


RESEARCH

Open Access



DNA methylation changes that precede onset of dysplasia in advanced sessile serrated adenomas

Cheng Liu^{1,2,3*} , Lochlan J. Fennell^{1,2}, Mark L. Bettington^{2,3}, Neal I. Walker^{2,3}, Joel Dwine³, Barbara A. Leggett^{1,2,4} and Vicki L. J. Whitehall^{1,2,5}

Abstract

Background: Sessile serrated adenomas (SSAs) are common polyps which give rise to 20–30% of colorectal cancer (CRC). SSAs display clinicopathologic features which present challenges in surveillance, including overrepresentation in young patients, proclivity for the proximal colon and rarity of histologic dysplasia (referred to then as SSAs with dysplasia, SSADs). Once dysplasia develops, there is rapid progression to CRC, even at a small size. There is therefore a clinical need to separate the “advanced” SSAs at high risk of progression to SSAD and cancer from ordinary SSAs. Since SSAs are known to accumulate methylation over time prior to the development of dysplasia, SSAD backgrounds (the remnant SSA present within an SSAD) likely harbour additional methylation events compared with ordinary SSAs. We therefore performed MethyLight and comprehensive methylation array (Illumina MethylationEPIC) on 40 SSAD backgrounds and 40 matched ordinary SSAs, and compared the methylation results with CRC methylation, CRC expression and immunohistochemical data.

Results: SSAD backgrounds demonstrated significant hypermethylation of CpG islands compared with ordinary SSAs, and the proportion of hypermethylated probes decreased progressively in the shore, shelf and open sea regions. Hypomethylation occurred in concert with hypermethylation, which showed a reverse pattern, increasing progressively away from the island regions. These methylation changes were also identified in *BRAF*-mutant hypermethylated CRCs. When compared with CRC expression data, *SV2B*, *MLH1/EP300*, *C16orf62*, *RCOR3*, *BAIAP3*, *OGDHL*, *HDHD3* and *ATP1B2* demonstrated both promoter hypermethylation and decreased expression. Although SSAD backgrounds were histologically indistinguishable from ordinary SSAs, *MLH1* methylation was detectable via MethyLight in 62.9% of SSAD backgrounds, and focal immunohistochemical *MLH1* loss was seen in 52.5% of SSAD backgrounds.

Conclusions: Significant hyper- and hypomethylation events occur during SSA progression well before the development of histologically identifiable changes. Methylation is a heterogeneous process within individual SSAs, as typified by *MLH1*, where both *MLH1* methylation and focal immunohistochemical *MLH1* loss can be seen in the absence of dysplasia. This heterogeneity is likely a generalised phenomenon and should be taken into account in future methylation-based studies and the development of clinical methylation panels.

Keywords: Sessile serrated adenoma, Dysplasia, Methylation, Colorectal cancer, Methylation array, *MLH1*

* Correspondence: john.liu@qimrberghofer.edu.au

¹The Conjoint Gastroenterology Laboratory, QIMR Berghofer Medical Research Institute, 300 Herston Road, Herston, Brisbane, QLD 4006, Australia

²Faculty of Medicine, University of Queensland, Brisbane, QLD, Australia

Full list of author information is available at the end of the article



Background

Colorectal cancer (CRC) is the third most common cancer worldwide, responsible for approximately 860,000 deaths in 2018 [1]. CRC arises from two precursor pathways. The conventional pathway, responsible for 70–80% of CRC, follows the adenoma-carcinoma sequence and is initiated by conventional adenomas [2]. These adenomas progress slowly and increase in size and dysplasia over time, with the risk of CRC development being directly related to these parameters [3]. This provides an extended period for colonoscopic detection.

The serrated pathway is responsible for the remaining 20–30% of cases [4] and is initiated by sessile serrated adenomas (SSAs). SSAs differ from conventional adenomas in several respects. They arise at any age and are overrepresented in young patients [5, 6]. Histologically, they vary little in appearance for the majority of their dwell time prior to the development of dysplasia. However, once dysplasia develops (referred to then as SSA with dysplasia, SSAD), they rapidly progress to CRC [7]. SSADs are thought to be responsible for many “interval” CRCs which arise within the colonoscopy surveillance interval [8].

These characteristics present special challenges in surveillance. The criteria used in conventional adenomas are suboptimal, as SSAs do not increase in size significantly with age [5], and the SSAs which progress to dysplasia do so at a small size (less than 10 mm) [7, 9]. SSADs are rare, vary in morphology, are readily mistaken for other entities and progress too rapidly to be useful as a risk marker [7, 10]. On the other hand, because SSAs comprise 20% of resected colorectal polyps [11] and the majority do not develop dysplasia, treating all lesions as potentially malignant will require following up an unacceptably large number of patients.

Ideally high-risk SSAs should be identified prior to the development of dysplasia, but this is not possible based on histologic appearance alone. From a molecular aspect, SSAs are initiated by *BRAF* mutation and accumulate methylation at CpG sites over time. This accumulation is not random and can be quantified by assessing the methylation status of preselected genes. When these genes are methylated, it is referred to as the CpG island methylator phenotype (CIMP). CIMP is strongly associated with *BRAF* mutation in CRCs [12] and is almost universal in SSADs [7], while its incidence in SSAs is much more variable and correlates with increased patient age [13]. Methylation in SSAs occurs over many years, eventually reaching a threshold where critical tumour-suppressor genes are silenced by promoter methylation [14]. *MLH1* is the best characterised of these tumour-suppressor genes, and loss of *MLH1* protein function results in microsatellite instability and further accumulation of

mutations [15]. Histologically, detectable dysplasia then develops, and the SSA becomes an SSAD.

From this model, the residual non-dysplastic SSA in an SSAD (hereby referred to as “SSAD background”) represents the most “advanced” SSA prior to development of dysplasia and should harbour similarly advanced molecular alterations, despite being histologically indistinguishable from an ordinary SSA. *BRAF* mutation and CIMP are not useful in this context as they are present in the majority of ordinary SSAs, while *MLH1* silencing occurs too late in progression and is essentially restricted to SSADs. Identification of significant methylation events which accompany CIMP, but precede *MLH1* silencing, may translate into clinical markers which allow for identification of high-risk SSAs. We therefore analysed a large series of residual non-dysplastic SSA in an SSAD (i.e. SSAD backgrounds) and ordinary SSAs using the Infinium MethylationEPIC platform, a methylation microarray which interrogates 866,836 CpG sites across the genome.

Results

Patient demographics

As the SSAD backgrounds and SSAs were matched, the demographics were essentially identical (Table 1). For SSAD backgrounds, the mean age was 75.1 years, 30 were female and 36 were proximal. For SSAs, the mean age was 75.0 years, 30 were female and 39 were proximal.

Methylation of CIMP genes

The methylation status of the five CIMP genes (*NEUROG1*, *SOCS1*, *CACNA1G*, *IGF2* and *RUNX3*) [16] was examined by both MethyLight and the MethylationEPIC array.

With MethyLight, all 35 SSAD backgrounds with sufficient DNA were CIMP-high. As part of the selection criteria, all 40 SSAs were also CIMP-high. Specifically, *NEUROG1*, *SOCS1*, *CACNA1G*, *IGF2* and *RUNX3* were hypermethylated in 35 of 35 (100.0%), 28 of 35 (80.0%), 35 of 35 (100.0%), 35 of 35 (100.0%) and 35 of 35 (100.0%) SSAD backgrounds, respectively; they were hypermethylated in 39 of 40 (97.5%), 20 of 40 (50.0%), 35 of 40 (87.5%), 37 of 40 (92.5%) and 35 of 40 (87.5%) SSAs, respectively. Only *SOCS1* methylation was significantly different between the two groups ($P < 0.01$). These results are summarised in Additional file 1: Supplementary Table 1.

With the MethylationEPIC array, for *NEUROG1*, 16 of 20 promoter-associated probes were hypermethylated in SSAD backgrounds compared with SSAs, and one of these differences was significant ($P < 0.05$, Additional file 2: Supplementary Figure 1A). For *SOCS1*, 13 of 19 probes were hypermethylated, but none of the differences were significant (Additional file 2: Supplementary

Table 1. Patient demographics

	SSAD background (N = 40)	SSA (N = 40)	P value
Age (mean ± SD)	75.1 ± 8.3	75.0 ± 7.9	N.S.
Gender			
Female	30	30	N.S.
Male	10	10	
Colonic site			
Proximal	36	39	N.S.*
Distal	1	1	
Unknown	3	0	

*Excluding SSAD backgrounds of unknown site. SSAD sessile serrated adenoma with dysplasia, SSA sessile serrated adenoma, N.S. not significant

Figure 1B). For *CACNA1G*, 19 of 23 probes were hypermethylated, but none of the differences were significant (Additional file 2: Supplementary Figure 1C). For *IGF2*, 9 of 20 probes were hypermethylated, but none of the differences were significant (Additional file 2: Supplementary Figure 1D). For *RUNX3*, 18 of 20 probes were hypermethylated, and five of these differences were significant ($P < 0.05$, Additional file 2: Supplementary Figure 1E).

Methylation and expression of *MLH1*

With MethyLight, *MLH1* was methylated in 22 of 35 (62.9%) SSAD backgrounds with sufficient DNA, compared with 0 of 40 (0.0%) SSAs (Additional file 1: Supplementary Table 1). Although this difference was statistically significant ($P < 0.01$), it was likely due to the selection criterion which required all SSAs to be *MLH1* unmethylated. With the MethylationEPIC array, 19 of 24 *MLH1* promoter-associated probes were hypermethylated in SSAD backgrounds compared with SSAs, and two of these differences were significant ($P < 0.05$, Additional file 2: Supplementary Figure 1F). Of note, all 11 island-associated probes were hypermethylated, which included the two significantly different probes.

Immunohistochemical *MLH1* loss in isolated crypts was identified in 21 of 40 (52.5%) SSAD backgrounds (example shown in Fig. 1). These crypts were focal (less than 5% of all crypts in a given SSAD background) and not associated with any appreciable H&E differences compared with adjacent crypts with retained *MLH1*. Of the subset of SSAD backgrounds with sufficient DNA for MethyLight, 19 of 35 (54.3%) showed focal *MLH1* loss. Fourteen of 19 (73.7%) SSAD backgrounds with *MLH1* loss were *MLH1* methylated, compared with 8 of 16 (50.0%) SSAD backgrounds with retained *MLH1*. This difference was not significant (Additional file 3: Supplementary Table 2). As per convention, this analysis used a percentage of methylated reference (PMR) of 10% as the minimal cut off for significant methylation [16]. However, because *MLH1* loss involved only a small number of crypts, whole-lesion *MLH1* methylation was unlikely to reach this threshold. We thus relaxed the

cut off and examined cases with any detectable *MLH1* methylation (i.e. PMR > 0%). In SSAD backgrounds with *MLH1* loss, 17 of 19 (89.5%) showed detectable *MLH1* methylation, compared with 10 of 16 (62.5%) SSAD backgrounds with retained *MLH1*. This difference was also not significant (Additional file 3: Supplementary Table 2).

Global methylation differences between SSAD backgrounds and SSAs

Overall, 740,033 of 866,836 (85.4%) probes on the MethylationEPIC array met the quality control and filtering threshold and were included in the analyses. Using an adjusted $P < 0.05$, 51,304 of these probes were significantly differentially methylated between SSAD backgrounds and SSAs. The majority of the probes showed hypomethylation, rather than hypermethylation, in SSAD backgrounds (36,744 of 51,304, 71.6%). To further investigate this finding, we divided the probes into island, shore, shelf and open sea groups. With this division, SSAD backgrounds were more hypermethylated in the islands (4274 of 7561, 56.5%). However, as one moved away from the islands, the proportion of hypermethylated probes in SSAD backgrounds decreased progressively from 4195 of 9504 (44.1%) shore probes to 676 of 4097 (16.5%) shelf probes and to 5415 of 30,142 (18.0%) open sea probes. These data are summarised in Fig. 2.

Promoter methylation differences between SSAD backgrounds and SSAs

Due to the known role of promoter methylation in SSA progression, we then limited analyses to promoter-associated probes, defined as within 1500 bp upstream of the transcription start site of a gene [17]. A total of 19,938 probes were significantly differentially methylated between SSAD backgrounds and SSAs. On dividing the probes into island, shore, shelf and open sea groups, SSAD backgrounds were more hypermethylated in the islands (3047 of 4248, 71.7%). However, as one moved away from the islands, the proportion of hypermethylated probes in SSAD backgrounds decreased progressively from 2809 of

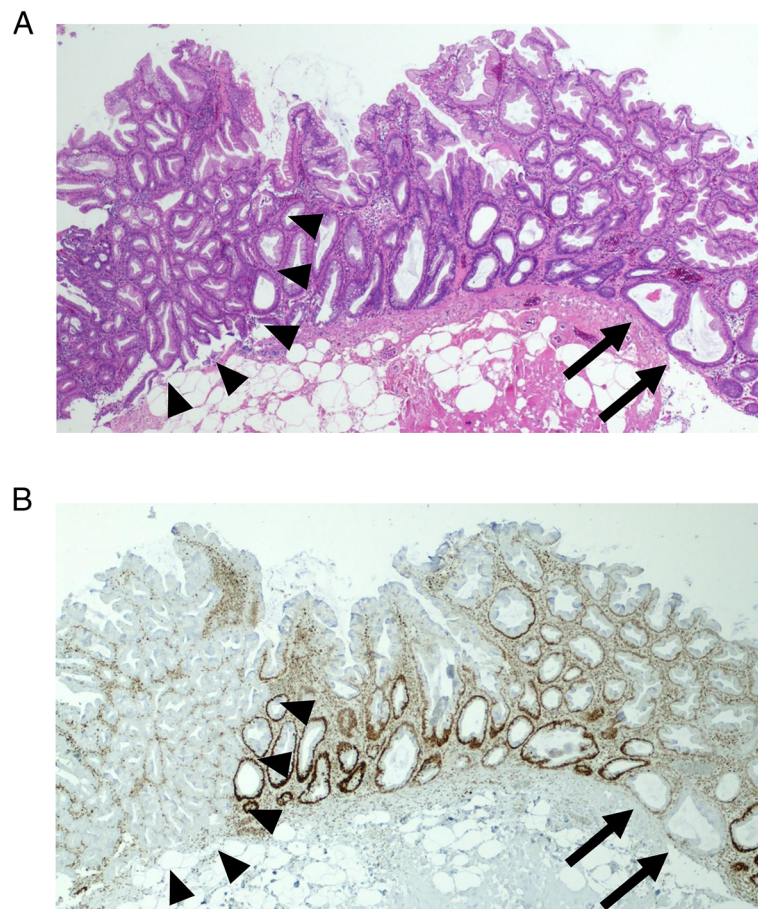


Fig. 1. Focal MLH1 loss in the SSAD backgrounds. **a** Haematoxylin and eosin-stained section of an SSAD, with the SSAD background on the right (crypts of interest indicated by arrows) and the dysplastic portion on the left (arrowheads). **b** MLH1 immunohistochemistry for the same SSAD. There is loss of nuclear MLH1 staining in a proportion of crypts in the SSAD background (arrows), while the dysplastic portion shows complete loss of staining (arrowheads). SSAD, sessile serrated adenoma with dysplasia; SSA, sessile serrated adenoma.

5384 (52.2%) shore probes to 180 of 961 (18.7%) shelf probes and to 1516 of 9345 (16.2%) open sea probes. These data are summarised in Fig. 2.

Comparison with methylation in *BRAF*-mutant CIMP-high microsatellite unstable CRC

Biologically significant methylation differences which occur from SSA to SSAD background should persist into the resultant CRC; that is, hypermethylated probes in the SSAD background should remain hypermethylated in the CRC, and similarly for hypomethylated probes. Because all of our SSADs showed loss of MLH1 staining on immunohistochemistry in the dysplastic component, they were expected to give rise to *BRAF*-mutant CIMP-high microsatellite unstable CRCs if allowed to progress. We had previously performed methylation array on 21 formalin-fixed, paraffin-embedded CRCs, using the Illumina HumanMethylation450 platform (Illumina, San Diego, USA) [18]. Six of these were *BRAF*-mutant

CIMP-high microsatellite unstable CRCs, and they were compared with the 40 SSAs in the current study.

Of the 51,304 probes which showed a significant difference between SSAD backgrounds and SSAs on the MethylationEPIC platform, 27,748 were also present in the HumanMethylation450 platform. They represented 9630 hypermethylated probes and 18,118 hypomethylated probes in the SSAD background versus SSA comparison, and comparison between CRC and SSA was limited to these probes. For the SSAD background hypermethylated probes, 7173 (74.5%) were also hypermethylated in CRC, 138 (1.4%) were hypomethylated in CRC and 2319 (24.1%) were not significantly different. For the SSAD background hypomethylated probes, 10,476 (57.8%) were also hypomethylated in CRC, 874 (4.8%) were hypermethylated in CRC and 6768 (37.4%) were not significantly different. Considering only probes which showed a significant difference in the CRC versus SSA comparison, 17,649 of 18,661 (94.6%) were concordant between SSAD background and CRC, and only 1012 of 18,661 (5.4%) were discordant.

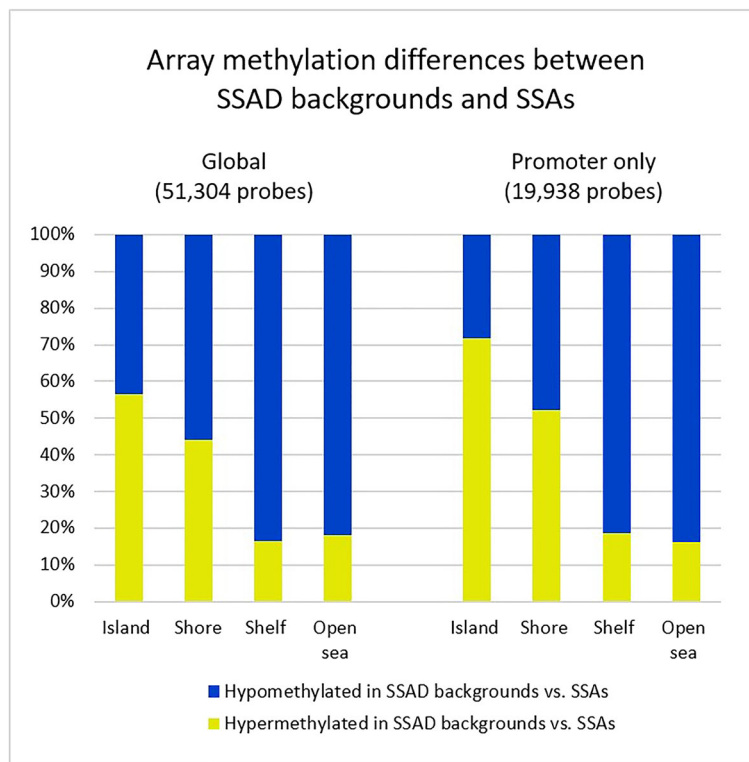


Fig. 2. Array methylation differences between SSAD backgrounds and SSAs. Globally SSAD backgrounds are more methylated than SSAs in island regions, but the proportion of hypermethylated probes decreases towards shore, shelf and open sea regions. The same trend is accentuated when only promoter probes are analysed. SSAD, sessile serrated adenoma with dysplasia; SSA, sessile serrated adenoma.

The great majority of methylation changes identified in SSAD backgrounds were therefore also present in *BRAF*-mutant CIMP-high microsatellite unstable CRC.

Identification of potential genes involved in SSA progression

To identify candidate tumour-suppressor genes silenced in the progression of SSA to SSAD background, we searched for hypermethylated islands within promoters of protein-coding genes. This identified 2936 probes, corresponding to 2290 unique protein-coding genes. The 20 most hypermethylated (highest logFC) and 20 most significantly different (lowest adjusted *P* value) probes and their corresponding genes are shown in Table 2. The complete list of all significant probes is given in Additional file 4: Supplementary Table 3.

Because promoter methylation does not always correlate with decreased gene expression [19, 20], we then compared our list of 2290 candidate genes with publicly available CRC expression data (Fennell et al. [21], ArrayExpress E-MTAB-7036). When 18 *BRAF*-mutant CIMP-high microsatellite unstable CRCs (i.e. the CRCs expected to arise from *MLH1*-deficient SSADs) were compared with 32 normal mucosa samples, 870 of the 2290 protein-coding genes were differentially expressed,

with 341 (39.2%) showing decreased expression. Of the top hypermethylated probes in Table 2, eight probes were associated with decreased gene expression in CRC, corresponding to *SV2B*, *MLH1/EPM2AIP1*, *C16orf62*, *RCOR3*, *BALAP3*, *OGDHL*, *HDHD3* and *ATP1B2*. The complete list of all differentially expressed genes is given in Additional file 5: Supplementary Table 4.

Discussion

Based on endoscopic and pathologic studies in the past decade, the natural history of SSAs is becoming increasingly clear. Once initiated by a *BRAF* mutation, SSAs accumulate methylation at a gradual rate over decades [13]. Despite the increase in methylation, there is little change in size and histologic appearance. After methylation reaches a certain threshold, tumour-suppressor genes such as *MLH1* are silenced by promoter methylation and histologic dysplasia, or an SSAD, develops [4]. SSADs progress rapidly to CRC which, based on microsatellite instability status, can have a superior prognosis or a poor prognosis [22].

These features present special challenges in colonoscopic surveillance, since criteria used to identify high-risk conventional adenomas are poorly applicable in SSAs. SSAs do not increase significantly in size over time [5], and dysplasia progresses too rapidly to be useful as a

Table 2. Hypermethylated promoter island probes showing largest differences between SSAD backgrounds and SSAs

Probe	Gene	Methylation logFC (SSAD background vs. SSA)*	Adjusted P value
Most hypermethylated probes			
cg21016956	<i>NOL4</i>	1.493751	0.000787
cg16916433	<i>ADAMTS19</i>	1.392486	0.002431
cg07733457	<i>ZNF594</i>	1.363048	0.002957
cg20388206	<i>NOL4</i>	1.290463	0.002242
cg08624472	<i>ARPC1B</i>	1.253215	0.040197
cg08326075	<i>SV2B</i>	1.251781	0.003033
cg16552945	<i>ARHGDI3</i>	1.233535	0.007083
cg14933485	<i>PGR</i>	1.200440	0.000873
cg06520273	<i>CCDC39</i>	1.198272	0.004478
cg07352001	<i>ZNF32</i>	1.197676	0.010383
cg24713878	<i>P2RX5</i>	1.195933	0.037569
cg27331401	<i>MLH1/EPM2AIP1</i>	1.194737	0.022763
cg15353810	<i>C16orf62</i>	1.190115	0.005143
cg17210604	<i>HIC1</i>	1.180987	0.031011
cg11224603	<i>MLH1/EPM2AIP1</i>	1.163986	0.004474
cg22308600	<i>NOL4</i>	1.163516	0.005631
cg22715021	<i>CLPSL2</i>	1.157144	0.035127
cg27586588	<i>MLH1/EPM2AIP1</i>	1.154898	0.048775
cg12609243	<i>CACNA2D3</i>	1.148465	0.002208
cg11555122	<i>ZNF594</i>	1.142647	0.005650
Most statistically significant probes			
cg16709874	<i>RCOR3</i>	1.017873	0.000382
cg21016956	<i>NOL4</i>	1.493751	0.000787
cg14933485	<i>PGR</i>	1.200440	0.000873
cg19809077	<i>GTF2H4/VARS2</i>	0.751390	0.000897
cg23546619	<i>BCAS3</i>	0.717448	0.001064
cg17360299	<i>MYCN</i>	0.783056	0.001130
cg06925115	<i>ATP2B2</i>	0.718108	0.001319
cg26691477	<i>EN1</i>	0.714066	0.001326
cg05116343	<i>BAIAP3</i>	0.902199	0.001433
cg24727182	<i>ETNPPL</i>	0.748564	0.001561
cg02357389	<i>CCDC59/METTL25</i>	0.958562	0.001634
cg09647147	<i>OGDHL</i>	0.959377	0.001654
cg07774938	<i>CCDC180</i>	0.880026	0.001682
cg11680300	<i>AUTS2</i>	1.135098	0.001693
cg02945056	<i>GCC1</i>	0.831327	0.001795
cg00142257	<i>LHX6</i>	0.789183	0.001795
cg04987474	<i>HDHD3</i>	0.561859	0.001948
cg21267231	<i>DUOX2</i>	0.557276	0.002008
cg09982069	<i>XPO4</i>	0.909124	0.002033
cg10043101	<i>ATP1B2</i>	0.940714	0.002133

*A methylation logFC value of > 0 indicates hypermethylation in SSAD background compared with SSA. SSAD sessile serrated adenoma with dysplasia, SSA sessile serrated adenoma

marker [7]. Methylation testing is a promising surrogate marker due to its long period of accumulation, and applicability is supported by the observation that virtually all SSADs are hypermethylated [7]. However, past age 60 more than 70% of ordinary SSAs are also hypermethylated [13], and the great majority of these do not develop dysplasia; the indiscriminate use of methylation panels in unselected SSAs is therefore impractical.

The non-dysplastic SSA component in an SSAD, or SSAD background as referred to in this study, represents the most “advanced” SSA possible prior to development of dysplasia. Although morphologically identical, SSAD backgrounds should differ in methylation from ordinary SSAs. To clarify these differences, we examined two homogeneous groups most likely to harbour biologically significant methylation differences. The SSAD backgrounds were all from *MLH1*-deficient SSADs. The SSAs were age-, gender- and site-matched to the SSAD backgrounds, with the additional requirements of also being CIMP-high and *MLH1* unmethylated via MethyLight. The CIMP-high requirement ensured these SSAs were similar to the SSAD backgrounds (which were all CIMP-high in samples with sufficient DNA for MethyLight), while the *MLH1* unmethylated requirement ensured the SSAs did not represent undersampled SSADs.

With these stringent clinical, histologic and methylation criteria, the MethylationEPIC array still identified a large number of significantly differentially methylated probes between SSAD backgrounds and SSAs (Fig. 2). Globally, although only 28.4% of probes were hypermethylated in SSAD backgrounds, they were concentrated within CpG islands, where 56.5% of probes were hypermethylated. The degree of methylation decreased as one moved away from the islands to the shore, shelf and sea regions. Due to the known importance of promoter hypermethylation-induced gene silencing in SSA progression, we then restricted analysis to the promoter regions. The trend was accentuated, where 71.7% of island-associated probes were hypermethylated, with the same decrease as one moved away from the islands. This concentration of hypermethylation in promoter-associated regions, and a tendency towards hypomethylation elsewhere, is also identified in studies comparing unselected CRC with normal mucosa [20, 23–25].

In contrast, there are few comparable studies of SSAs. Most previous methylation studies of SSAs have focused on a single gene or a panel of preselected genes [26–35]. Two array-based studies and one sequencing study have compared SSAs to normal mucosa. Dehghanizadeh et al. [36] used the HumanMethylation450 array to identify methylated genes in SSAs compared with normal mucosa, followed by RNA sequencing. Three genes were methylated and downregulated in SSAs (*BMP3*, *EPB41L3* and *CBS*). Using a customised methylation microarray, Inoue

et al. [37] identified 32 genes which were methylated in SSAs compared with normal mucosa, and performed immunohistochemistry for the six most promising candidates. *HDHD3*, which was also identified in our study (see Table 2), was one of the candidates. Parker et al. [38] compared SSAs with normal mucosa using the SeqCap Epi CpGiant sequencing platform, followed by RNA sequencing. They were able to construct a six-gene methylation panel which separated SSAs from normal mucosa.

The genes identified in the studies above are useful in understanding the earliest steps in SSA formation. However, the clinical challenge is in identifying markers which separate advanced SSAs from ordinary SSAs, as the former is capable of rapid progression to CRC. In this vein, Andrew et al. [39] used the HumanMethylation450 array to identify 15 CpG probes methylated in CIMP-high CRCs as well as SSAs, and validated the results in several public datasets. Although these probes were all located within islands they were not restricted to promoters, and interestingly none corresponded to *MLH1*. In our study, methylation of these probes as assessed by the MethylationEPIC array did not differ significantly between SSAs and SSAD backgrounds (Additional file 6: Supplementary Table 5). This suggests the candidate probes identified by Andrew et al. are methylated early in SSA development and are not useful in separating SSAs from SSAD backgrounds.

Notably, these four studies also identified significant hypomethylation with lesion progression, whether it was from normal mucosa to SSA [36–38], or from SSA to CIMP-high CRC [39]. In our study, we have shown it also occurs from SSA to SSAD background, at an even higher frequency than hypermethylation. This indicates hypomethylation and hypermethylation are changes that accumulate gradually at all steps of SSA progression to CRC. Although the functional consequences of hypomethylation could not be explored via the methods employed in our study, it is known hypomethylation does not represent a bystander process or the simple mirror image of hypermethylation [40]. Global hypomethylation can induce gross chromosomal abnormalities [41], while hypomethylation of repetitive sequences such as long interspersed nuclear elements-1 (LINE-1) facilitates their insertion into regulatory regions and tumour-suppressor genes [42]. In contrast, direct activation of proto-oncogenes by hypomethylation appears to be an uncommon event [43].

Together, the four studies described above utilised small numbers of lesions and several analysis methods to identify differentially methylated genes in the development of non-dysplastic SSAs. The obtained data showed minimal overlap, and no constant methylation event was identified across all studies. Although this may have been due to methodology differences, it is also possible

SSAs do not differentially methylate a set of predefined genes during its long dwell time prior to developing dysplasia. Notably, none of these studies identified *MLH1* as a candidate, consistent with the view that *MLH1* methylation represents an advanced alteration seen only in SSADs.

Since our methylation array data demonstrated only small fold changes (logFC range between -2.0 and +2.1 for all significantly different probes), we examined genes with known methylation statuses, to ensure data consistency across methods. Because all SSAD backgrounds and SSAs were CIMP-high via MethyLight, the five CIMP genes were expected to be similarly methylated on the MethylationEPIC array across all lesions. Indeed, when promoter-associated probes corresponding to the five CIMP genes were considered, the majority showed no significant difference between SSAD backgrounds and SSAs (Additional file 2: Supplementary Figure 1). The six probes which did show a significant difference were all hypermethylated in SSAD backgrounds, consistent with the observation they represented more advanced SSAs. As a further means of validating our data, we also compared our methylation data with *BRAF*-mutant CIMP-high microsatellite unstable CRCs. These CRCs were the expected end result if the *MLH1*-deficient SSADs in our study were allowed to progress. Of the assessable probes, 94.6% were concordant between SSAD background and CRC, suggesting the methylation changes which occur from SSA to SSAD background persist into CRC.

To identify potential tumour-suppressor genes silenced in the progression of SSA to SSAD background, we limited analyses to hypermethylated, island-located probes within the promoters of protein-coding genes. We then compared the results to expression data derived from *BRAF*-mutant CIMP-high microsatellite unstable CRCs, to further narrow down the candidates to genes which were downregulated by promoter hypermethylation.

Of the genes corresponding to the top probes, *ADAMTS19*, *PGR*, *MLH1*, *HIC1*, *ENI*, *OGDHL* and *LHX6* methylation have been described in CRCs, and *HDHD3* methylation has been described in SSAs. *ADAMTS19* methylation is associated with mucinous differentiation, *BRAF* mutation, microsatellite instability and metastasis [44]. *PGR* codes for the progesterone receptor. Because hormone replacement therapy reduces CRC risk [45], progesterone receptor likely functions as a tumour suppressor. This is supported by cell line [46] and murine [47] data. However, hormone replacement therapy does not significantly affect *PGR* methylation [48], and the protective effect of hormone replacement therapy may be due to oestrogen receptor rather than progesterone receptor function. *HIC1* has been investigated mainly as part of a CIMP panel [49], but its methylation is not associated with clinicopathologic features [50]. *ENI* methylation is

also not associated with clinicopathologic features [51]. *OGDHL* is methylated and downregulated in a proportion of unselected CRC [52]. *LHX6* methylation is described in CRC [53], but its functional significance has not been investigated. *HDHD3* is methylated in SSAs, where decreased expression is confirmed by immunohistochemical staining [37].

MLH1 warrants separate discussion as its role in SSA progression is well established. Previous studies have demonstrated *MLH1* methylation is detectable in otherwise unremarkable SSAs, ranging from 0.0 to 89.5% [13, 28, 31, 32, 54–65]. The cause for this large variation is uncertain but has been attributed to patient ethnicity [58], patient age [13, 60], lesion site [31, 55, 58] and primer choice [58, 66]. In SSADs, there is a very high concordance rate between *MLH1* methylation and immunohistochemical *MLH1* loss in the dysplastic portion [7]. These prior studies have led to the view that *MLH1* methylation occurs late in SSA progression, where once a critical threshold is reached, *MLH1* protein function is lost and there is concurrent development of dysplasia.

Our current data further clarify the role of *MLH1* in the progression of SSA to SSAD. Because all our SSAD backgrounds were associated with *MLH1*-deficient dysplasia, it was not surprising to identify *MLH1* methylation within this background. SSAD backgrounds were much more likely to be *MLH1* methylated via MethyLight compared with SSAs (62.9% vs. 0.0%), but this difference was enhanced by the requirement that all SSAs were *MLH1* unmethylated. Our results were different to a previous study of macrodissected SSADs, which found *MLH1* methylation did not differ significantly between ordinary SSAs and SSAD backgrounds [67]. This was likely due to their inclusion of both *MLH1*-deficient and *MLH1*-retained SSADs.

In our study, 52.5% of SSAD backgrounds showed *MLH1* loss in isolated crypts (Fig. 1). However, it should be noted that because these foci represented less than 5% of the crypts in each histologic section, undersampling was possible and the true incidence would likely be higher if additional sections were taken. Compatible with our result, a recent study demonstrated focal *MLH1* loss in 65.7% of SSAD backgrounds from *MLH1*-deficient SSADs [9]; furthermore, although we did not perform *MLH1* immunohistochemistry in our SSAs, focal *MLH1* loss was identified in 7.0% of randomly selected SSAs in the same study [9]. Focal *MLH1* loss in SSAs and SSAD backgrounds had also been illustrated in other studies [32, 68–71], mainly as a pathologic curiosity.

The *MLH1* methylation model can be improved as follows. *MLH1* methylation occurs late in SSA progression, concentrated within promoter-associated CpG islands. This methylation is heterogeneous across the lesion, and different crypts accumulate methylation at varying rates,

with some eventually losing protein function, reflected by *MLH1* loss on immunohistochemistry. This explains the discordant methylation and immunohistochemistry results, because not all sections from a given SSA background will include the hypermethylated crypts. It is unknown how long these *MLH1*-deficient crypts can persist before development of dysplasia. However, as they are rarely seen in the absence of overt dysplasia, the rate of progression is likely rapid.

The main limitation of our study was lack of expression data from SSAD backgrounds and SSAs, which were required to identify the genes silenced by hypermethylation. Because SSAs are CRC precursors, the entire lesion should be submitted for pathologic analysis to exclude a microscopic malignant component, and obtaining fresh tissue for RNA or protein extraction is difficult. As a surrogate, we utilised expression data derived from *BRAF*-mutant CIMP-high microsatellite unstable CRC. However, this involved comparing methylation data derived from formalin-fixed paraffin-embedded SSAD backgrounds and SSAs, and expression data derived from fresh-frozen CRCs and normal mucosa. Differences in tissue type, DNA/RNA quality, platform choice and data analysis were confounding factors in this part of the study.

Conclusion

We have shown significant methylation changes occur during SSA progression, well before CRC and even before development of histologic dysplasia. In addition to expected hypermethylation of promoter-associated CpG islands, there is also hypomethylation of shore, shelf and sea regions. These methylation changes likely persist into the resulting CRC. Furthermore, methylation is a heterogeneous process even within individual SSAs, where isolated *MLH1*-deficient crypts can occur in a *MLH1*-retained background, in the absence of histologically detectable dysplasia. This heterogeneity is likely a generalised phenomenon involving multiple genes in addition to *MLH1* and should be taken into account in future methylation-based studies. Furthermore, this result has implications in the development of clinical methylation panels, as these panels generate methylation data averaged across entire lesions, where small foci of significant methylation may be diluted by the background.

Methods

A schematic of our experimental design is shown in Fig. 3.

Case selection

A total of 40 SSADs were sourced from Envoi Specialist Pathologists, a gastrointestinal pathology practice in Brisbane, Australia. For each SSAD, haematoxylin and eosin (H&E)-stained slides were retrieved and reviewed by two gastrointestinal pathologists (C. L. and M. L. B.) to

confirm the SSAD fulfilled the World Health Organization diagnostic criteria. Furthermore, the SSADs were required to (1) have sufficient SSAD background component for macrodissection (at least 3 mm) and (2) demonstrate loss of *MLH1* staining on immunohistochemistry in their dysplastic component. The second criterion was introduced because *MLH1*-deficient SSADs were more common than *MLH1*-proficient SSADs, and it restricted analysis to a homogeneous group. Patient demographic data and lesion site (proximal or distal colon) were obtained from the pathology request form. Of these 40 SSADs, 33 had been reported in two previous studies [7, 10].

DNA extraction

For each SSAD, the corresponding paraffin block was retrieved. Ten 10- μ m unstained sections were cut onto uncharged slides, followed immediately by a single 4 μ m H&E section. This additional H&E section was used to confirm the SSAD background remained in the intervening unstained sections. The unstained sections were deparaffinised, and the SSAD background was macrodissected using a sterile needle and the original H&E section as a guide; to avoid contamination by the dysplastic focus, a thin rim of SSAD background immediately adjacent to the dysplastic focus was not included in the macrodissection. Lesional mucosa content was at least 90% for all cases. DNA was extracted via Chelex as previously described [72].

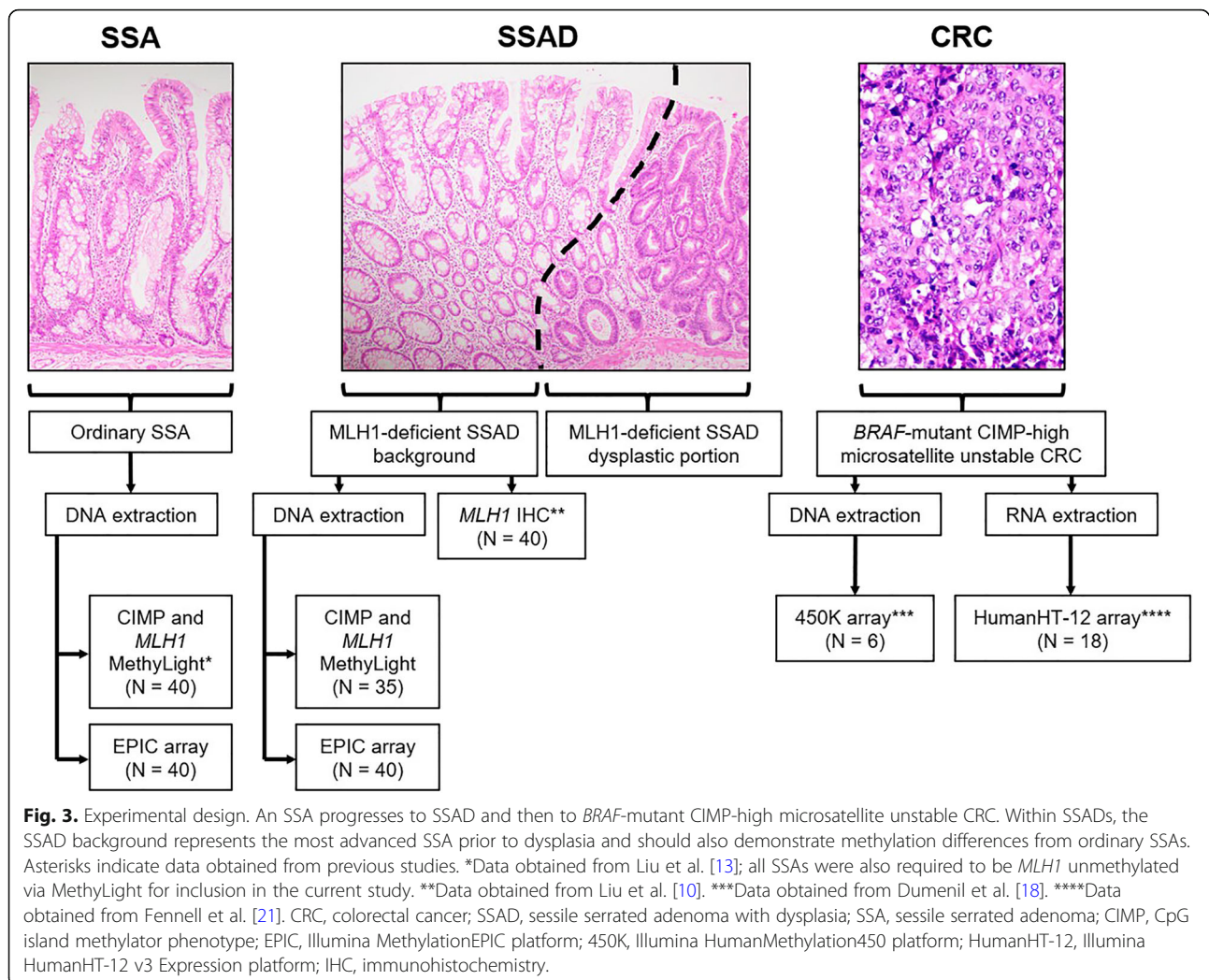
CIMP and *MLH1* methylation

After prioritisation of SSAD background DNA for methylation microarray, there was sufficient sample remaining to assess CIMP and *MLH1* methylation status by MethyLight in 35 of 40 SSAD backgrounds. This was performed as previously described [73]. CIMP used the five markers of Weisenberger et al. [16] (*NEUROG1*, *SOCS1*, *CACNA1G*, *IGF2* and *RUNX3*), where CIMP-high was defined as three or more markers methylated.

Control selection

For each SSAD, an age-, gender-, site- and CIMP status-matched ordinary SSA was selected as control. The SSAs were also required to be *MLH1* unmethylated. All 40 selected SSAs had been included in a previous study [13], which used identical methods of DNA extraction and MethyLight analysis. Because SSADs are overwhelmingly proximal colonic and CIMP-high [7], lesions without a stated site of origin (three SSADs) were matched with a proximal SSA, and lesions with insufficient DNA for CIMP (five SSADs) were matched with a CIMP-high SSA.

With this method of control selection, the study utilised two very closely related, but distinct, groups. The SSAD backgrounds represented the most advanced SSAs prior to development of *MLH1*-deficient dysplasia, while



the matched SSAs represented the highly methylated lesions typically seen in an older population [13]. The *MLH1*-unmethylated requirement in the SSAs enhanced the difference between the groups, which was intended to exclude undersampled SSADs not present in the planes of the section. Any differences between these two groups would therefore represent late events in SSA progression, occurring just prior to *MLH1* loss.

Methylation microarray

For all SSAD backgrounds and SSAs, 1.0 µg of extracted DNA was submitted to Macrogen, Inc. (Seoul, South Korea) for methylation microarray. DNA quality control used the Infinium FFPE QC Kit (Illumina, San Diego, USA), DNA restoration used the Infinium HD FFPE DNA Restore Kit (Illumina, San Diego, USA), bisulfite conversion used the EZ-96 DNA Methylation Kit (Zymo Research, Irvine, USA) and methylation microarray used

the Infinium MethylationEPIC BeadChip Kit (Illumina, San Diego, USA).

Immunohistochemistry

MLH1 immunohistochemistry had been previously performed on all SSADs as part of a previous study [10].

Methylation array data analysis

MethylationEPIC array data were imported into the R environment using *minfi* (v1.28.3) [74]. Probes were annotated using the mapping by Zhou et al. [75], and probes that mapped to single nucleotide polymorphisms or sex chromosomes were not included in downstream analyses. Data were normalised using the functional normalisation method [76]. We masked β values for individual positions with detection $P \geq 0.05$ and discarded any probes with detection $P \geq 0.05$ in more than 50% of samples. Differential methylation analyses

were performed on the M-transformed β values using the Limma package (v3.38.3) [77]. *P* values were adjusted to account for errors in multiple testing using the Benjamini-Hochberg false discovery rate method [78].

Comparison with CRC data

We compared our global SSAD background and SSA methylation data with CRC methylation data from a previous study [18], which included six *BRAF*-mutant CIMP-high microsatellite unstable CRCs and also utilised DNA extracted from formalin-fixed, paraffin-embedded tissue. Because an earlier version of the methylation array was used (HumanMethylation450 instead of MethylationEPIC), we limited the comparison to probes present on both platforms. Raw data were normalised using functional normalisation and quality control filtering as above.

Because promoter methylation does not always lead to transcriptional silencing [19, 20], we investigated whether significantly methylated genes in SSAD backgrounds correlated with reduced expression in CRCs. To this end, we examined differential gene expression between 18 *BRAF*-mutant CIMP-high microsatellite unstable CRCs and 32 normal colonic mucosa samples from a publicly accessible gene expression dataset (Fennell et al. [21], ArrayExpress E-MTAB-7036). Raw data were imported into the R environment, quantile normalised and background corrected. Where multiple probes mapped to a single gene, we computed the median to assess gene-level expression. Differential gene expression was performed using the Limma package (v3.38.3), and *P* values were corrected for false discovery using the Benjamini-Hochberg false discovery rate method as above.

Statistical analysis

Continuous variables were analysed using an unpaired *t*-test. Categorical variables were analysed using a chi-squared test. A *P* value of <0.05 was regarded as significant.

Availability of methylation array data

The dataset supporting the conclusions of this article is available in the ArrayExpress repository (E-MTAB-7854).

Additional files

Additional file 1: Supplementary **Table 1**. Detailed CIMP and *MLH1* MethyLight results. (DOCX 13 kb)

Additional file 2: Supplementary **Figure 1**. Methylation of CIMP gene and *MLH1* promoters as assessed by the MethylationEPIC array. Each data point represents a promoter-associated probe, plotted as the β value difference between SSAD backgrounds and SSAs ($\beta_{SSAD\ backgrounds} - \beta_{SSAs}$). A positive value indicates hypermethylation in the SSAD background, and a negative value indicates hypomethylation in the SSAD background.

Significantly different probes are highlighted in red. TSS, transcription start site. (DOCX 40 kb)

Additional file 3: Supplementary **Table 2**. Correlation between MethyLight *MLH1* methylation and immunohistochemical MLH1 loss in SSAD backgrounds. (DOCX 12 kb)

Additional file 4: Supplementary **Table 3**. Genes with at least one hypermethylated promoter island probe. (XLSX 670 kb)

Additional file 5: Supplementary **Table 4**. Differentially expressed genes in *BRAF*-mutant microsatellite unstable CRCs, as compared with our list of hypermethylated gene promoter-associated probes. (XLSX 42 kb)

Additional file 6: Supplementary **Table 5**. The probes identified by Andrew et al. [39], as assessed on the MethylationEPIC array in our study. (XLSX 11 kb)

Acknowledgements

Not applicable.

Authors' contributions

CL contributed to the study concept and design, acquisition of the data, analysis and interpretation of the data and drafting of the manuscript. LJF contributed to the analysis and interpretation of the data, bioinformatic analysis and critical revision of the manuscript. MLB contributed to the study concept and design, analysis and interpretation of the data, critical revision of the manuscript and study supervision. NIW contributed to the study concept and design, analysis and interpretation of the data, critical revision of the manuscript and study supervision. JD contributed to the acquisition of the data. BAL contributed to the study concept and design, analysis and interpretation of the data, critical revision of the manuscript and study supervision. VLW contributed to the study concept and design, analysis and interpretation of the data, critical revision of the manuscript and study supervision. All authors read and approved the final manuscript.

Funding

CL and LJF gratefully acknowledge funding support from the Australian Government (Australian Postgraduate Award), Pathology Queensland and Cancer Council Queensland (APP1160923). LJF is supported in part by a PhD scholarship from Australian Rotary Health.

Availability of data and materials

The datasets generated and analysed during the current study are available in the ArrayExpress repository (E-MTAB-7854).

Ethics approval and consent to participate

This study was approved by the ethics committee of the Queensland Institute of Medical Research Berghofer (P1298).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹The Conjoint Gastroenterology Laboratory, QIMR Berghofer Medical Research Institute, 300 Herston Road, Herston, Brisbane, QLD 4006, Australia. ²Faculty of Medicine, University of Queensland, Brisbane, QLD, Australia. ³Envoi Specialist Pathologists, Brisbane, QLD, Australia. ⁴The Royal Brisbane and Women's Hospital, Brisbane, QLD, Australia. ⁵Department of Chemical Pathology, Pathology Queensland, Brisbane, QLD, Australia.

Received: 5 April 2019 Accepted: 4 June 2019

Published online: 14 June 2019

References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394–424.

2. Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, et al. Genetic alterations during colorectal-tumor development. *N Engl J Med*. 1988;319:525–32.
3. Martinez ME, Baron JA, Lieberman DA, Schatzkin A, Lanza E, Winawer SJ, et al. A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy. *Gastroenterology*. 2009;136:832–41.
4. O'Brien MJ, Zhao Q, Yang S. Colorectal serrated pathway cancers and precursors. *Histopathology*. 2015;66:49–65.
5. Bettington M, Brown I, Rosty C, Walker N, Liu C, Croese J, et al. Sessile serrated adenomas in young patients may have limited risk of malignant progression. *J Clin Gastroenterol*. 2019;53:e113–6.
6. Kim HY, Kim SM, Seo JH, Park EH, Kim N, Lee DH. Age-specific prevalence of serrated lesions and their subtypes by screening colonoscopy: a retrospective study. *BMC Gastroenterol*. 2014;14:82.
7. Bettington M, Walker N, Rosty C, Brown I, Clouston A, McKeone D, et al. Clinicopathological and molecular features of sessile serrated adenomas with dysplasia or carcinoma. *Gut*. 2017;66:97–106.
8. Lee YM, Huh KC. Clinical and biological features of interval colorectal cancer. *Clin Endosc*. 2017;50:254–60.
9. Yozu M, Kem M, Cenaj O, Mino-Kenudson M, Odze RD, Misdraji J. Loss of expression of MLH1 in non-dysplastic crypts is a harbinger of neoplastic progression in sessile serrated adenomas/polyps. *Histopathology*. 2019. <https://doi.org/10.1111/his.13874> Epub ahead of print.
10. Liu C, Walker NI, Leggett BA, Whitehall VL, Bettington ML, Rosty C. Sessile serrated adenomas with dysplasia: morphological patterns and correlations with MLH1 immunohistochemistry. *Mod Pathol*. 2017;30:1728–38.
11. Bettington M, Walker N, Rahman T, Vandeleur A, Whitehall V, Leggett B, et al. High prevalence of sessile serrated adenomas in contemporary outpatient colonoscopy practice. *Intern Med J*. 2017;47:318–23.
12. Advani SM, Advani P, DeSantis SM, Brown D, VonVille HM, Lam M, et al. Clinical, pathological, and molecular characteristics of CpG island methylator phenotype in colorectal cancer: a systematic review and meta-analysis. *Transl Oncol*. 2018;11:1188–201.
13. Liu C, Bettington ML, Walker NI, Dwine J, Hartel GF, Leggett BA, et al. CpG island methylation in sessile serrated adenomas increases with age, indicating lower risk of malignancy in young patients. *Gastroenterology*. 2018;155:1362–5 e2.
14. Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. *Gastroenterology*. 2010;138:2088–100.
15. Kane MF, Loda M, Gaida GM, Lipman J, Mishra R, Goldman H, et al. Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. *Cancer Res*. 1997;57:808–11.
16. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet*. 2006;38:787–93.
17. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. *Genomics*. 2011;98:288–95.
18. Dumenil TD, Wockner LF, Bettington M, McKeone DM, Klein K, Bowdler LM, et al. Genome-wide DNA methylation analysis of formalin-fixed paraffin embedded colorectal cancer tissue. *Genes Chromosomes Cancer*. 2014;53:537–48.
19. Spainhour JC, Lim HS, Yi SV, Qiu P. Correlation patterns between DNA methylation and gene expression in The Cancer Genome Atlas. *Cancer Inform*. 2019;18:1176935119828776.
20. Klett H, Balavarca Y, Toth R, Gigic B, Habermann N, Scherer D, et al. Robust prediction of gene regulation in colorectal cancer tissues from DNA methylation profiles. *Epigenetics*. 2018;13:386–97.
21. Fennell L, Dumenil T, Wockner L, Hartel G, Nones K, Bond C, et al. Integrative genome-scale DNA methylation analysis of a large and unselected cohort reveals five distinct subtypes of colorectal adenocarcinomas. *Cell Mol Gastroenterol Hepatol*. 2019. <https://doi.org/10.1016/j.jcmgh.2019.04.002> Epub ahead of print.
22. Phipps AI, Limburg PJ, Baron JA, Burnett-Hartman AN, Weisenberger DJ, Laird PW, et al. Association between molecular subtypes of colorectal cancer and patient survival. *Gastroenterology*. 2015;148:77–87 e2.
23. Naumov VA, Generozov EV, Zaharjevskaya NB, Matushkina DS, Larin AK, Chernyshov SV, et al. Genome-scale analysis of DNA methylation in colorectal cancer using Infinium HumanMethylation450 BeadChips. *Epigenetics*. 2013;8:921–34.
24. Timp W, Bravo HC, McDonald OG, Goggins M, Umbricht C, Zeiger M, et al. Large hypomethylated blocks as a universal defining epigenetic alteration in human solid tumors. *Genome Med*. 2014;6:61.
25. Luo Y, Wong CJ, Kaz AM, Dzieciatkowski S, Carter KT, Morris SM, et al. Differences in DNA methylation signatures reveal multiple pathways of progression from adenoma to colorectal cancer. *Gastroenterology*. 2014;147:418–29 e8.
26. Caruso M, Moore J, Goodall GJ, Thomas M, Phillis S, Tyskin A, et al. Over-expression of cathepsin E and trefoil factor 1 in sessile serrated adenomas of the colorectum identified by gene expression analysis. *Virchows Arch*. 2009;454:291–302.
27. Caruso M, Fung KY, Moore J, Brierley GV, Cosgrove LJ, Thomas M, et al. Claudin-1 expression is elevated in colorectal cancer precursor lesions harboring the BRAF V600E mutation. *Transl Oncol*. 2014;7:456–63.
28. Dhir M, Yachida S, Van Neste L, Glockner SC, Jeschke J, Pappou EP, et al. Sessile serrated adenomas and classical adenomas: an epigenetic perspective on premalignant neoplastic lesions of the gastrointestinal tract. *Int J Cancer*. 2011;129:1889–98.
29. Gaiser T, Meinhardt S, Hirsch D, Killian JK, Gaedcke J, Jo P, et al. Molecular patterns in the evolution of serrated lesion of the colorectum. *Int J Cancer*. 2013;132:1800–10.
30. Patai AV, Bartak BK, Peterfia B, Micsik T, Horvath R, Sumanszki C, et al. Comprehensive DNA methylation and mutation analyses reveal a methylation signature in colorectal sessile serrated adenomas. *Pathol Oncol Res*. 2017;23:589–94.
31. Renaud F, Mariette C, Vincent A, Wacrenier A, Maunoury V, Leclerc J, et al. The serrated neoplasia pathway of colorectal tumors: identification of MUC5AC hypomethylation as an early marker of polyps with malignant potential. *Int J Cancer*. 2016;138:1472–81.
32. Sakai E, Fukuyo M, Ohata K, Matsusaka K, Doi N, Mano Y, et al. Genetic and epigenetic aberrations occurring in colorectal tumors associated with serrated pathway. *Int J Cancer*. 2016;138:1634–44.
33. Subramaniam MM, Chan JY, Soong R, Ito K, Yeoh KG, Wong R, et al. RUNX3 inactivation in colorectal polyps arising through different pathways of colonic carcinogenesis. *Am J Gastroenterol*. 2009;104:426–36.
34. Sugai T, Eizuka M, Fujita Y, Kawasaki K, Yamamoto E, Ishida K, et al. Molecular profiling based on KRAS/BRAF mutation, methylation, and microsatellite statuses in serrated lesions. *Dig Dis Sci*. 2018;63:2626–38.
35. Beggs AD, Jones A, Shepherd N, Arnaout A, Finlayson C, Abulafi AM, et al. Loss of expression and promoter methylation of SLIT2 are associated with sessile serrated adenoma formation. *PLoS Genet*. 2013;9:e1003488.
36. Dehghanizadeh S, Khoddami V, Mosbrugger TL, Hammoud SS, Edes K, Berry TS, et al. Active BRAF-V600E is the key player in generation of a sessile serrated polyp-specific DNA methylation profile. *PLoS One*. 2018;13:e0192499.
37. Inoue A, Okamoto K, Fujino Y, Nakagawa T, Muguruma N, Sannomiya K, et al. B-RAF mutation and accumulated gene methylation in aberrant crypt foci (ACF), sessile serrated adenoma/polyp (SSA/P) and cancer in SSA/P. *Br J Cancer*. 2015;112:403–12.
38. Parker HR, Orjuela S, Martinho Oliveira A, Cereatti F, Sauter M, Heinrich H, et al. The proto CpG island methylator phenotype of sessile serrated adenomas/polyps. *Epigenetics*. 2018;13:1088–105.
39. Andrew AS, Baron JA, Butterly LF, Suriawinata AA, Tsongalis GJ, Robinson CM, et al. Hyper-methylated loci persisting from sessile serrated polyps to serrated cancers. *Int J Mol Sci*. 2017;18:E535.
40. Ehrlich M. DNA hypomethylation in cancer cells. *Epigenomics*. 2009;1:239–59.
41. Eden A, Gaudet F, Waghmare A, Jaenisch R. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science*. 2003;300:455.
42. Kerachian MA, Kerachian M. Long interspersed nucleotide element-1 (LINE-1) methylation in colorectal cancer. *Clin Chim Acta*. 2019;488:209–14.
43. Hoffmann MJ, Schulz WA. Causes and consequences of DNA hypomethylation in human cancer. *Biochem Cell Biol*. 2005;83:296–321.
44. Alonso S, González B, Ruiz-Larroya T, Durán Domínguez M, Kato T, Matsunaga A, et al. Epigenetic inactivation of the extracellular matrix metalloproteinase ADAMTS19 gene and the metastatic spread in colorectal cancer. *Clinical Epigenetics*. 2015;7:124.
45. Lin KJ, Cheung WY, Lai JY, Giovannucci EL. The effect of estrogen vs. combined estrogen-progestogen therapy on the risk of colorectal cancer. *Int J Cancer*. 2012;130:419–30.
46. Tanaka Y, Kato K, Mibu R, Uchida S, Asanoma K, Hashimoto K, et al. Medroxyprogesterone acetate inhibits proliferation of colon cancer cell

- lines by modulating cell cycle-related protein expression. *Menopause*. 2008;15:442–53.
47. Meijer BJ, Wielenga MCB, Hoyer PB, Amos-Landgraf JM, Hakvoort TBM, Muncan V, et al. Colorectal tumor prevention by the progestin medroxyprogesterone acetate is critically dependent on postmenopausal status. *Oncotarget*. 2018;9:30561–7.
 48. Wu AH, Siegmund KD, Long TI, Cozen W, Wan P, Tseng CC, et al. Hormone therapy, DNA methylation and colon cancer. *Carcinogenesis*. 2010;31:1060–7.
 49. Noshu K, Irahara N, Shima K, Kure S, Kirkner GJ, Schernhammer ES, et al. Comprehensive biostatistical analysis of CpG island methylator phenotype in colorectal cancer using a large population-based sample. *PLoS One*. 2008;3:e3698.
 50. Bagci B, Sari M, Karadayi K, Turan M, Ozdemir O, Bagci G. KRAS, BRAF oncogene mutations and tissue specific promoter hypermethylation of tumor suppressor SFRP2, DAPK1, MGMT, HIC1 and p16 genes in colorectal cancer patients. *Cancer Biomark*. 2016;17:133–43.
 51. Mayor R, Casadome L, Azuara D, Moreno V, Clark SJ, Capella G, et al. Long-range epigenetic silencing at 2q14.2 affects most human colorectal cancers and may have application as a non-invasive biomarker of disease. *Br J Cancer*. 2009;100:1534–9.
 52. Fedorova MS, Kudryavtseva AV, Lakunina VA, Snezhkina AV, Volchenko NN, Slavnova EN, et al. Downregulation of OGDHL expression is associated with promoter hypermethylation in colorectal cancer. *Molecular Biology*. 2015;49:608–17.
 53. Estecio MR, Youssef EM, Rahal P, Fukuyama EE, Gois-Filho JF, Maniglia JV, et al. LHX6 is a sensitive methylation marker in head and neck carcinomas. *Oncogene*. 2006;25:5018–26.
 54. Sambuudash O, Kim HM, Jo H, Kim HS, Lee KJ, Park HJ, et al. Molecular characteristics of colorectal serrated polyps and hyperplastic polyps: a STROBE compliant article. *Medicine (Baltimore)*. 2016;95:e5592.
 55. Sandmeier D, Benhattar J, Martin P, Bouzourene H. Serrated polyps of the large intestine: a molecular study comparing sessile serrated adenomas and hyperplastic polyps. *Histopathology*. 2009;55:206–13.
 56. Maeda T, Suzuki K, Togashi K, Nokubi M, Saito M, Tsujinaka S, et al. Sessile serrated adenoma shares similar genetic and epigenetic features with microsatellite unstable colon cancer in a location-dependent manner. *Exp Ther Med*. 2011;2:695–700.
 57. Burnett-Hartman AN, Newcomb PA, Potter JD, Passarelli MN, Phipps AI, Wurscher MA, et al. Genomic aberrations occurring in subsets of serrated colorectal lesions but not conventional adenomas. *Cancer Res*. 2013;73:2863–72.
 58. Kim KM, Lee EJ, Ha S, Kang SY, Jang KT, Park CK, et al. Molecular features of colorectal hyperplastic polyps and sessile serrated adenoma/polyps from Korea. *Am J Surg Pathol*. 2011;35:1274–86.
 59. Murakami T, Mitomi H, Saito T, Takahashi M, Sakamoto N, Fukui N, et al. Distinct WNT/beta-catenin signaling activation in the serrated neoplasia pathway and the adenoma-carcinoma sequence of the colorectum. *Mod Pathol*. 2015;28:146–58.
 60. Noshu K, Igarashi H, Ito M, Mitsuhashi K, Kurihara H, Kanno S, et al. Clinicopathological and molecular characteristics of serrated lesions in Japanese elderly patients. *Digestion*. 2015;91:57–63.
 61. O'Brien MJ, Yang S, Clebanoff JL, Mulcahy E, Farraye FA, Amoroso M, et al. Hyperplastic (serrated) polyps of the colorectum: relationship of CpG island methylator phenotype and K-ras mutation to location and histologic subtype. *Am J Surg Pathol*. 2004;28:423–34.
 62. O'Brien MJ, Yang S, Mack C, Xu H, Huang CS, Mulcahy E, et al. Comparison of microsatellite instability, CpG island methylation phenotype, BRAF and KRAS status in serrated polyps and traditional adenomas indicates separate pathways to distinct colorectal carcinoma end points. *Am J Surg Pathol*. 2006;30:1491–501.
 63. Rau TT, Agaimy A, Gehoff A, Geppert C, Jung K, Knobloch K, et al. Defined morphological criteria allow reliable diagnosis of colorectal serrated polyps and predict polyp genetics. *Virchows Arch*. 2014;464:663–72.
 64. Kim YH, Kakar S, Cun L, Deng G, Kim YS. Distinct CpG island methylation profiles and BRAF mutation status in serrated and adenomatous colorectal polyps. *Int J Cancer*. 2008;123:2587–93.
 65. Muto Y, Maeda T, Suzuki K, Kato T, Watanabe F, Kamiyama H, et al. DNA methylation alterations of AXIN2 in serrated adenomas and colon carcinomas with microsatellite instability. *BMC Cancer*. 2014;14:466.
 66. Kaji E, Uraoka T, Kato J, Hiraoka S, Suzuki H, Akita M, et al. Externalization of saw-tooth architecture in small serrated polyps implies the presence of methylation of IGF1R. *Dig Dis Sci*. 2012;57:1261–70.
 67. Lee EJ, Chun SM, Kim MJ, Jang SJ, Kim do S, Lee DH, et al. Reappraisal of hMLH1 promoter methylation and protein expression status in the serrated neoplasia pathway. *Histopathology*. 2016;69:198–210.
 68. Hawkins NJ, Ward RL. Sporadic colorectal cancers with microsatellite instability and their possible origin in hyperplastic polyps and serrated adenomas. *J Natl Cancer Inst*. 2001;93:1307–13.
 69. Jass JR, Young J, Leggett BA. Hyperplastic polyps and DNA microsatellite unstable cancers of the colorectum. *Histopathology*. 2000;37:295–301.
 70. Oh K, Redston M, Odze RD. Support for hMLH1 and MGMT silencing as a mechanism of tumorigenesis in the hyperplastic-adenoma-carcinoma (serrated) carcinogenic pathway in the colon. *Hum Pathol*. 2005;36:101–11.
 71. Chino A, Nagayama S, Ishikawa H, Morishige K, Kishihara T, Arai M, et al. Cancer emerging from the recurrence of sessile serrated adenoma/polyp resected endoscopically 5 years ago. *Jpn J Clin Oncol*. 2016;46:89–95.
 72. Bettington ML, Walker NI, Rosty C, Brown IS, Clouston AD, McKeone DM, et al. A clinicopathological and molecular analysis of 200 traditional serrated adenomas. *Mod Pathol*. 2015;28:414–27.
 73. Eads CA, Danenberg KD, Kawakami K, Saltz LB, Blake C, Shibata D, et al. MethylLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Res*. 2000;28:E32.
 74. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014;30:1363–9.
 75. Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Res*. 2017;45:e22.
 76. Fortin JP, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol*. 2014;15:503.
 77. Smyth GK. limma: Linear Models for Microarray Data. In: Gentleman R, Carey VJ, Huber W, Irizarry RA, Dudoit S, editors. *Bioinformatics and computational biology solutions using R and Bioconductor*. New York, NY: Springer New York; 2005. p. 397–420.
 78. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1995;57:289–300.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

