**Описание найденных дифференциально экспрессированных генов и обоснование их отношения к гипоксии**

*Coding genes:*

1) **HMOX1** – ген, кодирующий фермент гемоксидазу 1, которая опосредует первый этап катаболизма гема, а именно превращение его в биливердин. Так как при гипоксии возможны изменения содержания гемоглобина, то дифференциальная экспрессия этого гена может быть вполне оправдана. Кроме того, регуляцию данного гена потенциально рассматривают как мишень для поддержки ткани во время гипоксического стресса.[1]

2) **PFKFB4** – ген, кодирующий фермент фруктоза-2,6-бисфосфатазу, который участвует в метаболизме одной из самых важных ключевых регуляторных молекул гликолиза – фруктоза-2,6-бисфосфата. При гипоксии происходит логичный переход от аэробного способа окисления глюкозы к преобладающему анаэробному. Более того, известно, что опухолевые клетки (на которых и производился анализ дифференциальной экспрессии) питаются преимущественно за счет гликолиза. Таким образом, дифференциальная экспрессия этого гена в условиях гипоксии так же заслуживает внимания. [2]

3) **NDRG1** – ген, кодирующий фермент, относящий к семейству гидролаз. Он играет важную роль в дифференцировании опухолевых клеток, их пролиферации и миграции, стрессовом ответе, иммунной реакции, ангиогенезе и других процессах. Примечательно, что данный ген может регулироваться гипоксией индуцированным фактором (HIF1 –a). С этой точки зрения его используют для моделирования гипоксии посредством непрямой активации тяжелыми металлами (никелем, кобальтом или железом).[3] В случае данной работы использовался хлорид кобальта.

4) **MAP1A** – ген, кодирующий белок, ассоциированный с микротрубочками. Несмотря на отсутствие очевидной прямой связи с гипоксическим состоянием, данный белок имеет большое значение для развития опухоли, влияя на ее цитоскелет и контролируя некоторые факторы роста (EGF – эпителиальный фактор роста).[4]

*Non-coding gene:*

1) **IGFL2-AS1** – длинная некодирующая РНК (lnRNA). Известно, что некодирующие РНК могут участвовать в различных клеточных процессах. Так и данная РНК наряду с другими РНК играет важную роль в биологических процессах в опухоли, включая гликолиз, активно функционирующий в условиях гипоксии.[5]

**Выбор потенциальных референсных генов**

Для начала необходимо провести анализ литературных источников, указывающих на наиболее часто используемые housekeeping genes.

Согласно одной из статей для анализа были отобраны следующие кандидаты референсных генов [6]:

ACTB – бета-актин, структурный компонент цитоскелета

B2M – бета-2-микроглобулин, компонент цитоскелета, участвующий в движении клетки, а также участнк презентации антигенов

GAPDH – глицеральдегид – 3 – фосфат – дегидрогеназа, фермент гликолиза

SDHA – сукцинатдегидрогеназный комплекс, участник ЭТЦ и ЦТК

HBG2 – гамма-глобулин, связывает и транспортирует кислород

По результатам исследования был сделан вывод, что выбор референсных генов индивидуален для каждого исследования с учетом особенностей используемой ткани или клеточной линии и конкретных условий эксперимента. Было показано, что даже экспрессии самых часто используемых housekeeping genes может значительно варьировать в зависимости от ситуации. Так, например, в данном исследовании наиболее стабильным оказался GAPDH, в отличие от наиболее популярного B2.

Это свидетельствует о том, что выбор референсных генов в случае нашей работы также не может достоверным на этом этапе и для дополнительной верификации требует дальнейших экспериментов о проверке стабильности экспрессии.

Однако можно отобрать 5 наилучших кандидатов из 10 найденных с помощью программы генов на основе сравнения их с наиболее популярными и исходя из общих представлений о требованиях к референсным генам.

Прежде всего необходимо отобрать гены с адекватной экспрессией (отбор по столцу baseMean в python. При отборе в группу приоритетных кандидатов попадут: WDR89, WDR24, BTN3A3, AC007566.1, FBXL13 (см. файл .ipynb)

Далее необходимо найти информацию о белках, кодируемых данными генами, не беря в расчет некодирующие гены, сделать выводы о вариабельности их экспрессии в зависимости от функции и сравнить их с каноническими референсными генами, приведёнными в статье, возможно объединяя их в определенные функциональные группы с переносом свойств известных housekeeping genes.

**WDR89(+WDR24**) – protein coding genes, значительно экспрессированные во многих тканях. Однако особенно примечательна их роль в развитии нервной системы, а именно неврных контактов – мутации в генах этой группы приводят к серьезным дефектам развития мозолистого тела. Более того, представители данной группы генов предположительно также участвуют в организации микротрубочкового аппарата [7]. Исходя из этого можно сделать косвенный вывод, что данный ген может иметь низкую вариабельность в экспрессии, так как отвечает за жизненно важные составляющие счуществования организма и ткани (клеток). Кроме того, возможное участие в организации цитоскелета клетки делает его похожим на ряд широко известных housekeeping genes, в том числе ACTB.

**BTN3A3** – protein coding gene, кодирует белок бутирофилин. Согласно информации, предоставленной в NCBI Protein, данный белок входит в группу главного комплекса гистосовместимости I (MHC I) и представляет собой мембранный рецептор, имеющий внеклеточный иммуноглобулиновый домен. Известно, что MHC взаимодействует с человеческим лейкоцитарным антигеном (HLA) и отвечает за базовые функции презентации антигенов иммунной системе. С этой точки зрения, он также может являться «оправданно» недифференциально экспрессированным, так как вариабельность экспрессии может потенциально привести к нарушению защитных механизмов клетки. Кроме того, один из наиболее часто используемых референсных генов – B2M – также участвует в иммунной реакции клеток, являясь «помощником» HLA.

**FBXL13** - protein coding gene, кодирует белок убиквитин-зависимой лигазы, участвующей в организации микротрубочкового аппарата посредством регуляции белков, присутствующих в центросоме клетки. Так он может участвовать в регуляции подвижности клетки, в том числе, миграции опухолевых клеток.[8] Следовательно, данный ген так же «оправдывает свою низкую вариабельность в экспрессии, кодируя белок с базовой функциональной нагрузкой.

В связи с тем, что функции **AC007566.1** ( также как и оставшихся генов среди отобранных 10 кандидатов) не были однозначно определены, было произведен увеличение количества первичных кандидатов до 11 и рассмотрен вновь пришедший кандидат **POLR2J2.**

**POLR2J2** - protein coding gene, согласно информации, предоставленной в NCBI Gene, кодирует субъединицу РНК-полимеразы II. Известно, что данная РНК-полимераза II отвечает за синтез матричной РНК (messenger RNA). Следовательно, данный ген жизненно необходим клетке, и значительная вариабельность его экспрессии может привести к значительным изменениям функционирования клетки. Однако стоит отметить, что данный ген относится к низкоэкспрессируемым, и не является идеальным кандидатом для использования в качестве референсного.

Таким образом, по результатам анализа и синтеза информации из литераькрных источников и логических выводов относительно свойств различных генов, были отобраны следующие референсные гены: **WDR89, WDR24, BTN3A3, FBXL13, POLR2J2(?)**

Стоит отметить, что данный анализ может содержать неточности, так как, безусловно, более основателбный и точный подбор референсных генов требует большего изучения.

*Список литературы:*

1) Wang W, Yang Q, Xie K, Wang P, Luo R, Yan Z, Gao X, Zhang B, Huang X, Gun S. Transcriptional Regulation of *HMOX1* Gene in Hezuo Tibetan Pigs: Roles of WT1, Sp1, and C/EBPα. Genes (Basel). 2020 Mar 26;11(4):352. doi: 10.3390/genes11040352. PMID: 32224871; PMCID: PMC7231170.

2) Yi M, Ban Y, Tan Y, Xiong W, Li G, Xiang B. 6-Phosphofructo-2-kinase/fructose-2,6-biphosphatase 3 and 4: A pair of valves for fine-tuning of glucose metabolism in human cancer. Mol Metab. 2019 Feb;20:1-13. doi: 10.1016/j.molmet.2018.11.013. Epub 2018 Dec 5. PMID: 30553771; PMCID: PMC6358545.

3) Park KC, Paluncic J, Kovacevic Z, Richardson DR. Pharmacological targeting and the diverse functions of the metastasis suppressor, NDRG1, in cancer. Free Radic Biol Med. 2020 Sep;157:154-175. doi: 10.1016/j.freeradbiomed.2019.05.020. Epub 2019 May 24. PMID: 31132412.

4) Lajoie-Mazenc I, Tovar D, Penary M, Lortal B, Allart S, Favard C, Brihoum M, Pradines A, Favre G. MAP1A light chain-2 interacts with GTP-RhoB to control epidermal growth factor (EGF)-dependent EGF receptor signaling. J Biol Chem. 2008 Feb 15;283(7):4155-64. doi: 10.1074/jbc.M709639200. Epub 2007 Dec 3. PMID: 18056259.

5) Cheng G, Liu D, Liang H, Yang H, Chen K, Zhang X. A cluster of long non-coding RNAs exhibit diagnostic and prognostic values in renal cell carcinoma. Aging (Albany NY). 2019 Nov 14;11(21):9597-9615. doi: 10.18632/aging.102407. Epub 2019 Nov 14. PMID: 31727869; PMCID: PMC6874440.

6) Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001 Dec;25(4):402-8. doi: 10.1006/meth.2001.1262. PMID: 11846609.

7) Kannan M, Bayam E, Wagner C, Rinaldi B, Kretz PF, Tilly P, Roos M, McGillewie L, Bär S, Minocha S, Chevalier C, Po C; Sanger Mouse Genetics Project, Chelly J, Mandel JL, Borgatti R, Piton A, Kinnear C, Loos B, Adams DJ, Hérault Y, Collins SC, Friant S, Godin JD, Yalcin B. WD40-repeat 47, a microtubule-associated protein, is essential for brain development and autophagy. Proc Natl Acad Sci U S A. 2017 Oct 31;114(44):E9308-E9317. doi: 10.1073/pnas.1713625114. Epub 2017 Oct 12. PMID: 29078390; PMCID: PMC5676932.

8) Fung E, Richter C, Yang HB, Schäffer I, Fischer R, Kessler BM, Bassermann F, D'Angiolella V. FBXL13 directs the proteolysis of CEP192 to regulate centrosome homeostasis and cell migration. EMBO Rep. 2018 Mar;19(3):e44799. doi: 10.15252/embr.201744799. Epub 2018 Jan 18. PMID: 29348145; PMCID: PMC583609