BBA - General Subjects 1862 (2018) 2140–2151

Contents lists available at ScienceDirect

BBA - General Subjects

journal homepage: www.elsevier.com/locate/bbagen

Review

A combined literature and in silico analysis enlightens the role of the NDRG family in the gut

Nathalie Vaes[a], Simone L. Schonkeren[a], Erwin Brosens[b], Alexander Koch[a], Conor J. McCann[c], Nikhil Thapar[c], Robert M.W. Hofstra[b], Manon van Engeland[a], Veerle Melotte[a,b,*]

[a] Department of Pathology, GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center, The Netherlands
[b] Department of Clinical Genetics, University of Rotterdam, EMC, Rotterdam, The Netherlands
[c] Stem Cells and Regenerative Medicine, UCL Great Ormond Street Institute of Child Health, London WC1N, UK

ARTICLE INFO

Keywords:
Biomarkers
Colorectal cancer
Intestinal tract
N-myc downstream-regulated gene
Tumor suppressors

ABSTRACT

Background: The N-Myc Downstream-Regulated Gene (*NDRG*) family comprises four members that function in cellular processes like proliferation and differentiation. While NDRG1 and NDRG2 are extensively studied, knowledge regarding NDRG3 and NDRG4, despite its recognition as a well-established early-detection marker for colorectal cancer (Cologuard®), is sparse.

Scope of review: To summarize expression, biomarker potential and functional mechanisms of the NDRGs in the developing, mature and cancerous gut, we combine current literature and *in silico* analyses from the TCGA-database, GTEX Project, E14.5 mouse intestine and enteric neural crest cells, and an RNA-sequencing time-series of human embryonic colonic samples.

Major conclusions: This study reveals that all members display a differential expression pattern in the gut and that NDRG1, NDRG2 and NDRG4 (1) can serve as biomarker for colorectal cancer and (2) have tumor suppressive properties mainly affecting cell proliferation and epithelial-mesenchymal transition.

General significance: Similar effects of the NDRGs on the key-hallmarks of cancer, could implicate analogous functions in other tissue/cancer types.

1. Introduction

The N-myc downstream-regulated gene (*NDRG*) family is composed of four members: *NDRG1*, *NDRG2*, *NDRG3* and *NDRG4*, and owes its name to the discovery of the first member, *NDRG1*, as being repressed by the *C-Myc* and *N-Myc* oncogenes [1]. In humans and mice, the four genes are all located on different chromosomes. Various aliases have been designated to each family member and each gene is transcribed into multiple alternatively spliced mRNA transcripts that encode several protein isoforms [2]. The encoded proteins of this family are all characterized by an NDR region and an α/β hydrolase fold, but lack the catalytic motif required to be enzymatically active [3,4]. The NDRG proteins which share about 52-65% sequence homology with an identical C-terminal sequence "MEVSC" and only very few sequence differences primarily in the C- and N-terminal regions, have been shown to be highly conserved in a variety of species [5,6]. More detailed information about the general knowledge on the NDRG family members, with respect to structure, origin and signaling in physiological processes can be found in our former review by Melotte et al. [2].

Previously, we and various independent groups identified *NDRG4* promoter CpG island methylation in fecal DNA as an accurate early-detection marker for colorectal cancer (CRC) [7–16]. The

Abbreviations: BDM1, Brain developmental-related molecule 1; Cav-1, Caveolin-1; CRC, Colorectal cancer; DAC, 5-aza-2'-deoxycytine; DFS, Disease free survival; EMT, Epithelial-mesenchymal transition; ENCC, Enteric neural crest cells; ENS, Enteric nervous system; EW, Embryonic week; FDR, False discovery rate; FIT, Fecal immunochemical test; GEO, Gene expression omnibus; GI, Gastrointestinal; GTEx, Genotype-tissue expression project; HMSNL, Hereditary motor and sensory neuropathy-Lom; IBD, Inflammatory bowel disease; LN, Lymph node; NDRG, N-myc downstream-regulated gene; OS, Overall survival; RKPM, Reads per kilobase million; SMAP8, Smooth muscle-associated protein 8; SNARE, Soluble NSF attachment protein receptor; SNP, Single nucleatoide polymorphism; TSA, Trichostatin A; TCGA, The cancer genome atlas; TSS, Transcription start site; UC, Ulcerative colitis; VAMP, Vesicle-associated membrane protein

* Corresponding author at: Dept. of Pathology, GROW – School for Oncology and Developmental Biology, Maastricht University Medical Center, P.O. Box 616, 6200 MD, Maastricht, The Netherlands.

E-mail addresses: n.vaes@maastrichtuniversity.nl (N. Vaes), s.schonkeren@student.maastrichtuniversity.nl (S.L. Schonkeren), e.brosens@erasmusmc.nl (E. Brosens), a.koch@maastrichtuniversity.nl (A. Koch), conor.mccann@ucl.ac.uk (C.J. McCann), n.thapar@ucl.ac.uk (N. Thapar), r.hofstra@erasmusmc.nl (R.M.W. Hofstra), m.vanengeland@maastrichtuniversity.nl (M. van Engeland), veerle.melotte@maastrichtuniversity.nl (V. Melotte).

https://doi.org/10.1016/j.bbagen.2018.07.004
Received 4 April 2018; Received in revised form 29 June 2018; Accepted 5 July 2018
Available online 07 July 2018

biomarker performance of *NDRG4* was further exploited by Exact Sciences, who integrated *NDRG4* promoter methylation into a multi-target stool DNA test: Cologuard® (Madison, USA) [16]. This FDA-approved stool DNA test detects significantly more CRCs and advanced precancerous lesions than the leading fecal immunochemical test (FIT) and is currently used in the USA for CRC screening [7]. Despite being incorporated into the Cologuard®, almost nothing is known about the function(s) of NDRG4 in the (diseased) intestinal tract. Whereas NDRG1 and NDRG2 are important in various cellular processes (e.g. proliferation, differentiation, apoptosis), little is known about NDRG3 and NDRG4.

Here, we performed a literature search (Pubmed, Web of Science, Embase and Medline) and supported this literature search with *in silico* data extracted from (i) publicly available transcriptome databases [17,18], (ii) previously generated transcriptome data of E14.5 mouse intestine and enteric neural crest cells (ENCCs) [19] and (iii) an RNA-sequencing time-series of human embryonic colonic samples, to gain more insights into the expression pattern, biomarker potential and common pathways of the NDRGs in the intestinal tract.

Our review indicates that *NDRG1*, *NDRG2* and *NDRG4* can act as biomarkers for CRC. Although there is no evidence that the different family members are major drivers in the development of CRC, data suggest that, except for NDRG3, each NDRG member has a protective role during intestinal carcinogenesis.

2. Expression pattern of the NDRGs in the developing and adult gut

To understand the functional importance of the NDRG family members in the intestinal tract, we first elaborate on their structure and summarize their (overlapping) intestinal-specific expression pattern in mammalian species like rat, mouse and human.

In humans, ***NDRG1*** is located on chromosome 8q24.3, while the mouse ortholog is located on chromosome 15D2. The encoded canonical protein sequence of human NDRG1 (DRG1, RTP, RIT42, CAP43) consists of 394 amino acids (AA), resulting in a 43 kDa protein. The two other isoforms are made up of 328 and 313 AA, with a respective mass of 35.5 and 34 kDa. In mouse and rat, only one isoform has been found, with a similar length as the human canonical protein sequence and minor AA differences.

The expression of the *NDRG1* gene usually coincides in tissues where *N-myc* is expressed, but only initiates at E9.5 as the intestinal epithelium begins to differentiate and *N-myc* expression decreases [20]. As further intestinal development requires retinoic acid-dependent activation of *NDRG1* and suppression of the Wnt/β-catenin pathway [21], *NDRG1* mRNA expression steadily increases in the mouse and human intestinal epithelium from E10.5 onwards [20]. Further intestinal maturation in rodents and humans seems to be controlled by circulating levels of glucocorticoids and cortisol, respectively [22], giving rise to similar levels of NDRG1 in the adult small and large intestine [5,22–24]. Moreover, the expression of NDRG1 mRNA and protein is induced by glucocorticoid treatment (i.e. by dexamethasone) and strongly overlaps with a gradual increase from crypt cells towards surface cells [20,22,24,25]. On a sub-cellular level, NDRG1 is primarily localized in the cytoplasm of epithelial cells, where it is often associated with the basolateral membrane adjacent to adherens junctions and desmosomes [20,24,26].

NDRG2 is located on chromosome 14q11,2 and 14C1 in human and mouse, respectively. The main functional form of human, mouse and rat NDRG2 (SYLD, NDR2) is either a 41 kDa protein, composed of 371 AA, or a 39 kDa protein of 357 AA. In addition, there are four other human isoforms with a length ranging from 328 to 367 AA and a variable mass between 36 and 40.3 kDa.

In early developmental stages (E8.5-9.5) the poorly differentiated intestinal epithelium expresses low levels of NDRG2 mRNA and protein. From E10.5 towards adulthood, the levels of NDRG2 mRNA and protein very slightly increase in the mucous forming, lining epithelial cells [20,27], leading to barely detectable levels in adult gut tissue. Even though most studies observed a slightly higher level of *NDRG2* mRNA in the (distal) colon compared to the small intestine [5,28,29], contradictory observations regarding NDRG2 protein expression have been described. Yamamoto et al. detected a marginally higher signal in colon than in the small intestine [30], whereas Hu et al. detected opposite results with weak to moderate NDRG2 immunoreactivity in the small intestine but not in the colon [31].

Human chromosome 20q11.21-11.32 and mouse chromosome 2H1 contain the ***NDRG3*** gene, which can be transcribed into three identified protein isoforms. The canonical sequence of *NDRG3* (no known aliases) encodes a 41.5 kDa protein of 375 AA in various mammalian species. The second and third isoform are only found in humans and encode a 40 kDa protein of 363 AA and a markedly lighter protein (31.5 kDa) of only 286 AA.

Only three studies include a brief description about the expression of *NDRG3*. In fact, *in situ* mRNA hybridization (ISH) and Northern Blotting reveal that *NDRG3* expression is already activated at E9.5 and that *NDRG3* is widely expressed during embryogenesis. However no *NDRG3* mRNA is detected in the developing intestinal tract [20]. During further development, overall NDRG3 mRNA and protein levels increase slightly, although they are hardly detectable in matured mouse and human gut tissues [5,30].

The ***NDRG4*** gene is situated on human chromosome 16q21-q22.1 and mouse chromosome 8D1. For NDRG4 (SMAP8, BDM1), the canonical protein sequence is referred to as NDRG4B^{var}, consisting of 352 AA (39 kDa). The two other main isoforms are NDRG4B and NDRG4H which contain 339 AA (37 kDa) and 371 AA (41 kDa), respectively. In addition, 5 other isoforms can be produced by alternative splicing, with sequences ranging from 339 to 391 AA (37-43 kDa). In mouse, we recently identified three isoforms, corresponding to human NDRG4B^{var} (long isoform), NDRG4B (short isoform) and NDRG4H [32]. Furthermore, rat tissues express an additional fourth isoform of 45kDa, encoded by the rat ortholog: *SMAP8/BDM1* [33].

So far, the pattern of NDRG4 expression during embryonic gut development has not yet been described. In adult tissues however, current literature is inconsistent regarding the intestinal-specific expression pattern of NDRG4 which can be attributed to the use of different, a-specific commercially-available antibodies [32]. We and Chu et al. previously described that NDRG4 is expressed within epithelial cells of the gut [12,34] and Qu et al. found high NDRG4 protein levels in smooth muscle cells of the stomach [35]. However, we recently observed that the anti-NDRG4 antibody used in these studies was not specifically targeting NDRG4 [32]. The experimental application of an antibody that specifically targets NDRG4 has demonstrated that NDRG4 expression in the gut is restricted to the cytoplasm of neuronal cell bodies and nerve fibers belonging to the nervous system of the gut, i.e. the enteric nervous system (ENS) [32].

2.1. *In silico expression analysis*

To clarify the discrepancies concerning the NDRG expression pattern in the developing/mature gut observed in the limited number of publications, we evaluated expression data from the (i) GTEx Project (06/19/2017), (ii) previously generated transcriptome data of E14.5 mouse intestine and ENCCs (available in the Gene Expression Omnibus (GEO); GSE34208) [19] and (iii) an RNA seq time-series of human embryonic full-thickness samples of proximal colon (Embryonic Week (EW) 12, EW14 and EW16; data generated by the group of Prof. Thapar (UCL, London) and deposited to GEO).

Compared to its family members, ***NDRG1*** is highly expressed in EW12, EW14 and EW16 human colonic samples, with levels (in reads per kilobase million, RPKM) being 10, 8 and 12 times higher than for *NDRG2, 3* and *4*, respectively (Fig. 1A). In addition, *NDRG1* shows the highest and intercomparable expression in sections of the mature small