Published in final edited form as:

Eur J Cancer. 2020 November ; 140: 119–129. doi:10.1016/j.ejca.2020.09.006.

The impact of *ARID1A* mutation on molecular characteristics in colorectal cancer

Ryuma Tokunaga^{1,8,*}, Joanne Xiu², Richard M. Goldberg³, Philip A. Philip⁴, Andreas Seeber⁵, Francesca Battaglin¹, Hiroyuki Arai¹, Jae Ho Lo¹, Madiha Naseem¹, Alberto Puccini¹, Martin D. Berger¹, Shivani Soni¹, Wu Zhang¹, Sting Chen², Jimmy J. Hwang⁶, Anthony F. Shields⁴, John L. Marshall⁷, Hideo Baba⁸, W. Michael Korn^{2,9}, Heinz-Josef Lenz¹

¹Division of Medical Oncology, Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, USA.

²Caris Life Sciences, Phoenix, USA

³West Virginia University Cancer Institute, Morgantown, USA

⁴Department of Oncology, Karmanos Cancer Institute, Wayne State University, Detroit, USA

⁵Department of Haematology and Oncology, Innsbruck Medical University, Innsbruck, Austria

⁶Levine Cancer Institute, Carolinas HealthCare System, Charlotte, USA

⁷Ruesch Center for The Cure of Gastrointestinal Cancers, Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Washington, USA

*Corresponding author: Ryuma Tokunaga, Division of Medical Oncology, Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, 1441 Eastlake Avenue, Los Angeles, CA 90033, USA. Tel: +1-323-865-3967; mappymap@hotmail.co.jp.

Authors' Contributions:

Conceptualization: R. Tokunaga, J. Xiu, H.-J. Lenz

Data curation: R. Tokunaga, J. Xiu

Formal analysis: R. Tokunaga, J. Xiu

Funding acquisition: R. Tokunaga, H.-J. Lenz

Investigation: R. Tokunaga, J. Xiu

Methodology: R. Tokunaga, H.-J. Lenz

Project administration: H.-J. Lenz

Resources: J. Xiu, S. Chen, W.M. Korn

Software: R. Tokunaga, J. Xiu, S. Chen, W.M. Korn

Supervision: H.-J. Lenz

Validation: R. Tokunaga, J. Xiu, R.M. Goldberg, P.A. Philip, A. Seeber, H. Arai, F.

Battaglin, W. Zhang, A.F. Shields, J.L. Marshall, W.M. Korn, H.-J. Lenz

Visualization: R. Tokunaga, J. Xiu, H.-J. Lenz

Roles/Writing - original draft: R. Tokunaga

Writing - review & editing: R. Tokunaga, J. Xiu, R.M. Goldberg, P.A. Philip, A. Seeber, M. Naseem, J.H. Lo, H. Arai, F. Battaglin, A. Puccini, M.D. Berger, S. Soni, W. Zhang, S. Chen, J.J. Hwang, A.F. Shields, J.L. Marshall, H. Baba, W.M. Korn, H.-J. Lenz

Use of standardized official symbols:

In this study, HUGO (Human Genome Organisation)-approved official symbols for genes and gene products were used: all of which are described at www.genenames.org.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

⁸Department of Gastroenterological Surgery, Graduate School of Medical Sciences, Kumamoto University, Kumamoto, Japan

⁹Division of Hematology/Oncology, University of California San Francisco, San Francisco, USA

Abstract

Background: ARID1A is a key subunit of the SWI/SNF complex which regulates dynamic repositioning of nucleosomes to repair DNA damage. Only small pilot studies have evaluated the role of *ARID1A* mutation in colorectal cancer (CRC). The aim of present study was to explore the potential impact of *ARID1A* mutation on clinicopathological and molecular characteristics in CRC.

Methods: We used integrated datasets of 7,978 CRC cases (one data set from a CLIA-certified laboratory and three independent published data sets). The associations of *ARID1A* mutation with molecular characteristics including immune profile [the status of microsatellite instability (MSI), tumor mutational burden (TMB), programmed death ligand 1 (PD-L1), and estimated-infiltrating immune cells], clinicopathological features, and related pathways were analyzed using next-generation sequencing, RNA-sequencing, and immunohistochemistry.

Results: *ARID1A* mutant samples had more genomically unstable tumor features (MSI-high and TMB-high) and exhibited more characteristics of a T-cell-inflamed microenvironment [PD-L1 expression, and high estimated-infiltrating cytotoxic T lymphocytes (CTLs)] than *ARID1A* wild-type samples in the discovery and validation cohorts. Even *ARID1A* mutant samples without MSI-high status were TMB-high, had high levels of PD-L1 expression, and high estimated-infiltrating CTLs. *ARID1A* mutations were more common with right-sided primary and earlier stage tumors. *ARID1A* mutant tumors mainly had co-occurring gene mutations related to Chromatin modifying, DNA repair, WNT signaling, and EGFR inhibitor resistance pathways, and *ARID1A* mutations strongly regulated DNA repair pathways. Key genes for chemo/radio-therapy sensitivity were suppressed in *ARID1A* mutant samples.

Conclusions: Our findings may provide novel insight to develop individualized approaches for treatment of CRC based on *ARID1A* mutation status.

Keywords

ARID1A; mutation; colorectal cancer; SWI/SNF complex; molecular profile; immune profile; right-sided location

Introduction

ARID1A (also known as BAF250a) is a subunit of the SWI/SNF complex, an ATPase chromatin remodeling/tumor suppressor complex [1]. The SWI/SNF complex regulates lineage-specific transcription via nucleosome remodeling-based control and promotes DNA repair through recruiting DNA repair proteins. Therefore, loss of the SWI/SNF complex function is a common feature of tumors and can drive genomic instability and tumor progression [2]. Preclinical evidence suggests that the alterations in the subunit genes of the SWI/SNF complex drive changes in tumor biology that may augment the antitumor immune response and chemo/radio-therapy response [3, 4]. Of the subunits of the SWI/SNF complex,

ARID1A has the highest mutation frequency in cancers and mutations (loss-of-function) are known to be associated with tumor activity and poor prognosis [5]. In colorectal cancer (CRC), *ARID1A* mutation frequency is enriched in tumors with microsatellite instability (MSI) status [6], however its impact on pathological and molecular features remains largely unclear.

The availability of immune checkpoint inhibitors has changed the landscape of treatment strategies in many cancer patients, including a subset of patients with CRC [7]. However, a minority of CRC patients benefit from these treatments and the identification of reliable biomarkers predictive of response is one of the most pressing missions [8]. In the clinic, tumors exhibiting high levels of microsatellite instability (MSI-H), tumor mutational burden (TMB), and expression of programmed death ligand 1 (PD-L1) are more likely to respond to immune-based treatment strategies [9–11]. Recently, Shen et al. reported that *ARID1A* deficiency were associated with MSI-high and increased TMB, and that *ARID1A* mutations could be a predictive biomarker of response to the anti-PD-L1 antibody in ovarian cancer [12]. However, the relationship between *ARID1A* mutations and molecular profiles, including immune profiles, of the tumor in CRC are unclear. Further understanding of tumor characteristics and genetic features related to *ARID1A* mutations may be critical in future tailored-treatment for CRC.

We performed this study to highlight the impact of *ARID1A* mutations on molecular characteristics to investigate their potential value to influence treatment strategies in CRC, using integrated datasets generated from analyses of a total of 7,978 samples. Our analysis demonstrates that *ARID1A* mutation status could be a predictive biomarker of immune therapies and a novel key step towards personalized treatment strategy for CRC.

Patients and methods

Tumor Samples—A data set from a commercial CLIA-certified laboratory (Caris Life Sciences, Phoenix, USA) from April, 2015 to January, 2018 (named CARIS) and three published data sets (named DFCI [13], MSK [14], and TCGA) available through cBioportal website (“DFCI, Cell Reports 2016”, “MSK, Cancer cell 2018”, and “TCGA, PanCancer Atlas” datasets) were used for our analysis [15]. In the CARIS dataset, human subjects were anonymized prior to analysis. The subjects were de-identified and this research was determined to exempt from IRB review per policy 45 CFR 46.101 (b). GSE 59857 available through NCBI’s Gene Expression Omnibus and *ARID1A* mutation data available through cancer cell line encyclopedia (CCLE) were used for pathway analysis,

Next-generation sequencing (NGS) and immunohistochemistry (IHC) in the CARIS dataset—Next-generation sequencing (NGS) was performed using the NextSeq platform (Illumina, Inc., San Diego, USA). A custom-designed SureSelect XT assay was used to enrich 592 whole-gene targets (Agilent Technologies, Santa Clara, USA; <https://www.precisiononcology.com.au/doctors/whats-included/>). The cases with unclassified genetic variants ($n = 194$) or indeterminate low coverage regions after exon 2 ($n = 33$) for *ARID1A* were excluded. MSI status was examined using over 7,000 target microsatellite loci and compared to the reference genome hg19 from the University of California, Santa

Cruz (UCSC) Genome Browser database, and the status was defined as MSI-H or MSI-low (MSI-L)/microsatellite stable (MSS). The threshold to define TMB-high (TMB-H) was greater than or equal to 10 mutations/megabases (Mb). Immunohistochemistry (IHC) analysis was performed for PD-L1 expression status on full slides of formalin fixed paraffin-embedded (FFPE) tumor specimens using automated staining techniques (Benchmark XT, Ventana, and Autostainer Link 48, Dako) and PD-L1 SP142 clone (R&D Systems, Minneapolis, USA).

Statistical analysis—All statistical analyses were performed with JMP v10.0 (SAS Institute Inc., Cary, USA), or R v.3.5.1 (R Development Core Team). The comparisons of continuous data were analyzed using Student's t-test, and those of categorical data were analyzed using Fisher's exact test. RNA-seq based MCP-counter, which allows the quantification of eight immune and two stromal cell populations in heterogeneous tissues, was performed for estimation of the association between *ARID1A* mutant tumor and tumor infiltrating immune/stromal cells population [16]. A downloadable R package (CMSclassifier: <https://github.com/Sage-Bionetworks/CMSclassifier>) from CRC Subtyping Consortium in the original CMS classification [17] was obtained and used to filter the dataset. GSEA v3.0 and MORPHEUS websites were used for pathway analysis. Materials and Methods are detailed in Supplementary materials and methods.

Results

ARID1A mutations and molecular characteristics in the discovery cohort

We studied samples from 7,978 patients with CRC from four independent cohorts. Figure 1 summarizes the workflow of this study. Accumulating evidences suggest that *ARID1A* does globally impact gene expression through regulating chromatin remodeling (Figure 2A). We therefore first evaluated the *ARID1A* mutation status and examined the association with molecular characteristics including immune profiles in the discovery cohort (CARIS cohort).

Figure 2B shows the observed mutation locations in the *ARID1A* gene. The mutation rate was 8.2% (468/5,726 samples); the most frequent location was D1850 (13.0%: 73/560 mutations); and the most frequent type of mutation was a truncating mutation (71.4%: 400/560), such as a frameshift mutation. The observed patterns of common gene mutations in CRC were similar to past reports (Figure 2C). The immune profile in *ARID1A* mutant samples was strikingly different from that in *ARID1A* wild-type samples. MSI-H, TMB-H, and PD-L1 expression rates were significantly higher in *ARID1A* mutant samples (45.3%, 60.1%, and 18.3%) than in *ARID1A* wild-type samples (2.9%, 21.7%, and 5.9%) (Figure 2D). The samples with somatic *POLE* mutation or with MSI-H exhibited a substantially higher TMB. Even in samples without MSI-H status, TMB-H rate was significantly higher in *ARID1A* mutant samples than in *ARID1A* wild-type samples (27.7% vs 19.5%; $P=0.003$) (Figure 2E). Figure 2F shows the patterns of co-occurring and mutual exclusive genes among *ARID1A* mutant samples. Among the top 20 genes mutated in the CARIS panel, we found 3 mutually exclusive genes mutation (*TP53*, *APC*, and *KRAS*), and 14 co-occurring gene mutations that have been implicated in the regulation of Chromatin modifying, DNA repair, WNT signaling, or EGFR inhibitor resistance pathways (Supplementary Table S1).

Relationship between ARID1A mutation and clinicopathological characteristics

We next evaluated the association between *ARID1A* mutation and clinicopathological characteristics in each cohort. Baseline patient characteristics are shown in Supplementary Table S2. The CARIS cohort had no information of the TNM stage, and the DFCI and TCGA cohorts had no information about the location of tumor sampling. The CARIS cohort had fewer MSI-H tumors. The DFCI cohort had more elderly, female, right-sided, and MSI-H tumors. The MSK cohort had more stage IV tumors. The TCGA cohort had more elderly, right-sided, MSI-H, and *ARID1A* mutant tumors. Table 1 showed that there were strong associations of *ARID1A* mutant samples with right-sided tumors (CARIS: $P < 0.001$, DFCI: $P = 0.053$, MSK: $P < 0.001$, TCGA: $P = 0.002$), earlier tumor stage (DFCI: $P = 0.004$, MSK: $P < 0.001$, TCGA: $P = 0.046$), more tumor sampling from primary tumor (CARIS: $P < 0.001$, MSK: $P < 0.001$), and MSI-H (CARIS: $P < 0.001$, DFCI: $P < 0.001$, MSK: $P < 0.001$, TCGA: $P < 0.001$).

ARID1A mutation and tumor microenvironment

We further evaluated the association of *ARID1A* mutation with *PD-L1* expression, estimated-tumor infiltrating immune and stromal cells using the MCP-counter method, and CMS classifier using the Random Forest classifier in the TCGA cohort (Figure 3, Supplementary figure S1A). *ARID1A* mutant samples had higher *PD-L1* expression level ($P < 0.001$) as well as more estimated-infiltrating cytotoxic T lymphocytes (CTLs) ($P < 0.001$) and monocyte lineage cells ($P < 0.001$) than *ARID1A* wild-type samples (Figure 3A). Even in *ARID1A* mutant samples without MSI-H status, *PD-L1* was highly expressed ($P < 0.001$) and more CTL infiltration was expected ($P = 0.003$) (Figure 3B). The pattern of CMS distribution was different between *ARID1A* mutant samples and *ARID1A* wild-type samples. *ARID1A* mutant samples showed a higher prevalence of CMS1 (36% versus 11%) than *ARID1A* wild-type samples, as did in the samples without MSI-H status (19% versus 7%) (Figure 3C, D).

ARID1A mutation and leading pathways

Finally, we assessed the *ARID1A* mutation status and molecular characteristics in the validation cohorts (DFCI, MSK, and TCGA cohorts) and examined the *ARID1A* related-pathways using gene set enrichment analysis in the TCGA and GSE 59857 cohorts.

The *ARID1A* mutation rate was 9.4% (212/2,252 samples); the most frequent location was D1850 (8.7%: 24/275 mutations): and the most frequent type was a truncating mutation (62.2%: 171/275 mutations), similar to the discovery cohort (Supplementary figure S1B). The mutated gene list is shown in Supplementary table S3. OncoPrint (Figure 4A) showed that *ARID1A* mutant samples had higher numbers of gene mutations, right-sided primary tumors, and MSI-H tumors. Figure 4B further pointed to co-occurring mutated genes that have been implicated in the regulation of Chromatin modifying, DNA repair, WNT signaling, or EGFR inhibitor resistance pathways (Supplementary table S4). Among the 20 genes evaluated in the discovery cohort, co-occurring genes mutation were all validated, and only *TP53* mutations and *ARID1A* mutations were mutually exclusive.

Gene set enrichment analysis identified signaling pathways that might relate to *ARID1A* mutant status (Figure 4C). *ARID1A* mutation was strongly associated with DNA repair pathway. However WNT and EGFR inhibitor resistance pathways were not upregulated in *ARID1A* mutant samples. Further analysis of genes expression in DNA repair pathway showed that the key genes mediating chemo/radio-therapy sensitivity, such as *ATM*, *ATR*, *MRE11A*, *RAD50*, and *BRCA1* were significantly suppressed in *ARID1A* mutant samples (Supplementary figure S2).

Discussion

We report for the first time the impact of *ARID1A* mutations on molecular characteristics in CRC. With a total of 7,978 clinical samples from four independent cohorts, we found that *ARID1A* mutation was significantly associated with a favorable immune profile indicative of a higher likelihood of response to immune check point inhibitors (MSI-H, TMB-H, PD-L1 expression, and high estimated-infiltrating CTLs). It was also associated with right-sided primary tumor location, and earlier tumor stage. In addition, we determined that *ARID1A* mutant tumors mainly had co-occurring gene mutations related to Chromatin modifying, DNA repair, WNT signaling, and EGFR inhibitor resistance pathways, and manifested mutual exclusivity with *TP53* mutation. Also, gene expression-based pathway analysis showed that *ARID1A* mutation strongly regulated DNA repair pathway related to chemo/radio-therapy sensitivity. Our findings support developing a personalized treatment strategy in CRC patients with *ARID1A* mutations.

ARID1A is a key subunit of the multiprotein SWI/SNF complexes that regulate the dynamic repositioning of nucleosomes to repair DNA damage for gene transcription [3]. Accumulating evidences suggest that the loss of ARID1A function decreases expression of other SWI/SNF subunits, impacting chromatin remodeling ability [18]. Therefore, *ARID1A* mutation could epigenetically impact genetic stability and diverse biological pathways [19]. We highlighted the association between *ARID1A* mutation and favorable immune profile for immune check point inhibitors (MSI-H, TMB-H, PD-L1 expression, and high estimated-infiltrating CTLs) in CRC. Importantly, even *ARID1A* mutant samples without MSI-H status were associated with TMB-H, PD-L1 expression, and high estimated-infiltrating CTLs. In addition, *ARID1A* mutant samples (even without MSI-H status) showed a higher prevalence of CMS1 than *ARID1A* wild-type samples, which means *ARID1A* mutation could lead to strong immune activation. Recent data showed that ARID1A-deficient tumors had a remarkable response to anti-PD-L1 antibody in ovarian cancer [12]. Since immune checkpoint inhibitors are more likely to be effective against genomically unstable tumors expressing neoantigens and tumors with a T-cell-inflamed microenvironment [20], *ARID1A* mutation, being associated with both, may be associated with good immunotherapy response.

ARID1A mutation results in loss of tumor-suppressor function, resulting in enhanced tumor proliferation, metastasis, and recurrence [5, 21]. In preclinical CRC mouse models, Mathur et al. demonstrated that ARID1A inactivation increased the invasive ability and presence of tumor-infiltrating lymphocytes, features associated with MSI-H tumors [22]. Further, conditional *ARID1A* knockout in mice promoted tumorigenesis of ovarian cancer, CRC, and

pancreatic ductal adenocarcinoma (in KRAS-activated model), and *ARID1A* mutations are associated with poor prognosis [5, 22, 23]. However, surprisingly, our current study showed that *ARID1A* mutation had a strong association with earlier tumor stage. Although loss of ARID1A results in invasiveness, at the same time it might improve natural antitumor immunity on its own in a manner analogous to observations in MSI-H tumors [12, 24]. Our data also demonstrated that *ARID1A* mutation had a strong association with right-sided primary tumor location. In CRC, differences in characteristics due to tumor location have been reported, and it seems to have a great impact on its treatment strategy [25]. Therefore, future studies are required to clarify precisely how *ARID1A* mutation-related inactivation contributes to the clinical treatment strategy.

In cancer, individual tumors typically have multiple somatic alterations, and the co-occurring or mutually exclusive alterations could suggest functional interactions [26]. Co-occurrence indicates functional synergies and reflects resistance to targeted therapy. Mutual exclusivity reflects alternative functions in tumor initiation and progression, which provide new insights into tumorigenesis and have potential implications for clinical management. Our data pointed to co-occurring gene mutations related to Chromatin modifying, DNA repair, WNT signaling, and EGFR inhibitor resistance pathways. Consistent with our findings, past studies showed the interdependent interaction between ARID1A and chromatin remodeling and DNA repair proteins [2, 3, 19, 27], a negative association of WNT signaling with lymphocyte infiltration which related to *ARID1A* mutation status [28], and EGFR inhibitor resistance in ARID1A deficient tumors through activation of PI3K/AKT pathway [24]. *ARID1A* mutations might have effects on both epigenetic regulation and immune profiles, and could be a critical determinant for the effect of anti-EGFR therapy. Our findings also show that mutations in *ARID1A* and *TP53* occur uncommonly in the same tumor. This may reflect distinct mechanisms of tumorigenesis between tumor subtypes. *ARID1A* mutation leads to defective cell cycle checkpoint activation, as well as *TP53* mutation in response to DNA damage [2]. Although strong chemoresistance has been reported for tumors with MSI-H positively associated with *ARID1A* mutation [29], additional studies of the impact of *ARID1A* mutation on chemo/radio-therapy sensitivity through regulating DNA repair pathways are needed.

Our findings suggest the potential for the development of more effective treatment strategies for CRC based on the presence of *ARID1A* mutations. Specifically, our findings indicate that *ARID1A* mutation could be a promising target for immune checkpoint based therapies and might be a predictive biomarker for anti-EGFR therapy. Recently, epigenetic regulator and DNA repair proteins having synthetic lethality interaction with *ARID1A* mutation have been reported, and inhibition of such proteins including EZH2, HDAC, BET, and PARP resulted in selective vulnerabilities in *ARID1A* mutant tumors [19, 27, 30–32]. Since tumors with *ARID1A* mutations have impaired SWI/SNF control of enhancers, leading to reduced H3K27ac levels and expression of nearest genes [22], and have the DNA repair pathway inhibited [27], inhibition of such proteins may extensively impair epigenetic regulation and DNA repair, which could result in synthetic lethality. In future studies, it will be of interest to investigate the integrated effects of SWI/SNF complex proteins and these synthetic lethal interactions on clinical and molecular characteristics in CRC.

Limitations exist in this study. Due to the lack of treatment and prognosis information, the direct effects of our findings on clinical outcomes are unclear. In addition, this study only showed the association between *ARID1A* mutation and molecular/immunological features, not the basic biological role of *ARID1A* in CRC. The retrospective study design could not exclude selection bias. However, using the biggest dataset ($N = 7,978$) from four independent cohorts, our findings supported the past evidence and determined the impact of *ARID1A* mutation on clinicopathological and molecular characteristics including immune profile in CRC.

Conclusion

ARID1A mutation status has clinical impact on CRC. Our findings suggest that tumor with *ARID1A* mutation may be sensitive to immune check point inhibitors and could be a predictive biomarker for chemo/radio-therapy and targeted therapy through co-occurring mutation and regulating pathways. Upon validation with clinical features, our findings may develop individualized approach for CRC treatment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We thank Dr. Masafumi Horie for his support of the analysis. We also thank all the specimen donors and research groups of data sets.

Funding: This work was supported by the Uehara Memorial Foundation, the National Cancer Institute (grant number P30CA014089), Dhont Family Foundation, San Pedro Peninsula Cancer Guild, and Daniel Butler Research Fund.

Conflict of interests: P.A. Philip reports receiving speakers bureau honoraria from Merck, Celgene, Halozyme, Ipsen, Bristol-Myers Squibb, and Bayer, and is a consultant/advisory board member for Halozyme, Caris, and Merck. A. Seeber is a consultant/advisory board member for Caris Life Sciences. A.F. Shields is a consultant/advisory board member for and reports receiving commercial research grants from Caris. J.L. Marshall is an employee of and has ownership interests (including patents) at Caris; reports receiving speakers bureau honoraria from Genentech, Amgen, Celgene, Taiho, and Bayer; and is a consultant/advisory board member for Caris SAB. H. Baba reports receiving speakers bureau honoraria from Eli Lilly Japan K.K. W.M. Korn has ownership interests (including patents) at Caris Life Sciences and is a consultant/advisory board member for Merck and Lilly. H.-J. Lenz reports receiving speakers bureau honoraria from and is a consultant/advisory board member for Merck Serono, Bayer, and Genentech. No potential conflicts of interest were disclosed by the other authors.

Abbreviations:

CRC	colorectal cancer
CTL	cytotoxic T lymphocyte
IHC	immunohistochemistry
Mb	megabase
MSI	microsatellite instability
MSI-H	MSI-high

MSI-L	MSI-low
MSS	microsatellite stable
NGS	next-generation sequencing
PD-L1	programmed death ligand 1
TMB	tumor mutational burden
TMB-H	TMB-high

References

1. Wu JN, Roberts CW. ARID1A mutations in cancer: another epigenetic tumor suppressor? *Cancer Discov* 2013;3(1):35–43. [PubMed: 23208470]
2. Watanabe R, Ui A, Kanno S, Ogiwara H, Nagase T, Kohno T, et al. SWI/SNF factors required for cellular resistance to DNA damage include ARID1A and ARID1B and show interdependent protein stability. *Cancer Res* 2014;74(9):2465–75. [PubMed: 24788099]
3. Wilson BG, Roberts CW. SWI/SNF nucleosome remodellers and cancer. *Nat Rev Cancer* 2011;11(7):481–92. [PubMed: 21654818]
4. Zhao J, Liu C, Zhao Z. ARID1A: a potential prognostic factor for breast cancer. *Tumour Biol* 2014;35(5):4813–9. [PubMed: 24430365]
5. Luchini C, Veronese N, Solmi M, Cho H, Kim JH, Chou A, et al. Prognostic role and implications of mutation status of tumor suppressor gene ARID1A in cancer: a systematic review and meta-analysis. *Oncotarget* 2015;6(36):39088–97. [PubMed: 26384299]
6. Jones S, Li M, Parsons DW, Zhang X, Wesseling J, Kristel P, et al. Somatic mutations in the chromatin remodeling gene ARID1A occur in several tumor types. *Hum Mutat* 2012;33(1):100–3. [PubMed: 22009941]
7. Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. *Cell* 2017;168(4):707–723. [PubMed: 28187290]
8. Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *Lancet Oncol* 2016;17(12):e542–e551. [PubMed: 27924752]
9. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med* 2015;372(26):2509–20. [PubMed: 26028255]
10. Hellmann MD, Ciuleanu TE, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, et al. Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden. *N Engl J Med* 2018;378(22):2093–2104. [PubMed: 29658845]
11. Garon EB, Rizvi NA, Hui R, Leighl N, Balmanoukian AS, Eder JP, et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med* 2015;372(21):2018–28. [PubMed: 25891174]
12. Shen J, Ju Z, Zhao W, Wang L, Peng Y, Ge Z, et al. ARID1A deficiency promotes mutability and potentiates therapeutic antitumor immunity unleashed by immune checkpoint blockade. *Nat Med* 2018;24(5):556–562. [PubMed: 29736026]
13. Giannakis M, Mu XJ, Shukla SA, Qian ZR, Cohen O, Nishihara R, et al. Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma. *Cell Rep* 2016;15(4):857–865. [PubMed: 27149842]
14. Yaeger R, Chatila WK, Lipsyc MD, Hechtman JF, Cercek A, Sanchez-Vega F, et al. Clinical Sequencing Defines the Genomic Landscape of Metastatic Colorectal Cancer. *Cancer Cell* 2018;33(1):125–136 e3. [PubMed: 29316426]
15. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 2013;6(269):p11.

16. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol* 2016;17(1):218. [PubMed: 27765066]
17. Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med* 2015;21(11):1350–6. [PubMed: 26457759]
18. Lu C, Allis CD. SWI/SNF complex in cancer. *Nat Genet* 2017;49(2):178–179. [PubMed: 28138149]
19. Mathur R. ARID1A loss in cancer: Towards a mechanistic understanding. *Pharmacol Ther* 2018;190:15–23. [PubMed: 29730444]
20. Cristescu R, Mogg R, Ayers M, Albright A, Murphy E, Yearley J, et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* 2018;362(6411).
21. Zhu YP, Sheng LL, Wu J, Yang M, Cheng XF, Wu NN, et al. Loss of ARID1A expression is associated with poor prognosis in patients with gastric cancer. *Hum Pathol* 2018;78:28–35. [PubMed: 29689245]
22. Mathur R, Alver BH, San Roman AK, Wilson BG, Wang X, Agoston AT, et al. ARID1A loss impairs enhancer-mediated gene regulation and drives colon cancer in mice. *Nat Genet* 2017;49(2):296–302. [PubMed: 27941798]
23. Wang SC, Nassour I, Xiao S, Zhang S, Luo X, Lee J, et al. SWI/SNF component ARID1A restrains pancreatic neoplasia formation. *Gut* 2019;68(7):1259–1270. [PubMed: 30315093]
24. Gao Xuyuan, Xu Hang, James CM Ho Oscar SH Chan, Xu Feng, Wang Junwen, et al. EGFR-mutant lung adenocarcinomas mutation profiles reveal ARID1A might be a novel tyrosine kinase resistance pathway. *Cancer Research* 2018;78(13):Supplement, pp. LB-196.
25. Stintzing S, Tejpar S, Gibbs P, Thiebach L, Lenz HJ. Understanding the role of primary tumour localisation in colorectal cancer treatment and outcomes. *Eur J Cancer* 2017;84:69–80. [PubMed: 28787661]
26. Remy E, Rebouissou S, Chaouiya C, Zinovyev A, Radvanyi F, Calzone L. A Modeling Approach to Explain Mutually Exclusive and Co-Occurring Genetic Alterations in Bladder Tumorigenesis. *Cancer Res* 2015;75(19):4042–52. [PubMed: 26238783]
27. Shen J, Peng Y, Wei L, Zhang W, Yang L, Lan L, et al. ARID1A Deficiency Impairs the DNA Damage Checkpoint and Sensitizes Cells to PARP Inhibitors. *Cancer Discov* 2015;5(7):752–67. [PubMed: 26069190]
28. Grasso CS, Giannakis M, Wells DK, Hamada T, Mu XJ, Quist M, et al. Genetic Mechanisms of Immune Evasion in Colorectal Cancer. *Cancer Discov* 2018;8(6):730–749. [PubMed: 29510987]
29. Kim SY, Choi YY, An JY, Shin HB, Jo A, Choi H, et al. The benefit of microsatellite instability is attenuated by chemotherapy in stage II and stage III gastric cancer: Results from a large cohort with subgroup analyses. *Int J Cancer* 2015;137(4):819–25. [PubMed: 25614197]
30. Fukumoto T, Park PH, Wu S, Fatkhutdinov N, Karakashev S, Nacarelli T, et al. Repurposing Pan-HDAC Inhibitors for ARID1A-Mutated Ovarian Cancer. *Cell Rep* 2018;22(13):3393–3400. [PubMed: 29590609]
31. Bitler BG, Aird KM, Garipov A, Li H, Amatangelo M, Kossenkova AV, et al. Synthetic lethality by targeting EZH2 methyltransferase activity in ARID1A-mutated cancers. *Nat Med* 2015;21(3):231–8. [PubMed: 25686104]
32. Berns K, Caumanns JJ, Hijmans EM, Gennissen AMC, Severson TM, Evers B, et al. ARID1A mutation sensitizes most ovarian clear cell carcinomas to BET inhibitors. *Oncogene* 2018.

Highlights:

- The impact of *ARID1A* mutations was analyzed using 7,978 colorectal cancer samples.
- *ARID1A* mutation was linked to immune activation, right-sided tumor, and early stage.
- *ARID1A* mutation may be a biomarker for chemo/radio-therapy and targeted-therapy.

Discovery cohort

ARID1A mutation status and

- ✓ clinicopathological factors
- ✓ immune profile
- ✓ other genes mutation status

CARIS, CRC samples with NGS data on 592 panel from April, 2015 to January, 2018, $N = 6,272$

Appendiceal cancer excluded, $n = 297$

Neuroendocrine tumor/squamous cell carcinoma excluded, $n = 22$

Unclassified and indeterminate status for *ARID1A* mutation excluded, $n = 194$ and $n = 33$

CARIS, CRC (adenocarcinoma) samples with NGS data on 592 panel, $N = 5,726$

Validation cohorts

ARID1A mutation status and

- ✓ clinicopathological factors
- ✓ immune profile
- ✓ other genes mutation status

DFCI, CRC (adenocarcinoma) samples with NGS data, $N = 619$

MSK, CRC (adenocarcinoma) samples with NGS data, $N = 1,134$

TCGA, CRC (adenocarcinoma) samples, $N = 594$

Second samples from same patients excluded, $n = 35$

Samples without NGS data excluded, $n = 60$

MSK, CRC (adenocarcinoma) samples with NGS data, $N = 1,099$

TCGA, CRC (adenocarcinoma) samples with NGS data, $N = 534$

ARID1A mutation status and

- ✓ related pathways

TCGA, CRC (adenocarcinoma) samples with NGS data, $N = 534$

GSE59857, CRC cell lines with RNA-seq data, $N = 155$

Samples without RNA-seq data excluded, $n = 2$

CRC cell lines without *ARID1A* mutation data from CCLE excluded, $n = 91$

TCGA, CRC (adenocarcinoma) samples with NGS and RNA-seq data, $N = 532$

GSE59857, CRC cell lines with *ARID1A* mutation and RNA-seq data, $N = 64$

Figure 1.
CONSORT diagram.

Flow chart showing the inclusion/exclusion criteria in this study.

Abbreviations: CCLE, cancer cell line encyclopedia; CRC, colorectal cancer; NGS, next-generation sequencing; RNA-seq, RNA sequencing.

Figure 2.
The impact of *ARID1A* mutation on tumor characteristics in the discovery cohort.
A: The function of *ARID1A* as a SWI/SNF complex subunit for chromatin remodeling. B: Mutation location and type of *ARID1A*. C: Frequency of gene mutations in right-sided and left-sided CRC. D: The characteristics of immune profile and tumor location in all samples, comparing *POLE* mutant, MSI-H, *ARID1A* mutant and *ARID1A* wild-type subtypes. E: The characteristics of immune profile and tumor location in *ARID1A* mutant without MSI-H status. F: X-axis indicates co-occurring mutations (right, positive log ratio) and mutually

exclusive mutations (left, negative log ratio), and Y-axis indicates the significance. Log ratio is Log2 based ratio of samples with mutation in *ARID1A* mutation group/samples with mutation in *ARID1A* wild-type group.

Abbreviations: CRC, colorectal cancer; Mb, megabase; MSI, microsatellite instability; MSI-H, MSI-high; MSI-L, MSI-low; MSS, microsatellite stable; PD-L1, programmed death ligand 1; POLE, polymerase ϵ ; TMB, tumor mutational burden.

Figure 3.

The impact of *ARID1A* mutation on tumor microenvironment in the TCGA cohort.

A: The characteristics of tumor microenvironment in *ARID1A* mutant samples, comparing *POLE* mutant, MSI-H, and *ARID1A* wild-type subtypes. B: The characteristics of tumor microenvironment in *ARID1A* mutant samples without MSI-H status. C: CMS classifier in *ARID1A* mutant samples, comparing *ARID1A* wild-type samples. D: CMS classifier in *ARID1A* mutant samples without MSI-H status.

Abbreviations: MSI, microsatellite instability; MSI-H, MSI-high; MSI-L, MSI-low; MSS, microsatellite stable; PD-L1, programmed death ligand 1; *POLE*, polymerase ϵ .

Figure 4.

The impact of *ARID1A* mutation on tumor characteristics in the validation cohorts.

A: OncoPrint representation of gene mutations in the validation cohorts within each sample for the 50 top mutated genes and for the tumor location and MSI status (indicated in orange and red on the bottom annotation heatmap). Top and side bar plots indicate the number of mutations per sample and the number of each gene mutation across all the samples, respectively. B: X-axis indicates co-occurring mutations (right, positive log ratio) and mutually exclusive mutations (left, negative log ratio), and Y-axis indicates the significance.

Log ratio is Log2 based ratio of samples with mutation in *ARID1A* mutation group/samples with mutation in *ARID1A* wild-type group. C: Gene set enrichment analysis with hallmarks gene sets in the TCGA and CSE59857 cohort (*ARID1A* mutant vs. *ARID1A* wild-type samples).

Abbreviations: MSI, microsatellite instability; MSI-H, MSI-high; MSI-L, MSI-low; MSS, microsatellite stable.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.
Association between *ARID1A* mutation and clinicopathological factors across four cohorts.

Characteristics total N = 7,978	CARIS			DFCI			MSK			TCGA		
	<i>ARID1A</i> mutant n = 468 (8%)	<i>ARID1A</i> wild-type n = 5,258 (92%)	P value	<i>ARID1A</i> mutant n = 50 (8%)	<i>ARID1A</i> wild-type n = 569 (92%)	P value	<i>ARID1A</i> mutant n = 104 (10%)	<i>ARID1A</i> wild-type n = 995 (90%)	P value	<i>ARID1A</i> mutant n = 58 (11%)	<i>ARID1A</i> wild-type n = 476 (89%)	P value
n (%)	n (%)	n (%)	P value	n (%)	n (%)	P value	n (%)	n (%)	P value	n (%)	n (%)	P value
Age												
Median (range)	62 (16–95)	60 (16–95)	0.095	70 (52–85)	71 (42–90)	0.778	57 (24–87)	54 (13–93)	0.144	66 (33–90)	67 (31–90)	0.330
Sex												
Male	226 (48%)	2848 (54%)	0.015	14 (28%)	225 (40%)	0.101	61 (59%)	536 (54%)	0.350	29 (50%)	248 (52%)	0.739
Female	242 (52%)	2411 (46%)		36 (72%)	344 (60%)		43 (41%)	459 (46%)		29 (50%)	226 (48%)	
Tumor location												
Right side	157 (47%)	1215 (32%)	<0.001	32 (64%)	283 (50%)	0.053	51 (50%)	275 (28%)	<0.001	37 (64%)	197 (42%)	0.002
Left side	177 (53%)	2546 (68%)		18 (36%)	285 (50%)		52 (50%)	707 (72%)		21 (36%)	274 (58%)	
TNM Stage												
I	-	-	-	9 (19%)	143 (28%)	0.004	7 (7%)	33 (3%)	<0.001	11 (19%)	84 (18%)	0.046
II	-	-		27 (57%)	160 (31%)		25 (24%)	103 (10%)		30 (52%)	172 (37%)	
III	-	-		7 (15%)	152 (29%)		36 (35%)	231 (23%)		14 (24%)	138 (30%)	
IV	-	-		4 (9%)	61 (12%)		36 (35%)	628 (63%)		3 (5%)	71 (15%)	
Tumor sampling												
Primary	282 (60%)	2743 (52%)	<0.001	-	-	-	80 (77%)	506 (51%)	<0.001	-	-	-
Metastasis	186 (40%)	2515 (48%)		-	-		24 (23%)	489 (49%)		-	-	
MSI status												
MSI-L/MSS	249 (55%)	4919 (97%)	<0.001	15 (37%)	423 (87%)	<0.001	47 (45%)	959 (96%)	<0.001	29 (50%)	433 (91%)	<0.001
MSI-H	206 (45%)	146 (3%)		26 (63%)	65 (13%)		57 (55%)	36 (4%)		29 (50%)	41 (9%)	

P value was based on Student's t-test for continuous data and Fisher's exact test for categorical data. Unknown group was not included in this analysis. MSI, microsatellite instability; MSI-H, MSI-high; MSI-L, MSI-low; MSS, microsatellite stable.