PTC but 4/81 unrelated subjects with MNG [3 F, age at thyroidectomy 27-59 years, no family history of MNG/PTC]. The InDel frequency is significantly higher in MNG subjects compared with controls; $X^2 = 5.076$, $p = 0.024$. *PLCBI* transcript levels were significantly higher in thyroids with the InDel than without ($p < 0.02$).

The intronic *PLCBI* InDel is the first variant found in familial multiple papilloid adenomata-type MNG and in a subset of patients with sporadic MNG. It may function through over-expression and increased PLC activity has been reported in thyroid neoplasms. The potential role of the deletion as a biomarker to identify MNG patients more likely to progress to PTC merits exploration.

49 **Introduction**

50 Euthyroid multinodular goiter (MNG) is common and affects at least 4% of the population, although the
51 prevalence varies with ethnicity and the detection method employed (1). Furthermore, nodular goiter is far
52 more prevalent in iodine deficient regions (2). Although solitary nodules are considered a risk for thyroid
53 cancer (3) the situation for MNG is more controversial (4); the reported increase in the incidence of some
54 thyroid cancers (5) may, in part, be due to increased use of diagnostic tools (6). *BRAF* mutations causing
55 constitutive activation are the most frequent driver of papillary thyroid cancer (PTC) (7). Several genetic
56 variations lead to sporadic thyroid cancers including, among others, *RET* chromosomal re-arrangements
57 (8), translocations between chromosome 2 and 3 generating a PPAR γ -PAX8 fusion protein (9), mutations
58 in *RAS* genes (10) and poly-alanine tract length variation in *FOXEI* (11, 12).

59 Familial non-medullary thyroid cancers account for about 5% of thyroid cancers and have a younger age
60 of onset than sporadic disease. They are associated with 4 susceptibility loci (13-16) on chromosomes
61 19p13.2, 2q21, 1q21 and 10q23 (*PTEN*). There is some overlap with familial goiter in which 8 predisposing
62 loci have been identified (12, 17-20) on chromosomes Xp22, 3q26, 2q, 3p, 7q, 8p 14q13.3 and 14q32, the
63 last two including the *NKX2.1* (21) and the RNase *DICER1* genes respectively (22). A role for the
64 predisposing loci on chromosomes 2q.35, 5q.24, 8p.12 and 14q.13 has been confirmed in Chinese families
65 (23). Genes implicated in familial goiter and cancer generally differ from those in sporadic disease, with
66 the exception of *NKX2.1* (21) and *FOXEI* (24).

67 Previously, we reported a family (25) exhibiting a type of euthyroid MNG with papillary adenomas of
68 adolescent onset affecting 8 individuals in 4 generations to date. MNG is known to have progressed to PTC
69 in 2 of the 8 affected family members. We applied microsatellite analysis to exclude loci described above
70 on chromosomes 14q, Xp, 3q 9p, 2q and 1q. Since one family member had co-existing breast cancer and
71 another co-existing kidney disease we investigated genes co-expressed in these tissues and the thyroid, *NIS*
72 and *PAX8* respectively. Sanger sequencing revealed no abnormality in either gene. Subsequently, the *PTEN*
73 gene has been fully sequenced in the family member with breast cancer and no mutations were detected.

74 The aim of this study was to apply genome-wide linkage analysis (GWLA) and next generation
75 sequencing to identify the gene variant(s) responsible for the observed phenotype in this family. We then
76 aimed to assess the frequency of any variant(s) detected in other relevant cohorts.

77

78 **Subjects and Methods**

79

80 *Genome-Wide Linkage Analysis (GWLA)*

81 We undertook a GWLA of the family described in (25) and summarized in figure 1.

82 All patient samples were obtained with informed consent and Local Research Ethics Committee (LREC)
83 approval. Genomic DNA was extracted from whole blood from 18 family members (those labelled in the
84 tree) of whom 8 were affected (7 females, 1 male), according to the manufacturer's instruction (Qiagen)
85 and quantified using a Nanodrop. Samples (250 ng) were processed following the manufacturer's protocol
86 and the DNA integrity monitored by agarose gel electrophoresis before being hybridized at 48°C for 18
87 hours to Affymetrix Genechip™ Human Mapping 10K 2.0 Arrays. The chips were scanned using an
88 Affymetrix GeneChip scanner 3000; data were acquired using GCOS and analyzed using GTYPE software
89 respectively.

90 Two quality control steps were performed; the first eliminated SNPs showing 'no call' in more than 4
91 individuals. The second step would have eliminated data from any individual with >10% 'no calls', but this
92 did not apply and the data of all 18 family members were retained. Graphical Representation of
93 Relationships (GRR) software was used to determine how many alleles are shared [identity by state (IBS)]
94 at each locus. Mendelian errors were tested using PedCheck software. PLINK, was used to merge family
95 data (founders) with HapMap to investigate ethnicity. Multidimensional scaling (MDS) was performed on
96 the family merged with HapMap data from 60 European individuals (CEU), 90 Chinese (CHB) & Japanese
97 (JPT), & 60 Yoruba (YRI). The family were closest to the European cluster (data not shown) thus allele
98 frequencies were based on CEU HapMap data. Using MERLIN software, the primary analysis was multi-
99 point non-parametric and the secondary analysis multipoint parametric dominant mode assuming 90%

100 penetrance in females, 50% in males and age of onset later than 12 years (based on clinical information
101 summarized in figure 1). Single point analyses were also used to support the findings of multipoint analysis.
102 Since data are derived from a single large family, there is considerable allele sharing and hence the Kong
103 and Cox exponential (--exp) model was used (for non-parametric analysis) (26).

104

105 *Haplotype Analysis*

106 MERLIN software (--best) was also used to perform a haplotype analysis in the region of maximum LOD
107 score on chromosome 20. The haplotype was also confirmed manually.

108

109 *Copy Number Variation Analysis (CNV)*

110 Genomic DNA for CNV analysis of the index patient was quantified and prepared for hybridization to
111 Illumina Human 660W-Quad BeadChips according to the manufacturer's instructions. Data were analyzed
112 using PennCNV (27) software; CNVs were required to be 1 kb long and cover at least 10 consecutive
113 markers (SNP or cnvi) to be considered positive. We focused on the region with a high LOD score identified
114 in the GWLA.

115

116 *Next Generation Sequencing (NGS)*

117 Primer pools for preparation of DNA libraries were designed using Ampliseq 3.0.1 software
118 (<https://ampliseq.com/>) according to the manufacturer's protocol. A total of 429 primers were designed
119 generating 100-300 bp amplicons. The primer pools (details in supplemental table 2) covered the exome
120 sequences (all coding regions, intron/exon boundaries, proximal promoters and 3' untranslated regions) of
121 a region spanning from chr20: 8113337 to 11907302. Approximately 10 ng of the genomic DNAs of interest
122 were amplified according to the manufacturer's instructions. The amplified samples were partially digested
123 by FuPa reagent (Life Technologies) and ligated with barcode/adaptor mix. DNA libraries were then
124 purified using Agencourt AMPure XP beads (Beckman Coulter), quantified by qPCR and adjusted to a final
125 concentration of 100 pM, combined and prepared for Emulsion PCR with Ion OneTouch 2 (Life

126 Technologies). Following enrichment, the ion sphere particles were loaded onto an Ion PI Chip V2 and
127 sequenced by Ion Torrent Proton sequencer. Sequencing data were analyzed by Ion Torrent Suite software
128 (4.4.2), using the plug-in variant caller (v 4.2.10) and configuration with generic Personal Genome Machine
129 (PGM) germ line settings and high stringency analysis mode.

130 NGS was performed on 98 individuals, all 18 family members plus 80 unrelated subjects with MNG (please
131 see below).

132 Other variants identified in the family using NGS were interrogated in the SHIP cohort (Study of Health in
133 Pomerania) (28). Relevant genotyping data were available from 986 individuals who were either unaffected
134 or presented with diffuse goiter (as defined in (29)) and/or MNG (nodules identified by ultrasound). Figure
135 2 details the filtering steps and evaluations undertaken to assess whether detected variants might be linked
136 with disease.

137

138 *Defining deletion frequency*

139 Primers within and flanking the deleted region were designed using Primer 3 software (supplemental table
140 2) for PCR amplification of genomic DNA from all family members and 105 unrelated euthyroid
141 individuals from the UK. PCR amplicons were analyzed by agarose gel electrophoresis and PEG
142 precipitated for Sanger sequencing using Big Dye Terminator Cycle Sequencing Ready Reaction (ABI
143 Prism, PE Biosystems) and analysis on an ABI 3100 Genetic Analyser.

144 Tissues from patients recruited in Australia (snap frozen and stored in liquid nitrogen) were also studied
145 and consisted of 70 PTC and 81 MNG patients. [Ethics approval from the Northern Sydney Area Health
146 Service Human Research Ethics Committee]. To avoid population stratification, only subjects with self-
147 reported white European ancestry were included; patient data and tissues were collected between 1992 and
148 2012 at the Kolling Institute of Medical Research. Genomic DNA for genotyping was obtained from thyroid
149 tissue using Qiagen kits and analyzed by PCR and Sanger sequencing as described above; these samples
150 also underwent NGS.

151

152 *High Throughput Screening of PLCB1 InDel, analysis of additional cohorts.*

153 We developed a qPCR based genotyping tool using primers within and flanking the *PLCB1* InDel as
154 described above (Supplementary table 2). The genotyping tool was used to screen 200 breast cancer
155 patients. Initial optimization experiments revealed that greatest specificity was obtained using primers
156 flanking the InDel. The qPCR obtained a difference of approximately 10 Ct for samples with and without
157 the InDel. The qPCR was performed with approximately 100 ng Genomic DNA Input, 1x SyBR green
158 master qPCR mix (Invitrogen) and 100 nM of each primer in a 25 µl reaction. QPCR conditions included
159 an initial hold step at 50°C for 2 minutes, then 95°C for 2 minutes followed by 40 cycles of 95°C for 15
160 seconds and 60°C for 30 seconds then a hold step at 95°C for 1 minute, 55°C for 30 seconds and 95°C for
161 30 minutes. Samples found to harbor the InDel by qPCR were confirmed by Sanger sequencing.

162

163 *Transcript measurements of PLCB1 isoforms*

164 Thyroid tissue was obtained from 3 affected family members heterozygous for the InDel and five subjects
165 undergoing thyroidectomy for autoimmune thyroid disease expressing two normal *PLCB1* alleles (all
166 confirmed by genotyping). Thyroid RNA was extracted, reverse transcribed using standard protocols and
167 qPCR (SYBR Green incorporation measured on a Stratagene MX 3000) was used to measure transcript
168 levels and evaluate proportions of *PLCB1-a* and *PLCB1-b* isoforms (primers in supplemental table 2, wild
169 type amplicon identity confirmed by Sanger sequencing). Comparison with standard curves for transcript
170 levels of isoform 1a and 1b permitted calculations of absolute values for each sample. Transcripts for a
171 housekeeping gene (*APRT*) were also measured and values were expressed relative to this (transcripts/1000
172 *APRT*). In a single qPCR experiment, all measurements were made in duplicate; the standard curve was
173 also run in each reaction. Transcript levels of the various *PLCB1* isoforms were compared between deletion
174 affected and non-affected thyroids using the Mann Whitney U test and differences where $p < 0.05$ taken to
175 be significant.

176

177 **Results**

178

179 *Genome wide linkage, haplotype & copy number variation analyses*

180 We obtained a multipoint nonparametric LOD score of 3.01 over 19.5 cM on chromosome 20p (figure 3
181 and supplementary figure 1). In secondary analysis, the same region gave a multipoint dominant LOD score
182 of 2.16, based on a disease model with 0.01 allele frequency, 50% penetrance for males and 90% for
183 females, both age >12. LOD scores on the remaining 21 autosomes and X chromosome were all below 1
184 (figure3). Single-point analyses supported the multipoint data for both nonparametric and model-based
185 linkage on all chromosomes (supplementary table1).

186 Haplotype analysis was employed to identify a possible disease locus and reduced the region of interest to
187 8.73 cM (3.7 Mbp), which includes 10 genes (supplemental figure 2 and 3). The haplotype was not found
188 in 503 individuals from the 1000 genome European dataset, although one individual missed only the last
189 marker suggesting a shorter version of the haplotype (red highlight in supplementary figure 3a).

190 Analysis of copy number variation in an affected individual revealed a deletion of ~900 bp located in the
191 3rd intron in one copy of *phospholipase-C B1 (PLCB1)* in the region of interest (supplementary figure 4;
192 the log R ratio mean was -0.451, over 14 markers, with at least one marker below -1.00).

193

194 *Defining the deletion frequency in the family and selected cohorts*

195 The length of the deletion was confirmed to be 1077 bp by standard PCR and Sanger sequencing, using
196 primers flanking and within the deletion, to reveal one copy of full-length and one deleted allele in all
197 affected and obligate carrier II-3 but only the full-length product in family members free of any sign of
198 MNG. The sequence of the allele bearing the deletion corresponds to that immediately upstream and
199 downstream of the deleted region but with an additional 'ATAA' inserted at the junction, hence it is an
200 InDel.

201 Standard PCR was applied to genotype a selected cohort of 105 Caucasians in whom thyroid function
202 testing was clinically indicated because of general fatigue. A woman in her forties, with no history of

203 thyroid disease, was heterozygous for the InDel. Further *in silico* analyses, using the database for genomic
204 variants (30) identified a report which detected the InDel (variation 67651, LRR -0.645) in 2 of 180
205 Caucasians but none in more than 450 people of other ethnicities (31). Combining our genotyping data with
206 that of Conrad *et al.* (31) reveals 3 in 285 Caucasians harboring the InDel, suggesting that it is relatively
207 rare (~1%).

208 Subsequently, genomic DNA was extracted from thyroid tissue from 70 patients undergoing surgery for
209 non-familial PTC and an additional 81 operated for non-familial MNG. We used PCR analysis to test for
210 the InDel, as described above. The InDel was not detected in any of the PTC patients but 4 of the 81 MNG
211 were heterozygous for the InDel and sequencing revealed the same ATAA insertion at the junction.
212 Comparison of the frequency of the InDel in the general population with that in MNG gives a X^2 value of
213 5.076 (1 degree of freedom), $p= 0.024$ (two-tailed). The 4 MNG patients (3 women, 1 man) are unrelated
214 and with no apparent family history of MNG or PTC at the time of their surgery. The age at thyroidectomy
215 was between 27 to 59 years and the pathology is variously described as ‘oncocytic neoplasm with variable
216 patterns of growth’ to ‘cystic degeneration with calcification’. We also investigated whether the *PLCBI*
217 InDel might be implicated in breast cancer using the qPCR-based screening protocol. Prevalence in this
218 cohort was similar to that of the general population, i.e. 1%, since just 2 breast cancer patients harbored the
219 *PLCBI* InDel.

220

221 *Next Generation Sequencing of the Chr20 high LOD score region*

222 The Proton Sequencer generated 9.9 Gbp of data, achieving 98% accurately mapped sequences with >88%
223 of the percentage of target bases covered by at least 0.2 times the average base read depth.

224 A total of 181 sequence variants between Chr20 8113405 and 11907285 were identified in the family with
225 the minor allele being on the disease risk haplotype in 12 of these. Given the rarity of PTC and the expected
226 high penetrance, we expect a pathogenic variant to have a very low population frequency. After referring
227 to the UCSC genome browser, only 1 of the 12 variants was found to have a minor allele frequency <1%;
228 its presence in affected family members was confirmed by Sanger sequencing. The variant is at Chr20

229 10036484 (rs56234782) with T (98.8%) or C (1.2%) in the 3' UTR of the *ANKRD5* gene. To investigate
230 whether it is implicated in goiter and/or thyroid nodule formation, we investigated its frequency in the SHIP
231 cohort. However, even though the minor allele was more prevalent in the entire cohort, the prevalence in
232 the affected population (goiters 1.9% and nodules 2.54%) was lower than in the unaffected populations
233 (2.79% and 2.85% respectively), thereby excluding a role for it in MNG.

234 The MNG cohort was also submitted to NGS analysis. This identified more than 300 different sequence
235 variants across the 80 patients, however, all were also present in the 1000 genomes cohort at a population
236 frequency >1%. We therefore considered it unlikely that any of these variants are pathologically relevant
237 to MNG, thereby confirming the relevance of the InDel.

238

239 *Transcript measurements of PLCB1 & effect of knock-down on thyroid growth*

240 Having confirmed that the InDel may contribute to the pathogenesis for MNG (perhaps in combination with
241 other factors), we investigated how it might promote thyrocyte proliferation. The InDel is in the large 3rd
242 intron of *PLCB1*, the phosphoinositide-specific enzyme which generates IP3 and DAG leading to PKC
243 activation and also links signaling between the MAPK cascade and G protein coupled receptors (32).
244 *PLCB1* is present in several isoforms including *PLCB1-a* and *PLCB1-b*, with the latter having a
245 predominantly nuclear location (33). To test the hypothesis that the InDel causes preferential transcription
246 of certain *PLCB1* isoforms, RNA was extracted from thyroids from the original family and from subjects
247 undergoing thyroidectomy for benign disease. In all cases genomic DNA from the donor thyroid was tested
248 for the *PLCB1* deletion.

249 QPCR analysis of InDel-affected thyroids did not indicate altered expression of the major *PLCB1* isoforms
250 a and b (sequenced to confirm they were wild type, data not shown). However, qPCR measurements
251 indicated significantly higher *PLCB1* transcript levels ($p < 0.02$) in thyroids from family members with the
252 InDel, compared with those from benign thyroid disease who do not harbor the variant (figure 5). Lack of
253 thyroid tissue precluded analyzing *PLCB1* protein levels.

254

255

256 **Discussion**

257 Our GWLA led to the identification of an InDel in the family with a type of MNG, located in the large third
258 intron of *PLCBI*, a gene encoding an enzyme with a central role in several signaling cascades involved in
259 regulating thyrocyte growth. Subsequent NGS in the family failed to identify any other disease-linked
260 variant, thus supporting a role for the *PLCBI* InDel in the pathogenesis of MNG in this family.

261 The InDel comprises the loss of 1077 bp with an ATAA inserted at the junction in all affected family
262 members and the 4 unrelated patients with MNG. We suggest that this may indicate a ‘cut and paste’ event
263 indicating transposon activity. Interestingly, a 11-kb transposon cluster has been identified immediately
264 upstream of the 3.7 Mbp section on chr 20 displaying a non-parametric LOD score of 3.01 in the current
265 study (34). Of note the LOD score of 3.01, whilst at the lower limit to be considered significant, is higher
266 than the maximum estimated for a kindred having 8 affected individuals (35).

267 We detected the same InDel in 1 subject of a selected cohort of 105 people in whom measuring thyroid
268 function was clinically indicated. We also consulted the database of genomic variants and found several
269 reports of relevance. Conrad et al. found the deletion in 2 of 180 Caucasians but insufficient detail is
270 provided to know whether it is a simple CNV or the same InDel identified in our studies. Combining our
271 genotyping data with that of Conrad et al. reveals that 3 in 285 Caucasians harbor the deletion, suggesting
272 that it is rare (31). Several other authors did not observe this deletion, but aware of the difficulty in detecting
273 small CNVs, we did not include these in our calculation. In addition, 200 patients with breast cancer have
274 been screened for the InDel with only two harboring this deletion. Hence, the prevalence was similar to the
275 general population suggesting that there is no connection of the InDel with breast cancer.

276 We then considered how the deletion or novel *PLCBI* InDel might exert its effects. The region was
277 explored using the Encyclopedia of DNA elements (ENCODE) (although compiled without inclusion of
278 thyroid tissue or cell lines) (36), which revealed the existence of a binding site for the estrogen receptor
279 alpha (ER α) within the deletion. This is of potential importance since all thyroid diseases are more prevalent
280 in women than men (1). The incidence of thyroid disorders increases in the years immediately following

281 puberty and *in vitro* studies have demonstrated that estrogen can promote thyrocyte proliferation (37) by
282 several mechanisms. The *PLCBI* InDel is located in an intron; while many functional transcription factor
283 binding sites are found in promoters, a systematic search for ER α binding sites in the human genome
284 identified >1000 with >95% of them residing in introns and not promoters (38).

285 We also conducted experiments to determine whether the deletion alters the ratio of *PLCBI-a* and *PLCBI-*
286 *b*, which are generated by alternative splicing. Differences in their C terminal sequence mean that only
287 *PLCBI-a* has a nuclear export signal. We found no alteration in the ratio of *PLCBI-a* and *b* isoforms but
288 in all cases transcript levels for *PLCBI* were higher in thyroids from people heterozygous for the InDel
289 than in thyroids with two full-length copies. This suggests that the InDel may contribute to MNG
290 development through overexpression of *PLCBI*. Furthermore, total PLC enzyme activity is elevated in
291 thyroid neoplasms (39) but unfortunately PLC inhibitors lack the specificity required to identify which
292 isoform is responsible. Increased *PLCBI* expression has also been reported in small cell lung carcinoma
293 (40) and expression of *PLCB2* is substantially increased in breast cancer and is used as a prognostic marker
294 (40).

295 As mentioned above, PLC enzymes activate PKC and genes implicated in this signal pathway are
296 upregulated in euthyroid MNG (41). They also link signaling via Gq (which can also be activated via the
297 thyrotropin receptor) to the MAPK cascade and in the thyroid disruption of this pathway, by thyrocyte-
298 targeted Cre/Lox P knock-down of the Gq α subunit, produces mice which are resistant to goiter formation
299 when fed a goitrogenic diet (42). However, when we performed western blots with protein extracts of
300 thyroid tissue from family members with the *PLCBI* InDel we were surprised to observe that pMAPK
301 levels were substantially lower than in thyroid tissue from patients with autoimmune thyroid disease or
302 MNG without the *PLCBI* InDel (Supplementary Figure 5).

303 In conclusion, the *PLCBI* InDel identified in this family with MNG also occurs in a proportion of sporadic
304 MNG, and may provide a biomarker to identify MNG patients more likely to progress to PTC. The *PLCBI*
305 InDel appears to predispose to goiter formation, possibly by increasing *PLCBI* transcription with
306 subsequent downstream effects.

307

308 **Supplemental Data**

309 The supplemental data comprises 5 figures and 2 tables;

310 Supplemental Figure 1; LOD scores of all Chromosomes

311 Supplementary Figure 2; Genes in high LOD score region chromosome 20

312 Supplemental Figure 3; Haplotype Frequency in 1000 genomes European dataset

313 Supplementary Figure 4; Copy number variation in high LOD score region chromosome 20

314 Supplementary Figure 5; Densitometry ratios for pERK/total ERK

315 Supplemental Table 1; Single point LOD scores all chromosomes

316 Supplemental Table 2; Primers used for NGS and to define deletion frequency

317

318

319 **Web Resources**

320 The March 2006 human reference sequence (NCBI Build 36.1) produced by the International Human

321 Genome Sequencing Consortium, was used as a reference genome (UCSC Genome Browser;[http://genome-](http://genome-euro.ucsc.edu/cgi-bin/hgGateway?hgsid=192302910&clade=mammal&org=Human&db=hg18)

322 [euro.ucsc.edu/cgi-bin/hgGateway?hgsid=192302910&clade=mammal&org=Human&db=hg18](http://genome-euro.ucsc.edu/cgi-bin/hgGateway?hgsid=192302910&clade=mammal&org=Human&db=hg18)).

323

324 **Acknowledgements**

325 We express our sincere gratitude to the members of the family who participated in this research.

326 The work was part funded by the Government of Saudi Arabia (ref A390), by the Medical Research Council

327 and the Onassis Foundation.

328 SHIP is part of the Community Medicine Research Network of the University Medicine Greifswald,

329 Germany (www.community-medicine.de).

330 Genomic DNA from patients with breast cancer was provided by Dr Florentia Fostira from the National

331 Center for Scientific Research Demokritos (Athens, Greece).

332

333 **Declaration of interest**

334 There is no conflict of interest that could be perceived as prejudicing the impartiality of the research
335 reported.

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358 **References**

- 360 1. Vanderpump MP, Tunbridge WM, French JM, Appleton D, Bates D, Clark F, Grimley
361 Evans J, Hasan DM, Rodgers H, Tunbridge F 1995 The incidence of thyroid disorders in
362 the community: a twenty-year follow-up of the Whickham Survey. *Clin Endocrinol* 43:55-
363 68.
- 364 2. Carle A, Krejbjerg A, Laurberg P 2014 Epidemiology of nodular goitre. Influence of iodine
365 intake. *Best Pract Res Clin Endocrinol Metab* 28:465-479.
- 366 3. Frates MC, Benson CB, Doubilet PM, Kunreuther E, Contreras M, Cibas ES, Orcutt J,
367 Moore FD, Jr., Larsen PR, Marqusee E, Alexander EK 2006 Prevalence and distribution
368 of carcinoma in patients with solitary and multiple thyroid nodules on sonography. *J Clin*
369 *Endocrinol Metab* 91:3411-3417.
- 370 4. Fiore E, Rago T, Provenzale MA, Scutari M, Ugolini C, Basolo F, Di Coscio G, Berti P,
371 Grasso L, Elisei R, Pinchera A, Vitti P 2009 Lower levels of TSH are associated with a
372 lower risk of papillary thyroid cancer in patients with thyroid nodular disease: thyroid
373 autonomy may play a protective role. *Endocr Relat Cancer* 16:1251-1260.
- 374 5. Kilfoy BA, Zheng T, Holford TR, Han X, Ward MH, Sjodin A, Zhang Y, Bai Y, Zhu C,
375 Guo GL, Rothman N, Zhang Y 2009 International patterns and trends in thyroid cancer
376 incidence, 1973-2002. *Cancer Causes Control* 20:525-531.
- 377 6. La Vecchia C, Negri E 2017 Thyroid cancer: The thyroid cancer epidemic - overdiagnosis
378 or a real increase? *Nat Rev Endocrinol* 13:318-319.
- 379 7. Kimura ET, Nikiforova MN, Zhu ZW, Knauf JA, Nikiforov YE, Fagin JA 2003 High
380 prevalence of BRAF mutations in thyroid cancer: Genetic evidence for constitutive
381 activation of the RET/PTC-RAS-BRAF signaling pathway in papillary thyroid carcinoma.
382 *Cancer Res* 63:1454-1457.
- 383 8. Grieco M, Santoro M, Berlingieri MT, Melillo RM, Donghi R, Bongarzone I, Pierotti MA,
384 Dellaporta G, Fusco A, Vecchio G 1990 PTC is a novel rearranged form of the RET proto-
385 oncogene and is frequently detected in-vivo in human thyroid papillary carcinomas. *Cell*
386 60:557-563.
- 387 9. Kroll TG, Sarraf P, Pecciarini L, Chen CJ, Mueller E, Spiegelman BM, Fletcher JA 2000
388 PAX8-PPAR gamma 1 fusion in oncogene human thyroid carcinoma. *Science* 289:1357-
389 1360.
- 390 10. Lemoine NR, Mayall ES, Wyllie FS, Farr CJ, Hughes D, Padua RA, Thurston V, Williams
391 ED, Wynfordthomas D 1988 Activated RAS oncogenes in human thyroid cancers. *Cancer*
392 *Res* 48:4459-4463.
- 393 11. Gudmundsson J, Sulem P, Gudbjartsson DF, Jonasson JG, Sigurdsson A, Bergthorsson JT,
394 He H, Blondal T, Geller F, Jakobsdottir M, Magnusdottir DN, Matthiasdottir S, Stacey SN,

- 405 Skarphedinsson OB, Helgadóttir H, Li W, Nagy R, Aguillo E, Faure E, Prats E, Saez B,
406 Martinez M, Eyjólfsson GI, Björnsdóttir US, Holm H, Kristjánsson K, Frigge ML,
407 Kristvinsson H, Gulcher JR, Jonsson T, Rafnar T, Hjartarsson H, Mayordomo JI, de la
408 Chapelle A, Hrafnkelsson J, Thorsteinsdóttir U, Kong A, Stefansson K 2009 Common
409 variants on 9q22.33 and 14q13.3 predispose to thyroid cancer in European populations.
410 Nat Genet 41:460-464.
411
- 412 **12.** Bullock M, Duncan EL, O'Neill C, Tacon L, Sywak M, Sidhu S, Delbridge L, Learoyd D,
413 Robinson BG, Ludgate M, Clifton-Bligh RJ 2012 Association of FOXE1 polyalanine
414 repeat region with papillary thyroid cancer. J Clin Endocrinol Metab 97:E1814-1819.
415
- 416 **13.** Canzian F, Amati P, Harach HR, Kraimps JL, Lesueur F, Barbier J, Levillain P, Romeo G,
417 Bonneau D 1998 A gene predisposing to familial thyroid tumors with cell oxyphilia maps
418 to chromosome 19p13.2. Am J Hum Genet 63:1743-1748.
419
- 420 **14.** Malchoff CD, Sarfarazi M, Tendler B, Forouhar F, Whalen G, Joshi V, Arnold A, Malchoff
421 DM 2000 Papillary thyroid carcinoma associated with papillary renal neoplasia: genetic
422 linkage analysis of a distinct heritable tumor syndrome. J Clin Endocrinol Metab 85:1758-
423 1764.
424
- 425 **15.** McKay JD, Lesueur F, Jonard L, Pastore A, Williamson J, Hoffman L, Burgess J, Duffield
426 A, Papotti M, Stark M, Sobol H, Maes B, Murat A, Kaariainen H, Bertholon-Gregoire M,
427 Zini M, Rossing MA, Toubert ME, Bonichon F, Cavarec M, Bernard AM, Boneu A, Leprat
428 F, Haas O, Lasset C, Schlumberger M, Canzian F, Goldgar DE, Romeo G 2001
429 Localization of a susceptibility gene for familial nonmedullary thyroid carcinoma to
430 chromosome 2q21. Am J Hum Genet 69:440-446.
431
- 432 **16.** Frisk T, Foukakis T, Dwight T, Lundberg J, Hoog A, Wallin G, Eng C, Zedenius J, Larsson
433 C 2002 Silencing of the PTEN tumor-suppressor gene in anaplastic thyroid cancer. Genes
434 Chromosom Cancer 35:74-80.
435
- 436 **17.** Capon F, Tacconelli A, Giardina E, Sciacchitano S, Bruno R, Tassi V, Trischitta V, Filetti
437 S, Dallapiccola B, Novelli G 2000 Mapping a dominant form of multinodular goiter to
438 chromosome Xp22. Am J Hum Genet 67:1004-1007.
439
- 440 **18.** Takahashi T, Nozaki J, Komatsu M, Wada Y, Utsunomiya M, Inoue K, Takada G, Koizumi
441 A 2001 A new locus for a dominant form of multinodular goiter on 3q26.1-q26.3. Biochem
442 Biophys Res Commun 284:650-654.
443
- 444 **19.** Bayer Y, Neumann S, Meyer B, Ruschendorf F, Reske A, Brix T, Hegedus L, Langer P,
445 Nürnberg P, Paschke R 2004 Genome-wide linkage analysis reveals evidence for four new
446 susceptibility loci for familial euthyroid goiter. J Clin Endocrinol Metab 89:4044-4052.
447
- 448 **20.** Bignell GR, Canzian F, Shayeghi M, Stark M, Shugart YY, Biggs P, Mangion J, Hamoudi
449 R, Rosenblatt J, Buu P, Sun S, Stoffer SS, Goldgar DE, Romeo G, Houlston RS, Narod
450 SA, Stratton MR, Foulkes WD 1997 Familial nontoxic multinodular thyroid goiter locus

- 451 maps to chromosome 14q but does not account for familial nonmedullary thyroid cancer.
452 Am J Hum Genet 61:1123-1130.
453
- 454 **21.** Barnett CP, Mencil JJ, Gecz J, Waters W, Kirwin SM, Vinette KMB, Uppill M, Nicholl J
455 2012 Choreoathetosis, congenital hypothyroidism and neonatal respiratory distress
456 syndrome with intact NKX2-1. Am J Med Genet Part A 158A:3168-3173.
457
- 458 **22.** Frio TR, Bahubeshi A, Kanellopoulou C, Hamel N, Niedziela M, Sabbaghian N, Pouchet
459 C, Gilbert L, O'Brien PK, Serfas K, Broderick P, Houlston RS, Lesueur F, Bonora E, Muljo
460 S, Schimke RN, Bouron-Dal Soglio D, Arseneau J, Schultz KA, Priest JR, Nguyen V-H,
461 Ruben Harach H, Livingston DM, Foulkes WD, Tischkowitz M 2011 DICER1 Mutations
462 in Familial Multinodular Goiter With and Without Ovarian Sertoli-Leydig Cell Tumors. J
463 Am Med Assoc 305:68-77.
464
- 465 **23.** Liao S, Song W, Liu Y, Deng S, Liang Y, Tang Z, Huang J, Dong D, Xu G 2013 Familial
466 multinodular goiter syndrome with papillary thyroid carcinomas: mutational analysis of
467 the associated genes in 5 cases from 1 Chinese family. BMC Endocr Disord 13:48.
468
- 469 **24.** Tomaz RA, Sousa I, Silva JG, Santos C, Teixeira MR, Leite V, Cavaco BM 2012 FOXE1
470 polymorphisms are associated with familial and sporadic nonmedullary thyroid cancer
471 susceptibility. Clin Endocrinol 77:926-933.
472
- 473 **25.** Bakhsh A, Kirov G, Gregory JW, Williams ED, Ludgate M 2006 A new form of familial
474 multi-nodular goitre with progression to differentiated thyroid cancer. Endocr Relat Cancer
475 13:475-483.
476
- 477 **26.** Kong A, Cox NJ 1997 Allele-sharing models: LOD scores and accurate linkage tests. Am
478 J Hum Genet 61:1179-1188.
479
- 480 **27.** Wang K, Li M, Hadley D, Liu R, Glessner J, Grant SF, Hakonarson H, Bucan M 2007
481 PennCNV: an integrated hidden Markov model designed for high-resolution copy number
482 variation detection in whole-genome SNP genotyping data. Genome Res 17:1665-1674.
483
- 484 **28.** Volzke H, Ludemann J, Robinson DM, Spieker KW, Schwahn C, Kramer A, John U, Meng
485 W 2003 The prevalence of undiagnosed thyroid disorders in a previously iodine-deficient
486 area. Thyroid 13:803-810.
487
- 488 **29.** Teumer A, Rawal R, Homuth G, Ernst F, Heier M, Evert M, Dombrowski F, Volker U,
489 Nauck M, Radke D, Ittermann T, Biffar R, Doring A, Gieger C, Klopp N, Wichmann HE,
490 Wallaschofski H, Meisinger C, Volzke H 2011 Genome-wide association study identifies
491 four genetic loci associated with thyroid volume and goiter risk. Am J Hum Genet 88:664-
492 673.
493
- 494 **30.** Iafrate AJ, Feuk T, Van Puymbroeck L, Rivera MN, Listewnik ML, Ying QP, Scherer SW,
495 Lee C 2004 Detection of large-scale variation in the human genome. J Mol Diagn 6:411-
496 411.

- 497
498 **31.** Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, Aerts J, Andrews TD,
499 Barnes C, Campbell P, Fitzgerald T, Hu M, Ihm CH, Kristiansson K, MacArthur DG,
500 MacDonald JR, Onyiah I, Pang AWC, Robson S, Stirrups K, Valsesia A, Walter K, Wei J,
501 Tyler-Smith C, Carter NP, Lee C, Scherer SW, Hurles ME, Wellcome Trust Case C 2010
502 Origins and functional impact of copy number variation in the human genome. *Nature*
503 464:704-712.
504
- 505 **32.** Kadamur G, Ross EM 2013 Mammalian Phospholipase C. In: Julius D, (ed) *Annu Rev*
506 *Physiol* 75:127-154.
507
- 508 **33.** Grubb DR, Vasilevski O, Huynh H, Woodcock EA 2008 The extreme C-terminal region
509 of phospholipase C beta 1 determines subcellular localization and function; the "b" splice
510 variant mediates alpha(1)-adrenergic receptor responses in cardiomyocytes. *Faseb Journal*
511 22:2768-2774.
512
- 513 **34.** Giordano J, Ge Y, Gelfand Y, Abrusan G, Benson G, Warburton PE 2007 Evolutionary
514 history of mammalian transposons determined by genome-wide defragmentation. *PLOS*
515 *Comput Biol* 3:1321-1334.
516
- 517 **35.** Ott J, Wang J, Leal SM 2015 Genetic linkage analysis in the age of whole-genome
518 sequencing. *Nature Rev Genet* 16:275-284.
519
- 520 **36.** Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis C, Doyle F, Epstein CB, Frietze S,
521 Harrow J, Kaul R, Khatun J, Lajoie BR, Landt SG, Lee B-K, Pauli F, Rosenbloom KR,
522 Sabo P, Safi A, Sanyal A, Shores N, Simon JM, Song L, Trinklein ND, Altshuler RC,
523 Birney E, Brown JB, Cheng C, Djebali S, Dong X, Dunham I, Ernst J, Furey TS, Gerstein
524 M, Giardine B, Greven M, Hardison RC, Harris RS, Herrero J, Hoffman MM, Iyer S, Kellis
525 M, Khatun J, Kheradpour P, Kundaje A, Lassmann T, Li Q, Lin X, Marinov GK, Merkel
526 A, Mortazavi A, Parker SCJ, Reddy TE, Rozowsky J, Schlesinger F, Thurman RE, Wang
527 J, Ward LD, Whitfield TW, Wilder SP, Wu W, Xi HS, Yip KY, Zhuang J, Bernstein BE,
528 Birney E, Dunham I, Green ED, Gunter C, Snyder M, Pazin MJ, Lowdon RF, Dillon LAL,
529 Adams LB, Kelly CJ, Zhang J, Wexler JR, Green ED, Good PJ, Feingold EA, Bernstein
530 BE, Birney E, Crawford GE, Dekker J, Elnitski L, Farnham PJ, Gerstein M, Giddings MC,
531 Gingeras TR, Green ED, Guigo R, Hardison RC, Hubbard TJ, Kellis M, Kent WJ, Lieb
532 JD, Margulies EH, Myers RM, Snyder M, Stamatoyannopoulos JA, Tenenbaum SA, Weng
533 Z, White KP, Wold B, Khatun J, Yu Y, Wrobel J, Risk BA, Gunawardena HP, Kuiper HC,
534 Maier CW, Xie L, Chen X, Giddings MC, Bernstein BE, Epstein CB, Shores N, Ernst J,
535 Kheradpour P, Mikkelsen TS, Gillespie S, Goren A, Ram O, Zhang X, Wang L, Issner R,
536 Coyne MJ, Durham T, Ku M, Truong T, Ward LD, Altshuler RC, Eaton ML, Kellis M,
537 Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde
538 J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C,
539 Rozowsky J, Roeder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer
540 MT, Batut P, Bell I, Bell K, Chakraborty S, Chen X, Chrast J, Curado J, Derrien T,
541 Drenkow J, Dumais E, Dumais J, Duttagupta R, Fastuca M, Fejes-Toth K, Ferreira P,
542 Foissac S, Fullwood MJ, Gao H, Gonzalez D, Gordon A, Gunawardena HP, Howald C,

543 Jha S, Johnson R, Kapranov P, King B, Kingswood C, Li G, Luo OJ, Park E, Preall JB,
544 Presaud K, Ribeca P, Risk BA, Robyr D, Ruan X, Sammeth M, Sandhu KS, Schaeffer L,
545 See L-H, Shahab A, Skancke J, Suzuki AM, Takahashi H, Tilgner H, Trout D, Walters N,
546 Wang H, Wrobel J, Yu Y, Hayashizaki Y, Harrow J, Gerstein M, Hubbard TJ, Reymond
547 A, Antonarakis SE, Hannon GJ, Giddings MC, Ruan Y, Wold B, Carninci P, Guigo R,
548 Gingeras TR, Rosenbloom KR, Sloan CA, Learned K, Malladi VS, Wong MC, Barber G,
549 Cline MS, Dreszer TR, Heitner SG, Karolchik D, Kent WJ, Kirkup VM, Meyer LR, Long
550 JC, Maddren M, Raney BJ, Furey TS, Song L, Grasfeder LL, Giresi PG, Lee B-K,
551 Battenhouse A, Sheffield NC, Simon JM, Showers KA, Safi A, London D, Bhinge AA,
552 Shestak C, Schaner MR, Kim SK, Zhang ZZ, Mieczkowski PA, Mieczkowska JO, Liu Z,
553 McDaniell RM, Ni Y, Rashid NU, Kim MJ, Adar S, Zhang Z, Wang T, Winter D, Keefe
554 D, Birney E, Iyer VR, Lieb JD, Crawford GE, Li G, Sandhu KS, Zheng M, Wang P, Luo
555 OJ, Shahab A, Fullwood MJ, Ruan X, Ruan Y, Myers RM, Pauli F, Williams BA, Gertz J,
556 Marinov GK, Reddy TE, Vielmetter J, Partridge EC, Trout D, Varley KE, Gasper C, Bansal
557 A, Pepke S, Jain P, Amrhein H, Bowling KM, Anaya M, Cross MK, King B, Muratet MA,
558 Antoshechkin I, Newberry KM, McCue K, Nesmith AS, Fisher-Aylor KI, Pusey B,
559 DeSalvo G, Parker SL, Balasubramanian S, Davis NS, Meadows SK, Eggleston T, Gunter
560 C, Newberry JS, Levy SE, Absher DM, Mortazavi A, Wong WH, Wold B, Blow MJ, Visel
561 A, Pennachio LA, Elnitski L, Margulies EH, Parker SCJ, Petrykowska HM, Abyzov A,
562 Aken B, Barrell D, Barson G, Berry A, Bignell A, Boychenko V, Bussotti G, Chrast J,
563 Davidson C, Derrien T, Despacio-Reyes G, Diekhans M, Ezkurdia I, Frankish A, Gilbert
564 J, Gonzalez JM, Griffiths E, Harte R, Hendrix DA, Howald C, Hunt T, Jungreis I, Kay M,
565 Khurana E, Kokocinski F, Leng J, Lin MF, Loveland J, Lu Z, Manthravadi D, Mariotti M,
566 Mudge J, Mukherjee G, Notredame C, Pei B, Rodriguez JM, Saunders G, Sboner A, Searle
567 S, Sisu C, Snow C, Steward C, Tanzer A, Tapanari E, Tress ML, van Baren MJ, Walters
568 N, Washietl S, Wilming L, Zadissa A, Zhang Z, Brent M, Haussler D, Kellis M, Valencia
569 A, Gerstein M, Reymond A, Guigo R, Harrow J, Hubbard TJ, Landt SG, Fietze S, Abyzov
570 A, Addleman N, Alexander RP, Auerbach RK, Balasubramanian S, Bettinger K, Bhardwaj
571 N, Boyle AP, Cao AR, Cayting P, Charos A, Cheng Y, Cheng C, Eastman C, Euskirchen
572 G, Fleming JD, Grubert F, Habegger L, Hariharan M, Harmanci A, Iyengar S, Jin VX,
573 Karczewski KJ, Kasowski M, Lacroute P, Lam H, Lamarre-Vincent N, Leng J, Lian J,
574 Lindahl-Allen M, Min R, Miotto B, Monahan H, Moqtaderi Z, Mu XJ, O'Geen H, Ouyang
575 Z, Patacsil D, Pei B, Raha D, Ramirez L, Reed B, Rozowsky J, Sboner A, Shi M, Sisu C,
576 Slifer T, Witt H, Wu L, Xu X, Yan K-K, Yang X, Yip KY, Zhang Z, Struhl K, Weissman
577 SM, Gerstein M, Farnham PJ, Snyder M, Tenenbaum SA, Penalva LO, Doyle F, Karmakar
578 S, Landt SG, Bhanvadia RR, Choudhury A, Domanus M, Ma L, Moran J, Patacsil D, Slifer
579 T, Victorsen A, Yang X, Snyder M, White KP, Auer T, Centanin L, Eichenlaub M, Gruhl
580 F, Heermann S, Hoeckendorf B, Inoue D, Kellner T, Kirchmaier S, Mueller C, Reinhardt
581 R, Schertel L, Schneider S, Sinn R, Wittbrodt B, Wittbrodt J, Weng Z, Whitfield TW,
582 Wang J, Collins PJ, Aldred SF, Trinklein ND, Partridge EC, Myers RM, Dekker J, Jain G,
583 Lajoie BR, Sanyal A, Balasundaram G, Bates DL, Byron R, Canfield TK, Diegel MJ, Dunn
584 D, Ebersol AK, Frum T, Garg K, Gist E, Hansen RS, Boatman L, Haugen E, Humbert R,
585 Jain G, Johnson AK, Johnson EM, Kutuyavin TV, Lajoie BR, Lee K, Lotakis D, Maurano
586 MT, Neph SJ, Neri FV, Nguyen ED, Qu H, Reynolds AP, Roach V, Rynes E, Sabo P,
587 Sanchez ME, Sandstrom RS, Sanyal A, Shafer AO, Stergachis AB, Thomas S, Thurman
588 RE, Vernot B, Vierstra J, Vong S, Wang H, Weaver MA, Yan Y, Zhang M, Akey JM,

589 Bender M, Dorschner MO, Groudine M, MacCoss MJ, Navas P, Stamatoyannopoulos G,
590 Kaul R, Dekker J, Stamatoyannopoulos JA, Dunham I, Beal K, Brazma A, Flicek P,
591 Herrero J, Johnson N, Keefe D, Lusk M, Luscombe NM, Sobral D, Vaquerizas JM, Wilder
592 SP, Batzoglou S, Sidow A, Hussami N, Kyriazopoulou-Panagiotopoulou S, Libbrecht
593 MW, Schaub MA, Kundaje A, Hardison RC, Miller W, Giardine B, Harris RS, Wu W,
594 Bickel PJ, Banfai B, Boley NP, Brown JB, Huang H, Li Q, Li JJ, Noble WS, Bilmes JA,
595 Buske OJ, Hoffman MM, Sahu AD, Kharchenko PV, Park PJ, Baker D, Taylor J, Weng Z,
596 Iyer S, Dong X, Greven M, Lin X, Wang J, Xi HS, Zhuang J, Gerstein M, Alexander RP,
597 Balasubramanian S, Cheng C, Harmanci A, Lochovsky L, Min R, Mu XJ, Rozowsky J,
598 Yan K-K, Yip KY, Birney E, Consortium EP 2012 An integrated encyclopedia of DNA
599 elements in the human genome. *Nature* 489:57-74.
600
601 **37.** Manole D, Schildknecht B, Gosnell B, Adams E, Derwahl M 2001 Estrogen promotes
602 growth of human thyroid tumor cells by different molecular mechanisms. *J Clin Endocrinol*
603 *Metab* 86:1072-1077.
604
605 **38.** Lin CY, Vega VB, Thomsen JS, Zhang T, Kong SL, Xie M, Chiu KP, Lipovich L, Barnett
606 DH, Stossi F, Yeo A, George J, Kuznetsov VA, Lee YK, Charn TH, Palanisamy N, Miller
607 LD, Cheung E, Katzenellenbogen BS, Ruan Y, Bourque G, Wei CL, Liu ET 2007 Whole-
608 genome cartography of estrogen receptor alpha binding sites. *PLOS Genet* 3:867-885.
609
610 **39.** Kobayashi K, Shaver JK, Liang W, Siperstein AE, Duh QY, Clark OH 1993 Increased
611 Phospholipase-C activity in neoplastic thyroid membrane. *Thyroid* 3:25-29.
612
613 **40.** Strassheim D, Shafer SH, Phelps SH, Williams CL 2000 Small cell lung carcinoma exhibits
614 greater phospholipase C-beta 1 expression and edelfosine resistance compared with non-
615 small cell lung carcinoma. *Cancer Res* 60:2730-2736.
616
617 **41.** Eszlinger M, Krohn K, Berger K, Lauter J, Kropf S, Beck M, Fuhrer D, Paschke R 2005
618 Gene expression analysis reveals evidence for increased expression of cell cycle-associated
619 genes and G(q)-protein-protein kinase C signaling in cold thyroid nodules. *J Clin*
620 *Endocrinol Metab* 90:1163-1170.
621
622 **42.** Kero J, Ahmed K, Wettschureck N, Tunaru S, Wintermantel T, Greiner E, Schuetz G,
623 Offermanns S 2007 Thyrocyte-specific G(q)/G(11) deficiency impairs thyroid function and
624 prevents goiter development. *J Clin Investig* 117:2399-2407.
625
626
627