

Supplementary information

An ignored pitfall in selecting reference genes and the solution with emphasis on subtypes and sex of thyroid cancer patients

Ghazal Esfandiarpour¹, Mohammad Mokhtari², Seyed-Morteza Javadirad*¹, Mohsen Kolahdouzan³, Ahmed Almuslimawi¹

Table 1: Primers information

Gene symbol (Genebank number)	Common name	Forward and reverse primers	Product length (bp)	Annealing temperature (°C)
GAPDH (NM_001256799.3)	Glyceraldehyde- 3-phosphate dehydrogenase	F:CCACTCCTCCACCTTTGACG R:CCACCACCCTGTTGCTGTAG	107	58
SYMPK (NM_004819.3)	Symplekin	F:ACGGTGCTGAGGGTCATTGA R:GAGGGTGGGACTTTGTCTGTGA	146	60

Table 1: Detailed information of Reference genes and their corresponding primers were listed. Genebank accession number, the sequences of forward (F) and reverse (R) primers, product length and annealing temperature of primers were presented.

GAPDH Glyceraldehyde-3-phosphate dehydrogenase, *SYMPK* Symplekin

Table 2: Detailed information of microarray datasets

Dataset Number	Accession Number	Tissue Samples	Platform	Country	Outliers
1.	GSE53157	27 samples: (5 PDTC, 7 cPTC,8 fvPTC, 4 FTC,3 Normal)	GPL570	Portugal	3 samples
2.	GSE65144	25 samples: (12 ATC, 13 Normal)	GPL570	USA	2 samples
3.	GSE53072	9 samples: (5 ATC, 4 Normal)	GPL6244	Portugal	1 samples
4.	GSE29315	71 samples: (9 Hurthle cell adenomas, 17 FA, 9 FTC, 13 FVPTC,6 Hashimoto thyroiditis,8 thyroid hyperplasias,9 PTC)	GPL8300	Belgium	34 samples: (including all hurthle cell adenomas, Hashimoto thyroiditis and thyroid hyperplasias)
5.	GSE27155	99 samples: (4 normal,10 FTA,13 FTC,7 oncocytic adenomas,8 oncocytic carcinomas,51 PTC,4 ATC, and 2 MTC)	GPL96	USA	26 samples: (including all oncocytic adenomas and oncocytic carcinomas)
6.	GSE33630	105 samples: (11 ATC, 49 PTC,45 Normal)	GPL570	Belgium	7 samples
7.	GSE104006	34 samples: (29 PTC, 5 Normal)	GPL14951	Italy	1 samples
8.	GSE82208	52 samples: (27 FTC, 25 FTA)	GPL570	Poland	14 samples
9.	GSE76039	37 samples: (20 ATC, 17 PDTC)	GPL570	USA	-
10.	GSE65074	38 samples: (PTC with different stages)	GPL13667	Denmark	-
11.	GSE9115	19 samples: (4 Normal, 10 PTC, 5 ATC)	GPL5917	USA	1 sample
12.	GSE6004	18 samples: (4 Normal, 14 PTC)	GPL570	USA	1 sample
13.	GSE3678	14 samples: (7 Normal, 7 PTC)	GPL570	USA	-
14.	GSE60542	92 samples: (29 lymph node,63 N and PTC)	GPL570	Belgium	30 sample

Table 2: Detailed information of microarray datasets including their accession numbers, tissue samples, platform, country and possible outliers were listed.

PTC papillary thyroid carcinoma, PDTC poorly differentiated thyroid cancer, cPTC classical papillary thyroid carcinoma, FVPTC follicular variant of papillary thyroid carcinoma, FTC follicular thyroid carcinoma, ATC anaplastic thyroid carcinoma, FTA follicular thyroid adenoma, MTC medullary thyroid carcinoma