Reviewer's report

Title: Methylation of Leukocyte DNA and Ovarian Cancer: Relationships with Disease Status and Outcome

Version: 1  
Date: 25 February 2014

Reviewer: Kristina Warton

Reviewer's report:

A number of studies have reported changes in the DNA methylation of peripheral blood leukocytes in cancer patients compared to controls. These are of great clinical interest, as they could potentially provide biomarkers for the non-invasive diagnosis of cancer. The study under review aims to answer an interesting question – which of the differences in peripheral blood DNA methylation are due to actual DNA methylation changes within leukocytes, rather than to fluctuations in the distribution of leukocyte populations. As such, it would shed light on the mechanistic response of leukocytes to cancer.

The strengths of the study are the large sample numbers (for what is a rare cancer), a well described patient cohort, and careful accounting for confounding variables. The manuscript is well written.

Major compulsory revisions:

1) The authors performed methylation analysis on peripheral blood DNA from three batches of cases and controls, using either the HumanMethylation27 BeadChip with 27,578 probes, or the HumanMethylation450 BeadChip, with 485,577 probes. Only probes in common between the two types of BeadChip were considered. As changes in peripheral blood methylation have been shown to parallel changes in leukocyte subtype population distributions, the authors then removed probes which are known to vary between leukocyte subtypes before undertaking the analysis of differential methylation. As the appropriate selection of subtype specific probes for removal is critical to the validity of the study, the authors need to provide clearer information about how this was done.

For example, paragraph 4 under Quality Control and Normalization in Materials and Methods states “…excluding 9,341 CpG probes on the Illumina Infinium HumanMethylation27 shown to associate with cell type distribution at q-value < 0.05…” referenced to Koestler et al. 2012, Houseman et al. 2012, and Chen et al. 2011. It’s not clear which of the referenced papers the number of 9,341 probes is derived from, though it’s obviously not Chen et al 2011. Koestler et al. 2012 do refer to 10,370 probes which were significantly differentially methylated between leukocyte subtypes, and this would seem the likeliest source of the 9,341 probes used, but this needs to be more clearly worded to avoid guess work. Perhaps even a list of the probes as a supplementary table would be appropriate.
2) (This comment assumes that the 2012 Koestler et al. paper is the source of the leukocyte subpopulation specific probes.) In the manuscript under review, 6% of the HumanMethylation27 BeadChip array probes did not pass quality control. This number seems quite low, reflecting good quality data. Assuming that the Koestler et al., 2012 data were similar, around 1600 of the probes from that study would not have been included in the list which differentiates leukocyte subpopulations. Since around 40% (10370/26486) of the good quality probes were differentially methylated among the leukocyte subtypes, one can assume that a similar proportion, or around 600, of the rejected poor quality probes would have been differentially methylated among leukocyte subtypes. Perhaps the Fridley et al. study is simply identifying additional sites which differentiate leukocyte subpopulations using probes which were not included in the Koestler et al. 2012 analysis due to not passing quality control? The description of probe selection by Fridley et al. should be presented with enough detail to be convincing that the above scenario is not the case, and to allow readers to evaluate it for other potential problems not identified by the current reviewer. One useful control would be to check that the 30 CpG’s identified as significant by Fridley et al. are actually present in the probe sets which passed QC and were used in the Koestler et al. study to differentiate leukocyte subpopulations.

3) The manuscript under review does not acknowledge that Teschendorff et al. clearly identified variation in the composition of peripheral blood leukocytes as the most likely source of the observed differences in methylation between controls and patient with ovarian cancer.

For example, the abstract states “In epithelial ovarian cancer (EOC), one report found substantial DNAm differences between cases and controls; however, results were later shown to be confounded by differences in white blood cell type distributions.”

The second paragraph of the introduction states “Previous work identified peripheral blood methylation signatures that predicted ovarian cancer case-control status using methylation measurements at more than 27,000 CpGs in 113 cases and 148 controls.[14 - Teschendorff et al. 2009] However, in re-analysis of this data, Koestler et al. and Houseman et al. found blood-based methylation measurements to be dependent on distribution of white blood cell (leukocyte) types, where the methylation level of CpG probes was highly associated with distribution of cell type and distribution of cell types was also associated with case-control status.[15 – Koestler et al., 2012, 16 – Houseman et al., 2012].

In a similar vein, the 8th paragraph of the discussion states “Prior work in a smaller set of cases and controls showed that blood-based DNAm associated with case-control status;[14 - Teschendorff et al. 2009] however, many of these associations were later found to be due to differential distributions of white blood cells types.[16 - Houseman et al., 2012].

Teschendorff et al. 2009 clearly identified differential distributions of white blood
cell types as being the likely source of the variation they observed. On p.4 of their manuscript they state “To understand these functional associations we hypothesized that some of these may reflect variations in blood cell type composition, as this is known to vary with both age and tumor presence” and go on to describe an enrichment in genes known to be upregulated in granulocytes, an effect also reported by Koestler et al., 2012.

In their abstract Teschendorff et al. 2009 say “…by comparing the pattern of methylation with gene expression data from major blood cell types, we here demonstrate that age and cancer elicit common changes in the composition of peripheral blood, with a myeloid skewing that increases with age and which is further aggravated in the presence of ovarian cancer”.

For these reasons I believe that wording that implies later papers showing an association between methylation changes and changes in leukocyte composition were correcting an earlier misinterpretation of data by Teschendorff et al. 2009 should be removed.

Minor Essential Revisions

A redundant comma appears in Table 2, for Probe ID cg02254461, under Location of nearest gene.

Discretionary revisions

Paragraph 1 of the Disease Status and Methylation section states that of the 30 sites identified as significantly differentially methylated all but one had lower methylation in cases compared to controls. It would be good if the name of the one hypermethylated site/nearest gene were provided in the text, rather than only in the table.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests.