

Transplantable human thyroid organoids generated from embryonic stem cells to rescue hypothyroidism



Open Access This file is licensed under a Creative Commons Attribution 4.0

International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to

the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. In the cases where the authors are anonymous, such as is the case for the reports of anonymous peer reviewers, author attribution should be to 'Anonymous Referee' followed by a clear attribution to the source work. The images or other third party material in this file are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Romitti et al present interesting progress on developing the thyroid organoid system from genetically modified human pluripotent stem cells. Transient overexpression of NKX2-1 and PAX8 expedite differentiation into thyroid hormone secreting follicular-like organoids. Given that thyroid organoid formation in humans is only successful from primary tissue, this PSC-based approach represents a major advance in the field. However, there are a substantial lack of experimental evidence to support the authors' conclusion mainly due to a lack of adequate sample size, appropriate control, and quantitative description of the results. My major concerns follow:

1. Comparator primary tissues

Throughout the manuscript, not primary tissue-derived samples were compared to thyroid organoids, prohibiting the fair assessment of maturity levels and its differentiation stage. It is increasingly recognized as important (almost imperative) to benchmark comparative-stage derived primary tissues. This comparison includes evaluation of gene expression level, spatial relationship, and functional output (secretion level). Additionally, it would be also informative to discuss their data compared to published literature of primary tissue-derived organoids.

2. Reproducibility and validation of scRNAseq

scRNAseq is certainly an informative tool to analyze the compositional profile of organoids however, it is only tested in one timepoint (day45) and one small number of samples (1874 cells), raising significant concern for reproducibility and robustness. How much extent cell compositions and their differentiation status are preserved across the different batches? It is highly encouraged to perform additional sample sets ideally with primary control.

Additionally, there is a substantial lack of scRNAseq validation to evaluate the significance of the findings. For example, the author's protocol derives a significant number of off-target cells by scRNAseq but is never evaluated in further figures. These need to be validated in later figures to determine their real existence, frequency, and spatial relationship to follicular tissues.

3. Quantitative assessment on the efficiency of hormone-secreting organoids in vitro and in vivo

One major drawback is the lack of a quantitative description of the authors' system. This is persistently problematic in all of the Figures. I wonder 1. the efficiency/yield of the hormone-secreting organoids, 2. what types of off-target cells are formed and again how much (as noted in #2), and 3. the level of hormone production capability in vivo (normalized by single-cell) and their discussion to normal adult follicular cells.

4. Lack of discussion with recent progress around thyroid organoids

Related to the #1 comment, the below literatures are recent examples of primary derived thyroid organoid system that was not mentioned nor discussed in the manuscript. It is highly encouraged to discuss pros/cons from multiple angles (differentiation levels, sample accessibility, transgene needs, immunogenicity, etc...).

Jelte van der Vaart, et al. Adult mouse and human organoids derived from thyroid follicular cells and modeling of Graves' hyperthyroidism. Proceedings of the National Academy of

Sciences 118 (51) e2117017118, (2021).

V. M. L. Ogundipe et al., Generation and differentiation of adult tissue-derived human thyroid organoids. *Stem Cell Reports* 16, 913–925 (2021).

Y. Saito et al., Development of a functional thyroid model based on an organoid culture system. *Biochem. Biophys. Res. Commun.* 497, 783–789 (2018).

A. Khoruzhenko et al., Functional model of rat thyroid follicles cultured in Matrigel. *Endocr. Connect.* 10, 570–578 (2021).

Reviewer #2 (Remarks to the Author):

Romitti et al perform T4-producing human thyrocytes from pluripotent stem cells. It seemed very difficult to achieve. The article lacks clarity (especially the first parts) and discussion. It's a shame because the work could be much more interesting. The part of single cell characterization of human thyroid organoid is very interesting.

There are several problems with the methodology (lack of controls), data and discussion, errors in the text. Whilst on a subject of potential interest, the paper is weak in several aspects.

Recommended points to address:

How do you explain difference of protocol between human and mice? Discuss please
Why it seems so difficult to obtain T4-producing human thyroid follicles?
Could you better explain and discuss each step of the protocol, and add controls?

Lines 38-39: Affirmation to be justified, publications? Too much affirmative, not real in clinical practice. It is not a strong argument. In comparison with immunosuppressive treatment needed for a transplant. Great interest for development biology, understanding... but for a replacement treatment for Levothyrox.

// Lines 24-25: This argument is not correct in abstract.

Line 79: Fig 1a: Add day of culture corresponding at each stick like Day 0 and Day 58

Line 80: Supp Fig1d: indicate stage of culture
Supp Fig 1e: line 804: replace l with e, add h

Line 90: Add co-immunostaining of NKX2-1 and PAX8
“large proportion”: could you give a percentage?

Lines 99-101: expression of endogenous NKX2-1: follows the PAX8 curve? And FOXE1, HHEX?

Could you conclude only with PAX8?

In Fig 1c: NKX2-1 seems to be very low in comparison with PAX8, why?

Line 101: direct addition of cAMP at D9 is not effective? Did you try without the proliferation

week?

Line 105: Add control: without cAMP

Could you show proliferation between D9 and D16? And difference with D20 after 2 weeks of cAMP?

Why control cells (stem cells) do not proliferate (supp fig 2a)?

Supp Fig 2b: 2 columns per stage?

Lines 109-110: why NKX2-1 increases and FOXE1 and PAX8 decrease between D16 and D23/30?

Fig 1c: no control on heatmap? And on all heatmaps

Controls are needed for each step, +/- dox, cAMP, TSH.....

Lines 115-117: PAX8 in the text and NKX2-1 in Supp 2d. Did you perform GFP co-immunostaining to check that TG-cells are the NKX2-1-GFP-cells?

Line 122: Could you explain more precisely the use of Dexamethasone?

Fig 1c: Could you give the quantification between D38 and D45? DUOXA2, TPO and TSHR decrease on heatmap. Supp 2e: strong increase between D45 +/-Dox but between D36 and D45, +/- Dox, +/- hTSH +/- Dexa, +/- SB

Controls are needed.

Line 130: supp Fig 2g: co-immunostaining with TG, TG-I, T4?

PAX8 in the text, NKX2-1 in figure?

Line 132: 25% of NKX2-1 cells, is it very efficiency? In comparison with other protocols?

Discuss please

What is the decrease in proliferation at D47 due to: differentiation, apoptosis, etc.

Line 139: it's not clear in the Fig 1c

Line 142: did you perform T4 immunostaining?

Line 147: why enrichment of 60% of NKX2-1-GFP+ cells? How did you choose this %?

Line 164: NKX2-1 GFP+ cells (60%) correspond to thyroid progenitors, immature thyrocytes and mature thyrocytes, airway cells

NKX2-1 neg cells (40%) correspond to fibroblasts, cardiovascular and endoderm epithelial cells? Are there other cells?

Line 169: in which figure?

Line 170: Did you check/validate crosstalk, that you observe with CellPhone-DB, between thyroid cells and the other cells? Did you perform culture without cardiovascular cells for example?

Line 203: quantifications of T4, TG and TPO are needed, comparison between D45 and D58?

Did you perform T4 assay in culture medium?

Did you analyze proliferation from D47 to D58? And apoptosis?

How long do T4-producing cells survive in culture in your model?

How many stem cells are needed to reproduce a functional human thyroid (11/12weeks)?
Same if it is a proof of concept. Did you compare T4-immunostaining? Quantification? (Fig 2d and compare with human thyroid tissue)

Line 235: Did you perform NIS immunostaining on graft? And analyze of airway, cardiovascular cells... by immunostaining on graft?

Line 236: human tissue at each stage?

Line 251: Did you perform TSH assay to validate the rescue?

Line 263: Could you compare your protocol and your results with those of Kurmanns? (human cells)

Line 528: Could you precise "blocking solution"

Line 549: Supp table 3? There is no Table 3.

Line 676: What is MTG? explain please

Mann and Whitney test is more appropriate than t-test

References to be checked

7,8: incorrect references

15: Ogundipe et al.

...

3

4 **Reviewer #1 (Remarks to the Author):**

5

6 Romitti et al present interesting progress on developing the thyroid organoid system
7 from genetically modified human pluripotent stem cells. Transient overexpression of
8 NKX2-1 and PAX8 expedite differentiation into thyroid hormone secreting follicular-
9 like organoids. Given that thyroid organoid formation in humans is only successful from
10 primary tissue, this PSC-based approach represents a major advance in the field.
11 However, there are a substantial lack of experimental evidence to support the authors'
12 conclusion mainly due to a lack of adequate sample size, appropriate control, and
13 quantitative description of the results. My major concerns follow:

14

15 1. Comparator primary tissues

16 Throughout the manuscript, not primary tissue-derived samples were compared to
17 thyroid organoids, prohibiting the fair assessment of maturity levels and its
18 differentiation stage. It is increasingly recognized as important (almost imperative) to
19 benchmark comparative-stage derived primary tissues. This comparison includes
20 evaluation of gene expression level, spatial relationship, and functional output
21 (secretion level). Additionally, it would be also informative to discuss their data
22 compared to published literature of primary tissue-derived organoids.

23 **R. Dear reviewer, thank you for your suggestions. As required, we compared our**
24 **results to primary thyroid tissue as described below:**

25 **qRT-PCR was used to compare our thyroid organoids (D58) and derived thyroid tissue**
26 **after transplantation (Graft) to human adult thyroid tissue. Fig. 4e. Lines 321-325.**

27 **Immunostaining was used to compare the resulted thyroid tissue after graft to the**
28 **human adult thyroid sample. The cell organization and protein expression show that**
29 **organoids-derived tissue resembles the architecture of adult human thyroid, however**
30 **the follicles size range is lower in graft compared to tissue. It might indicate that our**

31 organoids might represent an early stage of fetal thyroid development (Szinnai, 2007).
32 Fig. 4d and Supplementary Fig. 6e.

33 Despite the indications that at day 58 or after 5 weeks of transplantation, our organoids
34 resemble fetal thyroid, we aimed to compare our data to different stages of thyroid
35 development. Dom, et al, 2021, profiled the thyroids of human fetuses. We hoped
36 these data would be a good source of comparison. Unfortunately, the dramatic
37 technical variation between our RNAseq data and the Affymetrix arrays of Dom et al.
38 and the limited amount of data available for batch effect correction precluded any
39 meaning full comparison.

40 Furthermore, we have also contacted the authors of the article “Modeling Human
41 Thyroid Development by Fetal Tissue-Derived Organoid Culture, Adv. Sci, 2022”
42 aiming to compare their fetal thyroid organoids scRNAseq data to ours, but we didn’t
43 obtain any reply and since the data is not publicly available, we could not perform the
44 analysis.

45 Regarding comparison with adult thyroids, The Human Cell Landscape
46 (<https://db.cngb.org/HCL/>) includes two ‘normal’ single cell RNAseq thyroid profiles.
47 We processed both with our pipeline. One of them included a high percentage of
48 lymphocytes suggesting a thyroiditis. The other one had reasonable cell types
49 proportions. The Human Cell Landscape used the microwell-seq and lower per-cell
50 read counts, while we used 10X Chromium platform. After applying a batch effect
51 correction algorithm to resolve these technical differences, we found a significant
52 overlap between the cells of thyroid organoids and those of human thyroid tissue. The
53 analysis is presented at Fig. 3e and f.

54

55 2. Reproducibility and validation of scRNAseq
56 scRNAseq is certainly an informative tool to analyze the compositional profile of
57 organoids however, it is only tested in one timepoint (day45) and one small number of
58 samples (1874 cells), raising significant concern for reproducibility and robustness.
59 How much extent cell compositions and their differentiation status are preserved
60 across the different batches? It is highly encouraged to perform additional sample sets
61 ideally with primary control.

62 R: Dear reviewer, we agree that scRNAseq at another time point, and from human
63 thyroid tissue would bring new information to the article and reinforce the data from
64 day 45. For that we performed scRNAseq at day 58, when the system is functional,
65 and we also used the publicly available human cell landscape dataset from adult
66 thyroid tissue.

67 At Fig. 3a-c and Supplementary Fig. 3b, you can find the characterization of the
68 organoids at day 58 as well as the thyroid clusters identified. In addition, we performed
69 integration analysis initially with organoids from D45 and D58, showing strong overlap
70 between the time points (Fig. 3d); Secondly, we performed integration analysis among
71 the thyroid organoids datasets and the adult thyroid human cell landscape dataset (we
72 have excluded the sample in which most of cells were immune cells), which shows an
73 overlap between the hESC-derived thyrocytes and the adult thyroid cells (Fig. 3f).

74 As mentioned above the scRNAseq data from “Modeling Human Thyroid Development
75 by Fetal Tissue-Derived Organoid Culture, Adv. Sci, 2022” article is not publicly
76 available, so we could not perform the comparison analysis.

77

78 Additionally, there is a substantial lack of scRNAseq validation to evaluate the
79 significance of the findings. For example, the author's protocol derives a significant
80 number of off-target cells by scRNAseq but is never evaluated in further figures. These
81 need to be validated in later figures to determine their real existence, frequency, and
82 spatial relationship to follicular tissues.

83 R. We do share this concern as our study suggests that off-target cells are important
84 to effectively generate organoids that release thyroid hormones in the bloodstream.
85 As scRNAseq only provide a quantitatively biased picture of cell types composition
86 and does not inform spatial structures, we addressed this point with specific targeted
87 approaches. We first characterized each non-thyroid cluster by specific markers at
88 mRNA levels and performed IF to confirm the existence of such cell types and our
89 classification (Supplementary Fig. 4a-m). Furthermore, as predicted by connectome,
90 mesoderm might interact with thyroid cells through many distinct pathways. Hence,
91 we performed IF to determine the special location of such population and the proximity
92 with thyroid cells. Interestingly, we noticed that fibroblasts and cardiovascular cells

93 were in close contact with thyroid organoids, reinforcing the possible interaction
94 between them (Supplementary Fig. 5a-f). Even if at less extent, airway and epithelial
95 cells also seem to be close to thyroid cells as demonstrated at Supplementary Fig. 4k
96 and m.

97 Regarding the frequency of the non-thyroid cell types, due to technical issues we did
98 not have all markers (Ab) suitable for Flow cytometry quantitative analysis. In addition,
99 since we demonstrate by scRNAseq that compared to day 45, at day 58 the cluster of
100 fibroblasts has almost completely disappeared. The only conclusion we are able to
101 draw at this point is that the proportions of non-target cells can vary during the protocol.
102 Quantitative details will be addressed in future studies.

103 3. Quantitative assessment on the efficiency of hormone-secreting organoids *in vitro*
104 and *in vivo*
105 One major drawback is the lack of a quantitative description of the authors' system.
106 This is persistently problematic in all of the Figures. I wonder 1. the efficiency/yield of
107 the hormone-secreting organoids, 2. what types of off-target cells are formed and
108 again how much (as noted in #2), and 3. the level of hormone production capability *in vivo*
109 (normalized by single-cell) and their discussion to normal adult follicular cells.

110 R: Dear reviewer, as required we included the controls in the figures and included a
111 new Supplementary Fig. 2 where we show the controls for each change of the protocol
112 and how the gene expression changed according to the different treatments. We also
113 include the RNAseq heatmaps showing the expression profile using the new treatment
114 compared to cAMP control condition. For the time course bulk RNAseq (Fig. 1g), we
115 have not included -Dox condition as negative control, since the purpose is not to
116 compare it to an uninduced condition, but among the different stages of thyroid
117 development *in vitro*.

118 In order to measure TH yield *in vitro* we performed T4 and T3 dosage using the ELISA
119 kit and we could not detect clear values, which might be due to the limit of detection
120 of the techniques. In addition, following your suggestion, we tried to estimate the
121 percentage of functional follicles. Because we used paraffin sections for T4
122 immunostaining (5 μ m), counting follicles may not be appropriate because the same
123 follicle may be present on different sections.

124 The levels of THs *in vivo* were measured and compared to non-transplanted control
125 mice. Unfortunately, we can't compare the THs levels obtained from *in vivo*
126 transplantation to patients' levels, since the number of follicles, maturation degree and
127 TSH response are highly distinct. In addition, measurement of T4 production
128 (plasmatic level)/cell would require complete dissociation of all transplanted tissue and
129 counting of thyrocytes by flow cytometry. This could give us a T4 level/cell that would
130 be of little value anyway because of TSH regulation.

131

132 4. Lack of discussion with recent progress around thyroid organoids
133 Related to the #1 comment, the below literatures are recent examples of primary
134 derived thyroid organoid system that was not mentioned nor discussed in the
135 manuscript. It is highly encouraged to discuss pros/cons from multiple angles
136 (differentiation levels, sample accessibility, transgene needs, immunogenicity, etc...).

137

138 Jelte van der Vaart, et al. Adult mouse and human organoids derived from thyroid
139 follicular cells and modeling of Graves' hyperthyroidism. *Proceedings of the National*
140 *Academy of Sciences* 118 (51) e2117017118, (2021).

141

142 V. M. L. Ogundipe et al., Generation and differentiation of adult tissue-derived human
143 thyroid organoids. *Stem Cell Reports* 16, 913–925 (2021).

144

145 Y. Saito et al., Development of a functional thyroid model based on an organoid culture
146 system. *Biochem. Biophys. Res. Commun.* 497, 783–789 (2018).

147

148 A. Khoruzhenko et al., Functional model of rat thyroid follicles cultured in Matrigel.
149 *Endocr. Connect.* 10, 570–578 (2021).

150

151 R: Thank you for pointing it out. Indeed, those are very interestingly articles, and some
152 were not mentioned before since they were published after the first submission of the
153 paper. We have included a paragraph in the introduction (Page 3, L.54-67) and we
154 have discussed the suggested points with pros and cons of this model system
155 compared to ESCs-derived models. (Page 16, L.373-381; Page 17-18, L. 421-428).

156 **Reviewer #2 (Remarks to the Author):**

157 Romitti et al perform T4-producing human thyrocytes from pluripotent stem cells. It
158 seemed very difficult to achieve. The article lacks clarity (especially the first parts) and
159 discussion. It's a shame because the work could be much more interesting. The part
160 of single cell characterization of human thyroid organoid is very interesting.

161

162 There are several problems with the methodology (lack of controls), data and
163 discussion, errors in the text. Whilst on a subject of potential interest, the paper is weak
164 in several aspects.

165

166 Recommended points to address:

167

168 How do you explain difference of protocol between human and mice? Discuss please

169 R: Dear reviewer, the difference between both protocols might start by the fact that
170 mESC and hESC are different at basal state. As you can find in the discussion (Page
171 15, L.360-372) mESCs are considered as naïve while hESC are considered as primed
172 stem cells, it means that naïve mESCs can efficiently contribute to blastocyst
173 chimeras, whereas primed SCs have the potential to differentiate into primordial germ
174 cells *in vitro*. Moreover, not only pluripotency is not the same between the two stem
175 cells, but they also differ in timing of gene expression, levels and epigenetic
176 modifications. For example, studies have shown that neuron and sensory organ
177 develop differently from mESCs and hESCs (Gabdouline, 2015). Despite the
178 progresses made in the last years to induce a “more naïve” hESC culture method (as
179 used here by Stem Flex medium hESC culture), we cannot assume that both cells are
180 at the same stage of development and can differentiate similarly. The second aspect
181 is that for human the endoderm step is crucial, since spontaneously the hESC do not
182 generate substantial amount of endoderm cells. On the contrary of mouse model,
183 human thyrocytes cannot mature completely using only cAMP treatment, and the
184 addition of TSH is essential for the final maturation, which might indicate differences
185 in gene expression regulation between the two systems. Also, the timing of
186 development seems to follow the one observed *in vivo* probably the reason why the
187 human protocol takes around 60 days while mouse 21 days.

188 Why it seems so difficult to obtain T4-producing human thyroid follicles?

189 R: Based on our findings and from previous studies (Kurmann, 2015; Ma 2015, 2017,
190 2020; Serra, 2017) efficient generation of thyroid from pluripotent SCs is a challenge
191 and the protocols require several steps in order to generate thyroid cells. Most of the
192 studies show that this stage can be reached (thyrocytes derivation), but proper 3D
193 organization with thyroid hormone synthesis *in vitro* and *in vivo* was not achieved from
194 hPSCs. Many factors can lead to these difficulties such as, efficient endoderm-
195 induction and thyroid specification, 3D environment, cell culture media (most of the
196 studies don't use FBS, but defined supplementation, which can lead to the absence of
197 essential growth factors and other molecules) and the purified thyroid population, used
198 in most of the models can constitute a disadvantage since other non-thyroid cells can
199 play a role on thyroid development and maturation. An interesting example is shown
200 by Ogundipe, et al, in which adult tissue-derived thyroid organoids were transplanted
201 and needed 26 weeks to form functional tissue *in vivo*. Indicating that such pure
202 population might be inhibitory for thyroid formation and function, even when starting
203 from a mature and functional source of material.

204

205 Could you better explain and discuss each step of the protocol, and add controls?

206 R: Dear reviewer, we have included new nomenclature and experiments to provide
207 more information about the different stages of thyroid organoids development *in vitro*.

208 - At Fig. 1a, we renamed the stage from day 9-16 to thyroid fate commitment
209 instead of thyroid proliferation since the idea of this week without
210 supplementation was to allow the thyroid endogenous machinery activation.
211 And shown at new Fig. 1c, we performed flow cytometry for co-expressing
212 NKX2-1GFP and PAX8 cells and observed that around half of the cells that co-
213 expressed the two TF at day 9 keep the endogenous activation and day 16,
214 being the endogenous "progenitors". We also replaced the previous KI67
215 experiment by BrdU assay, and we observed that during this period of time,
216 cells uptake less BrdU, which reinforce the balance between differentiation and
217 proliferation (Fig. 1e). Of note, BrdU control, were cells treated with BrdU
218 without the Ab. Other controls were used for the verification of the analysis,

219 including -Dox condition (treated with BrdU and Ab) and cells not treated with
220 BrdU and incubated with the Ab.

221

222 - We also characterized by BrDU assay that from day 16 to 30 (cAMP treatment),
223 thyroid cells expansion occurs with a clear increase in the proportion of
224 proliferating cells expressing BrdU (Fig. 1e).

225

226 - We also included the Supplementary Fig. 2a-m, where we show by qRT-PCR
227 effect of the switch from cAMP supplementation to hrTSH +/-or Dexa +/-or
228 Sb431542 on thyroid gene expression (at day 38 and 45), showing all the
229 controls: -Dox, cAMP, cAMP+Dexa, hrTSH, hrTSH+Dexa, hrTSH+SB. In
230 addition, bulk RNAseq data was used to compare the inflammation, TGF-beta
231 and thyroid markers, between cAMP condition (control) and the new treatment.

232

233 - Immunostaining for TG was included starting from D30, showing the
234 progression of maturation after the medium supplementation switch (Fig. 1i).

235

236 - Day 45 was also now characterized by IF for the main thyroid protein markers
237 and to show the histological organization of the cells at this time point (Fig. 2a-
238 d).

239

240 - In addition, -Dox control was included in all qRT-PCR graphs as a column not
241 only as fold to -Dox as previously mentioned (Fig. 1c, d, h; Fig. 3g).

242

243 Lines 38-39: Affirmation to be justified, publications? Too much affirmative, not real in
244 clinical practice. It is not a strong argument. In comparison with immunosuppressive
245 treatment needed for a transplant. Great interest for development biology,
246 understanding... but for a replacement treatment for Levothyrox.

247 R: We modified this paragraph, and we added some references supporting this
248 discussion.

249

250 // Lines 24-25: This argument is not correct in abstract.

251 R: We deleted this sentence in this abstract because it seemed too ambitious. We
252 have replaced it with a few words that still leave room for an ambitious vision for the
253 medicine of the future.

254 One of the long-term therapeutic perspectives is, of course, to solve the problem of
255 immune tolerance. However, this problem, which is at the heart of regenerative
256 medicine, is already largely explored through approaches using patient-derived iPSC
257 or embryonic stem cells with universalization of the HLA system (*Riolobos L, Hirata*
258 *RK, Turtle CJ, Wang PR, Gornalusse GG, Zavajlevski M, Riddell SR, Russell DW.*
259 *HLA engineering of human pluripotent stem cells. Mol Ther. 2013 Jun;21(6):1232-41.*)
260 It therefore seems ambitious but scientifically relevant to investigate the same issues
261 in the context of thyroid medicine.

262

263

264 Line 79: Fig 1a: Add day of culture corresponding at each stick like Day 0 and Day 58

265 R: Thank you for your observation, days were included in the Fig.1a scheme of the
266 differentiation protocol.

267

268 Line 80: Supp Fig1d: indicate stage of culture

269 R: Dear reviewer, as suggested, we also included the stage of the culture to the figure
270 legend. It was already described in the methods “**Generation of tetracycline-induced**
271 **hESC line**” (Page 18, L. 460-462).

272

273 Supp Fig 1e: line 804: replace l with e, add h

274

275 R: Thank you for pointing it out. Figure, legend, and text were modified accordingly.

276

277 Line 90: Add co-immunostaining of NKX2-1 and PAX8
278 “large proportion”: could you give a percentage?

279 R: Dear reviewer, the proportion of NKX2-1 and PAX8 co-expressing cells at day 9
280 and 16 is shown in the new Fig. 1c.

281

282 Lines 99-101: expression of endogenous NKX2-1: follows the PAX8 curve? And
283 FOXE1, HHEX? Could you conclude only with PAX8?

284 R: Indeed, this is an interesting aspect and we explored how could we analyze NKX2-
285 1 expression by qRT-PCR, since in our Knock-in line GFP is replacing the exon 2 of
286 NKX2-1 gene (which has only 3 exons). We have tested several combinations of
287 primers, and due to the specificity (exon-exon junction design) we could not effectively
288 measure the mRNA. So, we decided to exclude the NKX2-1 gene expression from our
289 qRT-PCR analyses (Page 5, L104-106). On the other hand, we observed by IF that all
290 cells expressing PAX8 also expressed NKX2-1, which is not the case in the contrary
291 situation (Fig. 1b). So, we believe that the endogenous curve for PAX8 reflects the
292 thyroid fate commitment. In addition, as suggested, FOXE1 and HHEX expression
293 curve was included in the Supplementary Fig. 1g. Please note that HHEX expression
294 is not following the thyroid induction, since it is a classical endoderm marker, which in
295 our case is induced by AA in our protocol.

296

297 In Fig 1c: NKX2-1 seems to be very low in comparison with PAX8, why?

298 R: Thank you for the observation. There is indeed a technical effect pertaining to the
299 NKX2-1-GFP construct. As mentioned above in our NKX2-1-GFP line, one of the
300 alleles of NKX2-1 gene has GFP replacing exon 2, and since it interferes in mRNA
301 expression analysis it likely impacts the number of reads aligned on NKX2-1 in bulk
302 RNAseq which covers the entire transcript thanks to the random priming protocol.

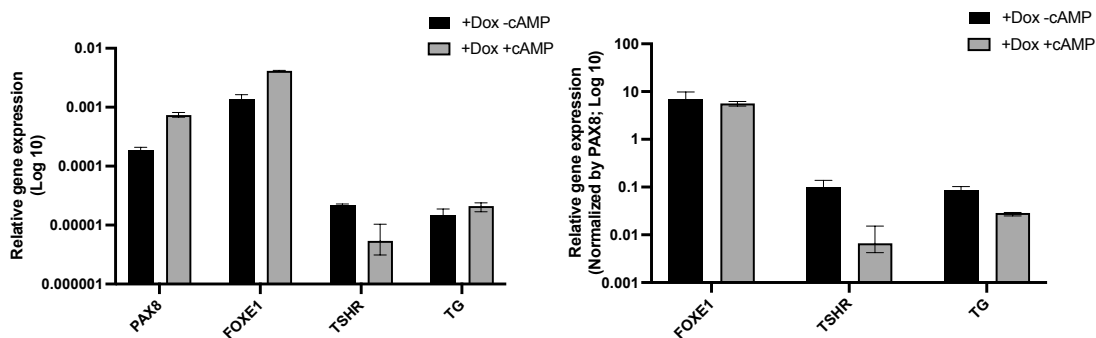
303 On the other hand, as demonstrated in the Fig. 2g, Fig. 3c and Supplementary Fig. 3
304 a-b, scRNAseq did not reveal any clear difference in the levels, very likely since in this
305 approach only the last 100-150bp from the 3' UTR are sequenced.

306 To the best of our knowledge, no study provides a strong reference to compare PAX8
307 and NKX2-1 levels during thyroid development *in vitro* or *in vivo*. We opted to show
308 the data as neutrally as possible, and refrain to emphasize any dataset.

309 Line 101: direct addition of cAMP at D9 is not effective? Did you try without the
310 proliferation week?

311 R: Dear reviewer, indeed that is an interesting point and we have renamed this step
312 (from day 9-16) to thyroid fate commitment, since it is when the Dox-induced cells
313 should start to express the endogenous thyroid machinery and activate the thyroid
314 differentiation program, which does not reflect the proliferation stage of the protocol.
315 So, when we treated the cells with cAMP at day 9 we observed increase in thyroid
316 population proliferation measured by BrdU uptake at day 16 compared to the untreated
317 situation (+cAMP 8.9% and -cAMP 5,9%). This reflected the % of NKX2-1^{GFP} with
318 around 50% more cells in the cAMP treated condition compared to the control (+cAMP
319 4.6% and -cAMP 3.0%). Also, we saw by qRT-PCR that PAX8 and FOXE1 levels are
320 slightly higher while TG is similar and TSHR is lower. Due to the higher number of
321 cells in the +cAMP condition we normalized once more the data using PAX8 levels
322 and we saw that the levels of the differentiation markers TSHR and TG are lower
323 compared to the untreated situation, indicating that the balance between of
324 proliferation and differentiation is important promoting thyroid commitment fate. Data
325 is show in the graphs bellow.

326



327

328

329 Line 105: Add control: without cAMP

330 R: Dear reviewer, the new version of the paper includes a BrdU experiment instead of
331 KI67 assessment. For that we used several different controls, such as, -Dox (to set
332 the gates), treated (all treatments) cells incubated of not with BrdU and/or with and
333 without anti-BrdU antibody. As the goal here was to determine the curve of proliferative

334 cells over time based on NKX2-1/GFP co-staining with BrdU the control is the treated
335 (BrdU) cells without anti-BrdU.

336

337 Could you show proliferation between D9 and D16? And difference with D20 after 2
338 weeks of cAMP?

339 R: As required, in the Fig. 1e we demonstrate the profile of BrdU+ among NKX2-1
340 expressing cells during this period and compared to the other time points/

341

342 Why control cells (stem cells) do not proliferate (supp fig 2a)?

343 R: Since for proliferation assay, we analyze BrdU positivity among NKX2-1 expressing
344 uninduced cells cannot be used as control (which in this case would be -Dox condition,
345 which comprises many differentiated cell types, not being at a stem cell state
346 anymore), so as control cells treated with BrdU solution and not stained were used as
347 control for the experiment. Please see Fig. 1f where we demonstrate that -Dox
348 condition has extremely low % of spontaneously differentiated NKX2-1 cells.

349 Supp Fig 2b: 2 columns per stage?

350 R: Dear reviewer, we have performed the experiment in duplicates, for which one we
351 have used at least 6 wells pooled together in order to overcome the possible variability
352 among wells (information is included in the “RNA-seq and analysis of bulk samples”
353 session described in Methods). The goal here was to see the changes in thyroid genes
354 expression overtime during the differentiation protocol, using the previous time point
355 as control, since we showed by qRT-PCR at several time points the thyroid gene
356 expression compared controls (Fig. 1. d; h; Fig. 3g; Supplementary Fig. 2 c-l). Also,
357 for the RNAseq data we included cAMP as control for TGF-beta, inflammation and
358 thyroid markers comparing to the medium changes from day 30 onwards
359 (Supplementary Fig. 2a, b, m).

360

361 Lines 109-110: why NKX2-1 increases and FOXE1 and PAX8 decrease between D16
362 and D23/30?

363 R: At day 16 we could expect still to have an effect of the exogenous induction of
364 NKX2-1 and PAX8, which could explain higher levels of PAX8 and FOXE1. NKX2-1
365 levels might be affected since it is a knock-in knock-out in our cell line.

366

367 Fig 1c: no control on heatmap? And on all heatmaps. Controls are needed for each
368 step, +/- dox, cAMP, TSH.....

369 R: As explained above, the meaning of the bulk RNAseq here is not to show induction
370 of differentiation but see the changes in gene expression overtime during the
371 differentiation protocol and the comparison it is made among time points. Besides that,
372 we included RNAseq data using cAMP as control to see changes in TGF-beta,
373 inflammation and thyroid markers comparing to the medium changes from day 30
374 onwards (Supplementary Fig. 2a, b, m).

375

376 Lines 115-117: PAX8 in the text and NKX2-1 in Supp 2d. Did you perform GFP co-
377 immunostaining to check that TG-cells are the NKX2-1-GFP-cells?

378 R: Dear reviewer, we apologize for the mistake, and it has been corrected, thank you
379 for pointing it out. We have observed by IF that all the TG-expressing cells are NKX2-
380 1/GFP positive, however not all NKX2-1/GFP cells are TG positive since we can have
381 airway and goblet cells.

382

383 Line 122: Could you explain more precisely the use of Dexamethasone?

384 R: Since studies have shown that high inflammation impact thyroid genes expression
385 (see list of references bellow) we evaluated the levels of inflammation by bulk RNAseq
386 at days 38 and 45 when cells were treated only with cAMP (Supplementary Fig. 2a).
387 Also, qRT-PCR data showed the dexamethasone effect on NIS (D38) and the addition
388 to SB increasing significantly the levels of most of the genes including NIS, TG and
389 TPO at day 45 compared to distinct control conditions (Supplementary Fig. 2c-l).

390 References:

391 Mori, K., Mori, M., Stone, S., Braverman, L. & DeVito, W. Increased expression of
392 tumor necrosis factor-alpha and decreased expression of thyroglobulin and thyroid
393 peroxidase mRNA levels in the thyroids of iodide-treated BB/Wor rats. *European*
394 *Journal of Endocrinology* 139, (1998).

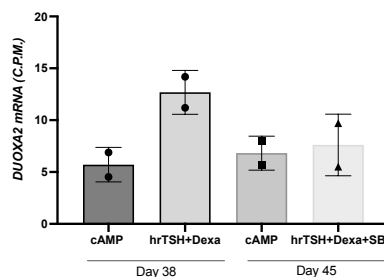
395 Dohán, O. et al. The sodium/iodide symporter (NIS): Characterization, regulation, and
396 medical significance. *Endocrine Reviews* vol. 24 48–77 (2003).

397 Faria, M. et al. TNF α -mediated activation of NF- κ B downregulates sodium-iodide
398 symporter expression in thyroid cells. *PloS one* 15, e0228794–e0228794 (2020).

399

400 Fig 1c: Could you give the quantification between D38 and D45? DUOXA2, TPO and
401 TSHR decrease on heatmap. Supp 2e: strong increase between D45 +/-Dox but
402 between D36 and D45, +/- Dox, +/- hTSH +/- Dexa, +/- SB
403 Controls are needed.

404 R: Dear reviewer, as required we have included the qRT-PCR data from days 38 and
405 45 for most of the thyroid genes including several controls, see Supplementary Fig. 2.
406 c-l. In addition, since our designed primers for quantification of DUOXA2 didn't provide
407 acceptable curve and trustful data, bellow you can find the data concerning DUOXA2
408 gene expression (C.P.M.) from bulk RNAseq at both time points, compared to the
409 cAMP control.



410

411

412 Line 130: supp Fig 2g: co-immunostaining with TG, TG-I, T4?

413 R: Dear reviewer, new confocal images from day 45 are included in the new Fig. 2 a-
414 d.

415

416 PAX8 in the text, NKX2-1 in figure?

417 R: Thank you for spotting the mistake, the figure was corrected accordingly.

418

419 Line 132: 25% of NKX2-1 cells, is it very efficiency? In comparison with other
420 protocols? Discuss please

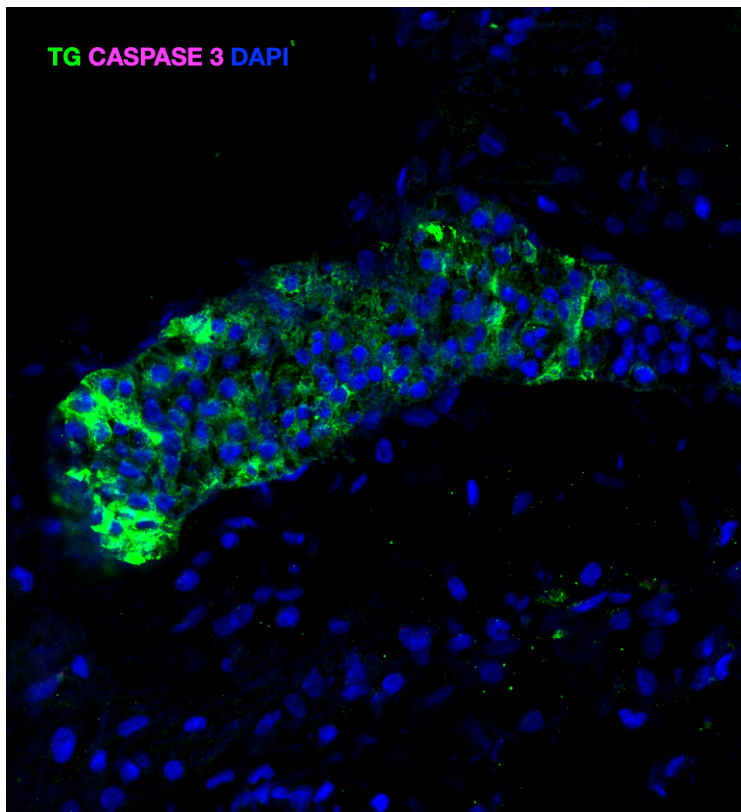
421 R: Dear reviewer, since obtained around 25% of the cells at the end of the protocol
422 which are NKX2-1+ and considering that our model is multicellular and that the other
423 cells are expected to be more proliferative than thyroid, the efficiency is significant
424 considering human models. Our thyroid mouse model and from other groups show
425 efficiency ranging from 24-60% of the cells (Posabella, et al, Frontiers in Endoc, 2021),
426 showing that our human model is close to this range of efficiency. However, proper
427 comparison to all protocols of human thyroid-derived (PSCs models) organoids is not
428 possible since the studies don't describe clearly the efficiency of thyroid generation.
429 Arauchi, et al, (Frontiers in Endoc, 2017) describes an efficiency of 16% of the thyroid
430 cells generation at day 20 of their protocol. Ma, et al, (Frontiers in Endoc, 2020) shows
431 that after Endoderm and specification induction 34% of the cells co-expressed NKX2-
432 1 and PAX8, compared to our protocol in which we observed 80% of the cells co-
433 expressing both TFs at day 9. In addition, they show that at day 21 of the protocol,
434 TSHR/NIS co-expressing cells were around 58% and after sorting could reach 97% of
435 purification.

436

437 What is the decrease in proliferation at D47 due to: differentiation, apoptosis, etc.

438 R: Dear reviewer, since we decided to change the assay for measuring precisely the
439 cells proliferating, at the new Fig. 1e, we show that proliferation is clearly reduced from
440 day 37 compared to day 30, which follows the change of media supplementation from
441 cAMP to hrTSH+Dexa. It is clear that cAMP is promoting proliferation in our system
442 as you can observe by the improve in % of proliferating cells from day 16 onwards. In
443 addition, we can observe that the media change at day 30 resulted in increase in
444 thyroid differentiation and maturation, shown at Fig. 1g, suggesting the balance
445 between differentiation and proliferation in our system. Also, it is important to notice,
446 that despite the decrease in the proliferation rate from D37, the % of NKX2-1 cells is
447 stable (Fig. 1f). Furthermore, to address the possible process of apoptosis, we

448 performed Caspase3 staining in our organoids (D58) and we did not observe any
449 expression among thyroid cells (see example below).



450

451

452 Line 139: it's not clear in the Fig 1c

453 R: The sentence was rewritten adequately (Page. 8, L.178). At Fig. 1d we
454 demonstrated that at day 9, *FOXE1*, *TSHR* and *TG* can be detected by qRT-PCR,
455 even if at lower levels compared to *NKX2-1* and *PAX8*. While by bulk RNAseq and
456 qRT-PCR (Fig. 1g and Supplementary Fig 2. c-l) we demonstrate that significant
457 expression of those genes and other maturation markers is observed from day 37
458 onwards.

459

460 Line 142: did you perform T4 immunostaining?

461 R: To better characterize the cells at day 45, you can find at the new Fig. 2 d the co-
462 staining for TG-I and T4 among NKX2-1 cells.

463

464 Line 147: why enrichment of 60% of NKX2-1-GFP+ cells? How did you choose this
465 %?

466 R: Dear reviewer, we enriched the thyroid population aiming to obtain enough thyroid
467 cells for characterization, since 25% GFP could result in very diluted distinct thyroid
468 population. Also, the aim was only identifying the other cells, not to estimate
469 proportion.

470

471 Line 164: NKX2-1 GFP+ cells (60%) correspond to thyroid progenitors, immature
472 thyrocytes and mature thyrocytes, airway cells NKX2-1 neg cells (40%) correspond to
473 fibroblasts, cardiovascular and endoderm epithelial cells? Are there other cells?

474 R: Dear reviewer, please see Fig. 2e-f and Fig. 3a-b. Since we performed a new
475 scRNAseq at day 58 we reanalyze our data set and reclassified the clusters, according
476 to the expression profile. However, we cannot exclude the presence of other cells in
477 very low proportion, such as fibroblasts at day 58. In addition, you can find the
478 characterization by immunostaining of those non-thyroid cells at Supplementary Fig.
479 4j-m and Supplementary Fig. 5c and f.

480

481 Line 169: in which figure?

482 R: Dear reviewer, this data is based in the cell counting from scRNAseq data. From
483 the total of NKX2-1 cells, more than 75% are also PAX8, this quantification is not
484 shown as graph, since the UMAPs from Fig. 2g and Fig. 3c show the clusters
485 expressing both TFs.

486

487 Line 170: Did you check/validate crosstalk, that you observe with CellPhone-DB,
488 between thyroid cells and the other cells? Did you perform culture without
489 cardiovascular cells for example?

490 R: Dear reviewer, in our culture condition, sorted GFP+ cells (day 23) which were
491 cultured alone did not survive. To show special distribution of thyroid cells and
492 mesodermal/cardiovascular cells, we used immunostaining which shows that both
493 populations are in close contact within our organoids (Supplementary Fig. 5c and f).

494 In addition, TGF-beta inhibition resulted in improvement of maturation as
495 demonstrated by several analysis within the paper. Concerning the role of
496 cardiovascular cells during thyroid development, we have a project ongoing aiming to
497 understand such relationship. For this reviewing, due to the complexity of the
498 experiment we could not perform it on time. For example, we observed that
499 dissociation of thyroid cells clusters, even at early stages, interfere in the thyroid 3D
500 generation, so removal of cardiovascular cells gets tricky.

501

502 Line 203: quantifications of T4, TG and TPO are needed, comparison between D45
503 and D58?

504 R: Dear reviewer, following your request we also performed immunostaining, qRT-
505 PCR and scRNAseq at day 58, in order to determine the differences in gene
506 expression, structures organization and determine the thyroid populations (see Fig. 2
507 and Fig. 3). T4 was assessed by IF and confocal images are included in both figures.
508 However, quantification of T4 levels in the medium was performed but not detected,
509 which might be due to the limit of detection of the T4 Elisa kit used in the study.

510

511 Did you perform T4 assay in culture medium?

512 R: As described above, despite the observation of T4 accumulation in the follicles, we
513 could not detect clear levels in the culture media, which can be due to the limit of
514 detection of the T4 Elisa kit used in the study.

515

516 Did you analyze proliferation from D47 to D58? And apoptosis?

517 R: Dear reviewer, please see Fig. 1e whereby BrdU uptake assay we describe the
518 proliferation rate from day 9 to day 58. Apoptosis was evaluated by IF using Caspase
519 3 Ab, and we did not observe apoptosis among thyroid cells (see above).

520

521 How long do T4-producing cells survive in culture in your model?

522 R: We have kept the cells in culture for 70 days and observe that the organoids keep
523 the expression of TG, TPO and accumulate TG-I in the lumen (see Supplementary
524 Fig. 3c).

525

526 How many stem cells are needed to reproduce a functional human thyroid
527 (11/12weeks)? Same if it is a proof of concept. Did you compare T4-immunostaining?
528 Quantification? (Fig 2d and compare with human thyroid tissue)
529

530 R: Dear reviewer, since it is not clear how to estimate the number follicles in the human
531 thyroid, we could use the transplantation experiment as example. Since for each
532 transplanted mice we used around 1,500 follicles and the generated tissue was
533 functional and almost recovered the levels in hypothyroid mice. We can estimate that
534 by using 6 wells of thyroid organoids we need initially 240,000 stem cells for EBs
535 generation. T4 immunostaining is compared at Fig. 4d, where grafted tissue is
536 compared to thyroid tissue.

537

538 Line 235: Did you perform NIS immunostaining on graft? And analyze of airway,
539 cardiovascular cells... by immunostaining on graft?

540 R: Dear reviewer, NIS staining was performed in organoids and grafted samples,
541 however the staining was not convincing since the localization was not corrected,
542 same staining was observed using human tissue, so we concluded that none of our
543 antibodies was efficiently staining NIS. Moreover, as shown at Fig. 4 f, ¹²³I uptake
544 images shows NIS presence since the cells could uptake iodine. The graft staining
545 with SMA demonstrates that all the non-thyroid cells are SMA+. We also did not
546 observe any E-Cadherin+ area without NKX2-1, TG staining, which can exclude the
547 presence of airway cells. In addition, Troponin T staining did not show any
548 cardiovascular cells.

549

550 Line 236: human tissue at each stage?

551 R: Dear reviewer, for the comparison we used adult thyroid tissue (see Methods: “RNA
552 extraction and quantitative real-time PCR” session).

553

554 Line 251: Did you perform TSH assay to validate the rescue?

555 R: Dear reviewer, as suggested, TSH was measured and the correlation with T4 levels
556 is demonstrated at Fig. 4i. In addition, we measured T3 levels and *Dio1* mRNA in the
557 liver, since *Dio1* is shown to be very sensitive to changes in TH levels (Fig. 4h and 4j;
558 Supplementary Fig. 6g).

559

560 Line 263: Could you compare your protocol and your results with those of Kurmanns?
561 (human cells)

562 R: Kurmann’s study shows a direct differentiation protocol to generate NKX2-
563 1+/PAX8+ thyroid follicular epithelial progenitors. This is the first aspect that differs
564 from our model, where we transiently induce NKX2-1 and PAX8 expression by
565 doxycycline incubation. In both cases, endoderm induction is promoted by Activin A
566 incubation, but Kurmann’s protocol includes an anteriorization step and 10 days of
567 incubation with FGF2 and BMP4 to induce thyroid specification. Also, the
568 differentiation/maturation period used is taken 22 days in their protocol, while in ours
569 it takes 49 days. Another important aspect is that our model uses FBS in the media
570 composition while the media used in Kurmann’s protocol is serum-free supplemented
571 with distinct compounds depending on the stage of the protocol. Since efficiency was
572 not evaluated in the study, we can compare to our organoid’s efficiency generation.
573 Finally, regarding the organization and functionality of the thyroid follicles we show the
574 3D follicular organization and accumulation of T4 inside the lumen, what is not
575 reported for Kurmann’s organoids.

576

577 Line 528: Could you precise “blocking solution”

578 R: Dear reviewer, since the assay used for proliferation assessment was changed in
579 the new version of paper, the sentence was removed, and the method rewritten
580 accordingly.

581 Line 549: Supp table 3? There is no Table 3.

582 R: Dear reviewer, sorry for the mistake, it referred to Supplementary table 3, which is
583 now corrected.

584

585 Line 676: What is MTG? explain please

586 R: Dear reviewer, MTG refers to Matrigel. It is now written when it appears for the first
587 time on the text (Page 4. L. 94).

588

589 Mann and Whitney test is more appropriate than t-test

590 R: Dear reviewer, all the analysis were performed again using non-parametric tests as
591 required.

592

593 References to be checked 7,8: incorrect references

594 R: Dear reviewer, the references are corrected accordingly.

595

596 15: Ogundipe et al.

597 R: Dear reviewer, thanks for noticing the mistake, the reference was corrected.

598

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Romitti et al has provided several new data to address the original concern that I've raised and partially addressed the issues.

#1. Thanks for providing the immunostaining data to compare adult thyroid tissues. However, details are lacking to assess the data (for example what day?) It seems awkward that panel C claimed "follicular structure" is not entirely consistent with panel d T4 positive area which raises a concern that the follicular area is not entirely a thyroid tissue.

SI Fig6 H&E (and scRNAseq datasets) nuances this concern and it could be just the other epithelial cystic structure that often comes out from PSC-derivatives. It would be informative to show what's off-target tissue came out of these grafts and discuss differences/similarities from primary thyroid.

#2. Reproducibility concern is not addressed as the authors just added one different timepoint. But probably fine without these data, given a higher workload to get new sample and BMI done.

#3. Quantifying the human thyroid function after transplantation is not formally addressed. (I believe) ELISA-based assay didn't distinguish human versus mouse T4 . So in vivo quantification of T4 could be just a variation of RAI treatment or other indirect protection of the native thyroid.

To quantify organoid graft-derived tissues, it would be highly encouraged to look at graft-derived T4 production capability and discuss their function relative to original tissues.

Reviewer #2 (Remarks to the Author):

Romitti et al have made many corrections in this article with many experiments to improve its quality. The proofreading is complicated in the text and the figures as the article has been so much modified, but to be better. Mainly and among others, the essential dosage of T4 in the grafted animals.

The authors responded well to all questions and comments. They added more experiences than requested.

Some comments:

Fig 2D: the T4 immunostaining is unclear: could you improve this figure for the final picture, T4 is very weak? Or is there a problem for this picture?

T4 assay in medium: Have you tried concentrating the medium to assess T4?

Organoids at day 70 day: can they be grafted? Do they have larger follicles?

Nis antibody: have you tried Nancy Carrasco's NIS antibody?

Line 252: could you explain further "epithelial cells 1" vs 2: with KRT14 and KRT7 and AQP5

but why?

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author):

Romitti et al has provided several new data to address the original concern that I've raised and partially addressed the issues.

#1. Thanks for providing the immunostaining data to compare adult thyroid tissues. However, details are lacking to assess the data (for example what day?) It seems awkward that panel C claimed "follicular structure" is not entirely consistent with panel d T4 positive area which raises a concern that the follicular area is not entirely a thyroid tissue.

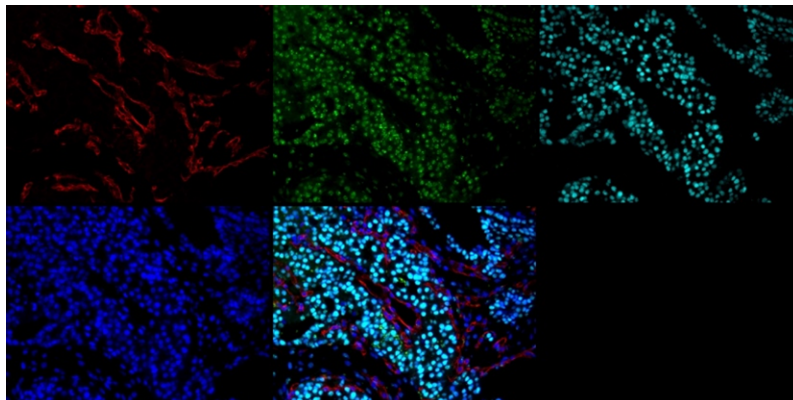
R: Figure 4 refers to D58 of organoids culture. Indeed, not all the thyroid follicles present in the grafted tissue produce detectable amount of T4, however, as showed at Fig 4C and D, epithelial cells presented in the grafted tissue express thyroid markers, such as NKX2-1, TG and TPO. Despite that a proportion of the follicles are not functional (T4+) after 5 weeks, it is very likely that the maturation and function is progressive and at this stage it might not have reached the full maturation.

SI Fig6 H&E (and scRNAseq datasets) nuances this concern, and it could be just the other epithelial cystic structure that often comes out from PSC-derivatives. It would be informative to show what's off-target tissue came out of these grafts and discuss differences/similarities from primary thyroid.

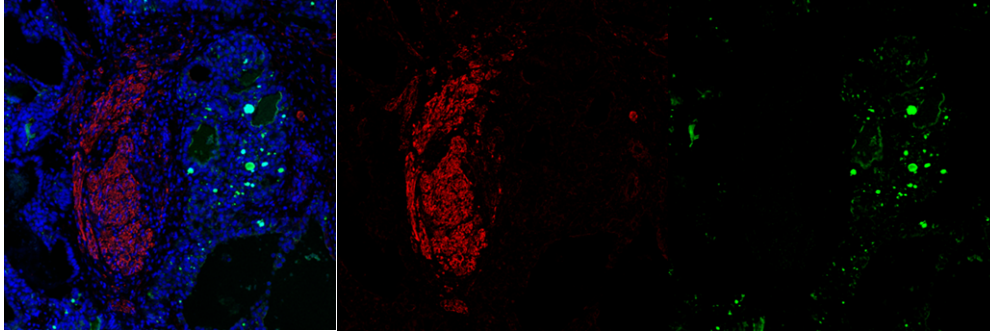
R: As described above, the epithelial structures present in the grafted tissue express markers such as NKX2-1, TG and TPO strongly indicating that those structures are thyroid tissue. Moreover, as described in the methods section, thyroid epithelial structures were enriched by filtration after dissociation, meaning that structures bigger than 100um were not used for the graft experiment. Also, the non-thyroid epithelial structures present within the organoids (lung airway mostly) are characteristically bigger than 100um, as you can appreciate at SI Fig 4 j and m. Moreover, as evidenced in our both single cell RNAseq, stem cells are not detected within the organoids at day 45 and 58, which reduces the chances of spontaneous generation of PSC-derived cystic structures after graft.

Concerning the off-target cells present within the grafted tissue, we detected by IF the presence of fibroblasts (SMA+) from which the origin is the human organoids (HNA+) and endothelial/vascular cells (CD31+) which form the blood vessels originated from the host (mouse cells). In addition, we have performed staining for lung markers (detected by scRNAseq in the organoids) and no cells were detected.

CD31 HNA NKX2-1 DAPI



SMA TG DAPI



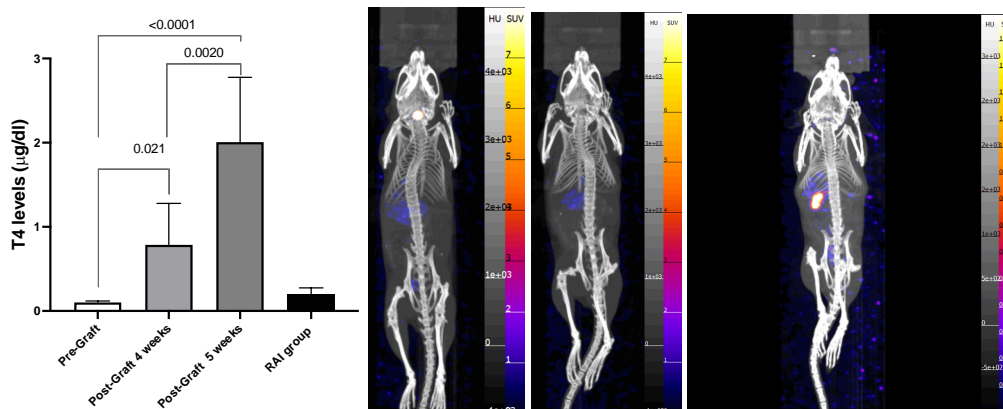
#2. Reproducibility concern is not addressed as the authors just added one different timepoint. But probably fine without these data, given a higher workload to get new sample and BMI done.

R: Dear reviewer, as mentioned previously, due to the time and technical challenges to get additional “normal” thyroid tissue and more samples from our organoids, additional scRNAseq were not performed, except for day 58. However, we believe that the data present here, strongly supports the generation of human functional thyroid in vitro and in vivo. We also believe that understanding the mechanisms in each step of thyroid formation in vitro and the comparison to human thyroid tissue is valuable and can bring insights about mechanisms involved in thyroid development which are still unknown, and this is being addressed in ongoing projects in the lab.

#3. Quantifying the human thyroid function after transplantation is not formally addressed. (I believe) ELISA-based assay didn't distinguish human versus mouse T4 . So in vivo quantification of T4 could be just a variation of RAI treatment or other indirect protection of the native thyroid.

R. Dear reviewer, Indeed ELISA does not distinguish T4 from mouse or human, however, we can observe that in the RAI group, the animals do not show recovery or increase in T4 levels, what is clearly observed among the grafted animals (Fig 4G), similar results are obtained when T3 levels were analyzed. Moreover, as previously demonstrated at SI Fig 6F we can see progressive increase in T4 levels from week 4 to 5 after transplantation. Also, here we included additional data showing the T4 levels among the same animals, before transplantation, and after 4 and 5 weeks. This clearly shows that the transplant is the cause of increase in T4 levels, especially if we compare to RAI group at week 5 where the levels are similar to pre-transplantation (time point) levels of grafted group.

We also include here, the 123I uptake images from controls, RAI and grafted animals (respectively) at the end of the protocol, showing that no sign of thyroid tissue (123I uptake) is observed in the neck region of the RAI and grafted groups. Taken together these set of evidence strongly indicate that the thyroid native tissue is not recovered after RAI treatment and that the increase in THs is due to the organoids transplantation.



To quantify organoid graft-derived tissues, it would be highly encouraged to look at graft-derived T4 production capability and discuss their function relative to original tissues.

R. Dear reviewer, as suggested we have counted the T4-producing follicles (expressing at least one thyroid marker: NKX2-1, TG or TPO) inside the grafted tissue and we observed that around 58% of them have detectable T4. This information is now included in the text L.316-319. We also discuss that it suggests that there is an ongoing maturation/functional progress and that possibly soon it could reach full maturation. Also, we believe that comparing our transplanted organoids to adult thyroid tissue is not the most appropriated since we could consider our follicles as fetal thyroid, in which the accumulation and synthesis of THs is not comparable to adult tissue.

Reviewer #2 (Remarks to the Author):

Romitti et al have made many corrections in this article with many experiments to improve its quality. The proofreading is complicated in the text and the figures as the article has been so much modified, but to be better. Mainly and among others, the essential dosage of T4 in the grafted animals. The authors responded well to all questions and comments. They added more experiences than requested.

Some comments:

Fig 2D: the T4 immunostaining is unclear: could you improve this figure for the final picture, T4 is very weak? Or is there a problem for this picture?

R. Dear reviewer, due to the overlap with TG-I indeed the signal for T4 was not very visual, so we improved the quality of the image, now included in the Fig 3.

T4 assay in medium: Have you tried concentrating the medium to assess T4?

R. Dear reviewer, we have been working on T4 measurements by MS since ELISA kits are not sensitive enough for low amounts as we expect in our culture media. However, even trying to concentrate T4 we could not detect clear amount of T4 in the media. For the future the idea is to try to isolate T4 accumulate in the lumen of the follicles and also set up a protocol in order to collect without using the classical differentiation media (which might contain T4 in FBS) and/or PBS, since it is not compatible to the MS. Since it is quite challenging and time/culture consuming, we could not include clear measurements in this paper.

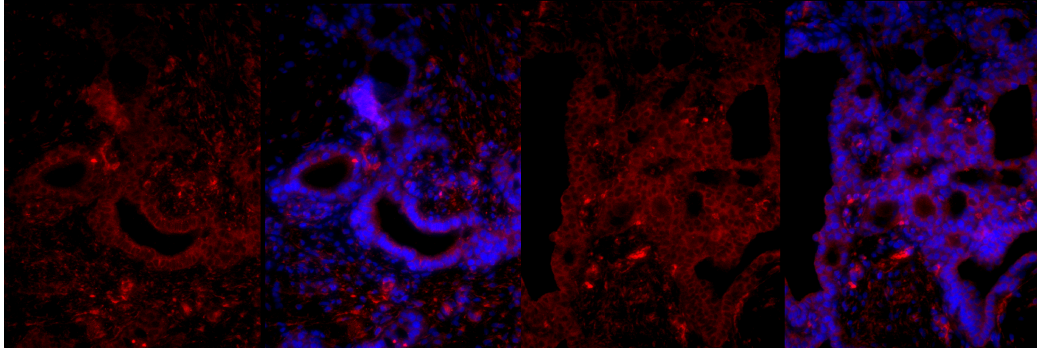
Organoids at day 70 day: can they be grafted? Do they have larger follicles?

R: The thyroid follicles kept in culture till day 70 show similar size and protein expression compared to day 58, which suggest that there is not major maturation/folliculogenesis occurring after day 58 in vitro. However, we have not grafted follicles from D70 into mice, so we cannot assert that it works as observed from D45. Also, using organoids from D45 is technically more advantageous since the culture period is considerably shorter and as we demonstrated the cells continue to mature in vivo.

Nis antibody: have you tried Nancy Carrasco's NIS antibody?

R: Dear reviewer, indeed we have tried some NIS antibodies provided by Nancy Carrasco but oppositely to mouse, in human organoids we don't see clear basal localization of the staining which also gave a lot of background making us to choose to do not show it (see pictures below). It might be that adjustments in the protocol are needed and we are working on that.

NIS (Carrasco) DAPI



Line 252: could you explain further “epithelial cells 1” vs 2: with KRT14 and KRT7 and AQP5 but why?

R: The clusters labeled "Epithelia cells" (1 and 2) express some lung cell type markers (such as KRT14, KRT7 and AQP5) but not a complete list of cell type markers otherwise expressed in the dataset. Thus, their differentially expressed genes could match many known cell types and they cannot be identified as belonging to any of them. The only label we can assign them with certainty is "epithelial" because they overexpress epithelial markers such as keratins (see the 50 first differentially expressed genes for clusters Epithelial cells 1 and Epithelial cells 2 below). The reasons to identify them as separate clusters (Epithelial cells 1 and Epithelial cells 2) and not giving them the same name (e.g. "Epithelial") are:

- 1) They are widely separated in the UMAP space
- 2) kNN clustering, even at low resolutions (resolution 0,1 in Seurat's FindClusters function) assigns them to 2 separate clusters
- 3) Their marker genes differ (see 50 first markers genes below where only 3 genes overlap)

Those 3 elements imply that they are two distinct cell populations that must be distinguished for downstream analyses such as ligand-receptor interactions.

Differentially expressed genes of cluster Epithelial cells 1 : MALAT1 S100A9 SPRR1B FABP5 MT-ND3 SPRR3 KRT5 MT-CO2 MT2A MT-ND4 MT-CO1 MT-CO3 MT-ATP6 NEAT1 MT-ND1 KRT13 MT-ND2 MT-CYB KRT15 KRT6A AGR2 COL3A1 S100P TMEM41B ORMDL1 TSPAN6 MTF2 TTC17 KRT17 MARCH6 SREK1 S100A6 ATP6V0D1 MIER1 BDH2 PLEKHJ1 MED10 PSMD10 POR RALY ATXN2 IARS2 USP48 VAMP2 WTAP PHF14 GSTK1 MKRN1 EMC4 GLOD4

Differentially expressed genes of cluster Epithelial cells 2 : DCD MUCL1 SCGB2A2 AZGP1 PIP SCGB1D2 SCGB1B2P AQP5 RPS4Y1 SFRP1 SAA1 LRRC26 ROPN1B STAC2 SAA2 TMEM213 CALML5 TFCP2L1 DNER KRT15 KRT7 SLC12A2 NCALD ADIRF RPL10 ZG16B CLDN10 RPL19 DBI PLA2R1 LGALS3 TSC22D1 COX4I1 RPS27A RBM3 SNORC FRZB PEBP1 RPL8 PPDPF GADD45B FAU SOD2 ITPR2 RPL30 NDUFB9 SLC26A2 KRT14 CD59 RPS5