Aberrant sperm DNA methylation predicts male fertility status and embryo quality

Kenneth I. Aston, Ph.D.,[a] Philip J. Uren, Ph.D.,[b] Timothy G. Jenkins, Ph.D.,[a] Alan Horsager, Ph.D.,[c] Bradley R. Cairns, Ph.D.,[d,e] Andrew D. Smith, Ph.D.,[b] and Douglas T. Carrell, Ph.D.[a,f]

[a] Department of Surgery, University of Utah Andrology and IVF Laboratories, University of Utah School of Medicine, Salt Lake City, Utah; [b] Molecular and Computational Biology, University of Southern California, Los Angeles, California; [c] Episona, Inc, Glendale, California; [d] Department of Oncological Sciences, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, Utah; [e] Howard Hughes Medical Institute, Chevy Chase, Maryland; and [f] Department of Obstetrics and Gynecology and Department of Human Genetics, University of Utah School of Medicine, Salt Lake City, Utah

Objective: To evaluate whether male fertility status and/or embryo quality during in vitro fertilization (IVF) therapy can be predicted based on genomewide sperm deoxyribonucleic acid (DNA) methylation patterns.
Design: Retrospective cohort study.
Setting: University-based fertility center.
Patient(s): Participants were 127 men undergoing IVF treatment (where any major female factor cause of infertility had been ruled out), and 54 normozoospermic, fertile men. The IVF patients were stratified into 2 groups: patients who had generally good embryogenesis and a positive pregnancy (n = 55), and patients with generally poor embryogenesis (n = 72; 42 positive and 30 negative pregnancies) after IVF.
Intervention(s): Genomewide sperm DNA methylation analysis was performed to measure methylation at >485,000 sites across the genome.
Main Outcome Measure(s): A comparison was made of DNA methylation patterns of IVF patients vs. normozoospermic, fertile men.
Result(s): Predictive models proved to be highly accurate in classifying male fertility status (fertile or infertile), with 82% sensitivity, and 99% positive predictive value. Hierarchic clustering identified clusters enriched for IVF patient samples and for poor-quality–embryo samples. Models built to identify samples within these groups, from neat samples, achieved positive predictive value ≥94% while identifying >one fifth of all IVF patient and poor-quality–embryo samples in each case. Using density gradient prepared samples, the same approach recovered 46% of poor-quality–embryo samples with no false positives.
Conclusion(s): Sperm DNA methylation patterns differ significantly and consistently for infertile vs. fertile, normozoospermic men. In addition, DNA methylation patterns may be predictive of embryo quality during IVF. (Fertil Steril® 2015;104:1388–97. ©2015 by American Society for Reproductive Medicine.)
Key Words: Sperm DNA, DNA methylation, IVF outcome, embryo, genomewide, microarray, male infertility

Discuss: You can discuss this article with its authors and with other ASRM members at http://fertstertforum.com/astonk-sperm-dna-methylation-infertility/

Use your smartphone to scan this QR code and connect to the discussion forum for this article now.*

* Download a free QR code scanner by searching for "QR scanner" in your smartphone's app store or app marketplace.

Received June 1, 2015; revised July 29, 2015; accepted August 18, 2015; published online September 8, 2015.
K.I.A. has a patent in preparation entitled Methods of Identifying Male Fertility Status and Embryo Quality. P.J.U. has a patent in preparation entitled Methods of Identifying Male Fertility Status and Embryo Quality. T.G.J. has a patent in preparation entitled Methods of Identifying Male Fertility Status and Embryo Quality. A.H. is a shareholder for Episona, Inc and is licensing this technology with Episona, Inc. B.R.C. has a patent in preparation entitled Methods of Identifying Male Fertility Status and Embryo Quality. A.D.S. has a patent in preparation entitled Methods of Identifying Male Fertility Status and Embryo Quality; and is a shareholder for Episona, Inc. D.T.C. is a shareholder for Episona, Inc.
Reprint requests: Douglas T. Carrell, Ph.D., Division of Urology-Andrology/IVF Laboratories, University of Utah School of Medicine, Andrology, 675 S Arapeen Dr, Ste #205, Salt Lake City, Utah 84108 (E-mail: douglas.carrell@hsc.utah.edu).

Fertility and Sterility® Vol. 104, No. 6, December 2015 0015-0282/$36.00

http://dx.doi.org/10.1016/j.fertnstert.2015.08.019

The mainstay of male infertility diagnosis is the standard semen analysis. With the exception of modification of criteria for morphologic grading, semen analysis has changed very little over the past several decades. Numerous studies have evaluated the prognostic value of the various semen parameters evaluated by the standard analysis (1–3). Except for severely diminished sperm count or motility, the predictive value of semen analysis for

fertility is modest at best. A milestone study of the predictive value of semen analysis concluded that although it is useful for classifying men as being either subfertile, of indeterminate fertility, or fertile, it is very ineffective for diagnosing infertility, owing to the fact that semen parameters for many infertile men fall within normal ranges (4).

The main parameters evaluated in semen analysis, namely sperm count, motility, viability, and morphology, are somewhat subjective; consequently, they are subject to technical error. Although continual training and assessment, quality control measures, and proficiency testing all minimize technical error, multiple studies have demonstrated that coefficients of variation (CVs) between labs and technicians commonly fall in the 20%–30% range, with higher CVs reported in some studies (5–7). In addition to the technical variability inherent in the testing, semen parameters for the same individual vary significantly among collections, with CVs of approximately 30% between any 2 collections from the same man, according to a recent study of >5,000 men (8). Given this inherent variability, the World Health Organization recommends that ≥2 semen analyses be performed before clinical decisions are made (9).

Lastly, the predictive value of the various semen parameters has been demonstrated to be severely limited. Two large and comprehensive studies have been performed to characterize semen parameters in healthy, fertile men; fertile and subfertile ranges have been defined for each parameter assessed (4, 10). Nevertheless, assessment using standard semen analysis is broadly accepted to fall far short of the goal of predicting fertility potential.

Adjunct tests have been developed over the years, such as sperm deoxyribonucleic acid (DNA) damage assessment, capacitation, and acrosome reaction tests, egg and zona penetration assays, antisperm antibody testing, aneuploidy screening, motile sperm organelle morphology examination, and hyaluronic acid binding (11–18). Although these tests can be helpful in characterizing fertility potential in select patients, the predictive values of the assays are generally accepted as being suboptimal (19). The need for additional diagnostic tools for evaluation of male infertility is widely acknowledged (2).

With advancing molecular diagnostic tools, the identification of novel genetic and epigenetic markers of male infertility is becoming a realistic option. A few genetic markers for male infertility have been identified, such as Y-chromosome microdeletions (20), Klinefelter syndrome (21), and DPY19L and SPATA16 mutations (22–25), among others. These genetic features are associated with extreme male infertility phenotypes, including severe oligozoospermia, nonobstructive azoospermia, and complete globozoospermia, accounting for a very small percentage of infertile men.

Although genetics refers to the DNA sequence itself, epigenetics refers to modifiable but generally stable modifications to the DNA or chromatin packaging. Direct DNA modifications consist of methylation alteration to the 5-carbon position of cytosine bases, usually in the context of cytosine-guanine dinucleotides (CpGs). Chromatin packaging modifications include covalent modifications to histones, and in sperm, the distribution of histones and protamines in the genome. Epigenomic information can be reliably replicated through cell divisions, and although the epigenome is largely reprogrammed in early gametogenesis and embryogenesis, some portion of the epigenome in parental gametes is transmitted to offspring. The amount of epigenomic information that is heritable in humans is largely unknown, as are the functions and consequences of epigenomic inheritance.

Sperm epigenetics is an emerging area of study, driven largely by early observations that the way DNA is packaged within a sperm head may affect the capacity of sperm to fertilize an egg and/or the developmental capacity of the zygote (26–29). Sperm DNA has long been recognized as being packaged differently, compared with other cell types (namely by protamines). Protamine packaging was generally assumed to serve utilitarian purposes, including compaction of the sperm nucleus to facilitate efficient motility, and protection of sperm DNA from the harsh environment of the female reproductive tract (30). However, more recent characterization of the epigenetic landscape of the sperm nucleus, specifically the localization of, and modification to, histones that remain associated with sperm DNA after replacement of most histones by protamines, indicates that sperm epigenetic architecture reflects the development of mature sperm from spermatogonial stem cells and likely contributes to early embryonic development (31, 32).

Numerous recent studies have reported abnormal sperm DNA methylation patterns associated with infertility, particularly in oligozoospermic patients. However, most of these studies evaluated just 1 or a few genes, or included very small cohorts using genomewide approaches (33–48). The aim of the current study was to evaluate genomewide sperm DNA methylation patterns in a large cohort of sperm donors and in vitro fertilization (IVF) patients, to determine whether DNA methylation patterns are predictive of fertility status or IVF prognosis.

The current study was motivated by 3 primary factors. (1) The current diagnostic standard for male infertility (i.e., semen analysis) provides minimal clinically actionable information. (2) Numerous exploratory studies have established that altered sperm DNA methylation in single or multiple genes is associated with male infertility (49, 50). (3) Sperm DNA methylation patterns are relatively stable within an individual (51). The goal of the current study was to determine whether sperm DNA methylation patterns can be used to predict male fertility status and IVF success.

MATERIALS AND METHODS

Sample Selection

Semen samples used for the current study were obtained from the University of Utah tissue bank, after informed consent had been provided according to IRB-approved protocols. Individuals were asked to adhere to general semen collection instructions, which included 2–5 days of abstinence immediately preceding collection. Collected samples were mixed in a 1:1 ratio with Test Yolk Buffer (Irvine Scientific) and stored in liquid nitrogen until they were used in this study.

Controls

Control samples (n = 54) were collected from normozoospermic men with proven fertility. Most control samples were composed of whole ejaculate (n = 36), whereas 12 were prepared by density-gradient centrifugation prior to cryopreservation. For 6 samples, the preparation method was not recorded.

Patient Samples

Samples were selected from 292 IVF patients based on embryo quality and pregnancy outcome (Supplemental Table 1, available online). Couples with moderate-to-severe female factor infertility, including advanced maternal age and severe endometriosis or polycystic ovarian syndrome, were excluded from the study. A total of 55 patients were selected because they had high embryo quality overall, and a confirmed chemical pregnancy. Embryos were scored based on previously reported criteria relating to blastomere number and fragmentation in early embryos, and trophectoderm and inner cell-mass quality in blastocysts (52).

Fetal heartbeat was detected in 49 patients (89.1%), undetectable in 4 patients (7.3%), and not recorded in 2 patients (3.6%). Seventy-two patients were selected who displayed generally poor embryogenesis, including increased rates of early or late developmental arrest, or reduced embryo quality due to blastomere fragmentation. Of these, 42 achieved a pregnancy; 30 did not. Of the 42 pregnant patients, fetal heartbeat was detected in 33 (78.6%), undetectable in 8 (19%), and not recorded in 1 patient (2.4%). Live-birth outcome data were not available for all patients, but for patients for whom data was recorded, 75% in the "good embryo" group delivered successfully, compared with only 31.4% in the "poor embryo" group.

Gradient-Prepared Patient Samples

A separate experiment was performed on a subset of the samples from IVF patients. For samples that contained $>5 \times 10^6$ progressively motile sperm (n = 44), DNA methylation was assessed on the whole ejaculate; separately, a portion of the sample was purified using a 45%/90% discontinuous isolate gradient, and the 90% fraction was subjected to DNA methylation analysis as well.

Patient Details

Supplemental Table 2 (available online) presents the frequency of male factor infertility (defined as being below the World Health Organization threshold in ≥ 1 of the semen parameters), female factor infertility, the presence of both male and female factors, and the designation of idiopathic infertility among both partners. Shown in addition is the nature of the various mild female factors within each group of IVF patients. The frequencies of all factors are statistically similar among groups ($P>.05$; χ^2 analysis).

Sample Preparation

For DNA isolation, sperm samples were thawed simultaneously and were subjected to a column-based DNA extraction protocol with sperm-specific modification to the DNeasy kit (Qiagen). Prior to DNA extraction, somatic cell lysis was performed by incubation in somatic cell lysis buffer (0.1% sodium dodecyl sulfate, 0.5% Triton X-100 in diethylpyrocarbonate H_2O) for 20 minutes on ice, to eliminate white blood cell contamination. After somatic cell lysis, the sperm were pelleted, and a visual inspection of each sample was performed to ensure the absence of all potentially contaminating cells before proceeding.

For bisulfite conversion and array processing, extracted sperm DNA was bisulfite converted with the EZ-96 DNA Methylation-Gold kit (Zymo Research), according to manufacturer recommendations specifically for use with array platforms. The converted DNA was delivered to the University of Utah Genomics Core Facility and hybridized to Infinium HumanMethylation450 BeadChip microarrays (Illumina) and analyzed according to manufacturer protocols.

Statistical Analysis

After the hybridization protocol, arrays were scanned. The minfi software package (Bioconductor; (53)) was used to generate β-values (so called because they are expected to follow a β-distribution, each is a value between zero and 1, for each probed CpG, representing the proportion of DNA molecules that are methylated at the given locus). The same package was used to apply subset-quantile within array normalization (54) to the extracted β-values. Statistical comparisons were made as follows: (1) all patients vs. controls; and (2) IVF patients with good vs. poor embryogenesis.

Hierarchic clustering was applied using the Euclidean distance between methylation profiles. For construction of discriminative models to differentiate IVF patient samples from fertile donor samples, and good from poor-quality–embryo samples, we evaluated 2 types of features to describe each sample: the Euclidean distance between the sample and all other samples in the training set, and a subset of the individual CpG methylation values for each sample. In the latter case, to select the subset, we first identified the 500 most-discriminative loci using Wilcoxon's rank sum test (i.e., the 500 loci with the lowest P values).

All models were constructed (including feature-selection steps) and evaluated using 10-fold cross-validation. Briefly, the samples are split into 10 stratified folds (10 disjoint sets in which the proportion of classes in each set approximates, as closely as possible, the proportion in the full dataset). Ten rounds of testing are conducted; for each round, 1 of the 10 folds is held out of the complete dataset, as testing data, and the remaining 9 folds are used to train the model. After model training is complete, the "held-out" testing data are classified, and the numbers of true vs. false positives and negatives are recorded. This process is repeated for all 10 folds of the data, such that every sample is tested (classified) exactly once, and every sample is unknown to the model that classifies it.

For gene ontology analysis, we ranked genes by the number of differentially methylated CpGs ($P<.05$, Wilcoxon's rank sum test). For grouping genes into sets that were either differentially methylated or not differentially methylated,

we first combined individual *P* values for the CpGs within the promoter region, which we define as ±5 kb of the transcription start site, of each RefSeq transcript (55), using the method of Stouffer et al. (56). We considered any gene to be differentially methylated if ≥1 of its transcripts showed differential methylation within this ±5-kb window around its transcription start site. Unless otherwise stated, all *P* values are corrected for multiple hypothesis testing using the method of Benjamini and Hochberg (57).

Global Classifiers

To capture global properties of DNA methylomes, we constructed models based on the Euclidean distance between samples. We used a simple rule-based classifier (which we call a "decision stump"), given as follows: "If distance to training instance X is greater than Y, classify the sample as class A; else classify it as class B," where A and B depend on the comparison (for example, good vs. poor embryo quality). We define the *cost* of any such rule as the weighted sum of the number of false positives and false negatives resulting from use of the rule to classify the training data. In the simplest case, each false positive or false negative is given a weight of 1, and the cost is simply the number of samples that are misclassified by the rule on the training data. The values of X and Y are learned by exhaustive enumeration of all possible training instances and discriminative thresholds on the training data; the one with the lowest cost is selected. We modulated the tradeoff between sensitivity and specificity in this learning process by adjusting the relative cost of false positives to false negatives, which we report in the results as the "false-positive cost" of a classification scheme. Schemes that have a high false-positive cost favor results with fewer false positives and hence high specificity. We considered such models to be more clinically valuable than those with high sensitivity.

Site-Specific Classifiers

These classifiers were constructed to explore sample classification based on the DNA methylation status (β-values) of a subset of individual CpGs. Here, as before, we used the decision-stump classifier, but in addition, we evaluated a selection of more-complex models and learning algorithms: a support vector machine, a nearest-neighbor classifier, a decision tree, and a naïve Bayes technique. For these models, we used the WEKA implementations (WEKA data mining software), run with default parameters, unless otherwise specified. Further, we augmented several of these with adaptive boosting (AdaBoost; an R software package), and bootstrap aggregating (bagging)—methods designed to avoid overfitting the model to the training data. As with the global classifiers, we favored high-specificity models. For models in which the false-positive cost could not be modulated during training, we weighted the training instances to achieve the same result.

RESULTS

Methylation in Purified vs. Unpurified Samples

Hierarchic clustering of samples showed that, although purified and unpurified samples were different, purified and unpurified methylomes that were derived from the same sample always clustered with one another; intersample variations were greater than those among purification methods (Supplemental Fig. 1, available online).

Is Aberrant DNA Methylation Predictive of Fertility Status and Embryo Quality?

We first performed hierarchic clustering of all 163 neat samples, based on the methylation level of all of the >485,000 CpGs interrogated by the array. This clustering revealed an "out-group" composed exclusively of IVF patient samples (Fig. 1A, labeled "patient-only cluster"), which contained within it a subgroup composed almost exclusively of samples that led to poor embryo quality (labeled "poor embryo-quality clusters" in Fig. 1A). However, most IVF patient samples and poor-quality–embryo samples were distributed evenly among fertile donor samples and good-quality–embryo samples outside this group.

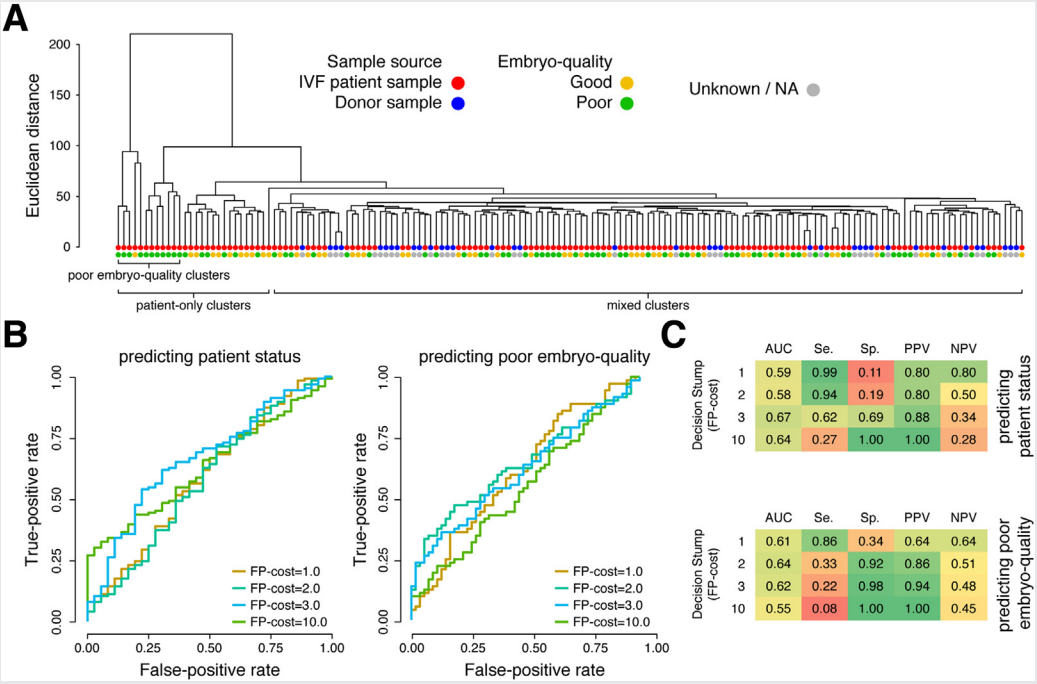
To verify the predictive power of the genomewide methylation differences observed from the clustering, we trained a simple decision-stump classifier using 10-fold cross-validation, in which each sample was described by the Euclidean distance to each of the samples in the training set. We adjusted the false-positive cost of the training algorithm (see Materials and Methods section for details) from 1 (i.e., a false positive was considered as costly as a false negative during training) to 10 (i.e., a false positive was considered 10 times more costly than a false negative when training). We observed that with high false-positive costs, the classifier could achieve specificity >0.9, while identifying one third (33%) of poor-quality–embryo samples, and 27% of IVF patient samples (Fig. 1B and C).

We replicated this analysis on the density gradient–purified samples and observed an even greater separation of good- vs. poor-quality–embryo samples, although little discernible separation of IVF patient and donor samples was observed, possibly because of the small number of purified donor samples (Fig. 2A). This lack of separation was reflected in the classification results from these samples, which showed even greater classifier identification of poor-quality–embryo samples, with just under one half of the poor-quality–embryo patients recovered, with zero false positives, whereas classifier differentiation between IVF patient and donor samples was largely eliminated (Fig. 2B and C).

Is Predictive Differential Methylation Concentrated Within Specific CpGs or Annotation Categories?

To better understand whether differences in methylation between groups were concentrated at particular individual CpGs or regions, we called differentially methylated positions (DMPs) using Wilcoxon's rank sum test. As a control, we repeated this process using randomly shuffled sample labels. The number of differentially methylated CpGs that were identified between neat IVF patient vs. donor samples, both before and after correction for multiple hypothesis testing, is shown in Figure 3A. The histogram of *P* values from all CpGs is

FIGURE 1



Aberrant global methylation profiles are indicative of fertility status and poor embryo quality. (**A**) Hierarchic clustering of 163 neat samples based on a global methylation profile; an "out-group" of exclusively patient samples, with a subgroup strongly enriched for poor embryo samples is apparent. (**B**) The ROC curve (*left*) for prediction of patient status (IVF patient or fertile donor). Classification was performed by building a 1-level decision tree (Supplemental Methods) based on Euclidean distance between samples. Training and testing is performed using 10-fold cross-validation. The same information is shown for the task of predicting embryo quality (*right*). In both cases, high confidence predictions (*bottom right of plots*) have a high probability of being correct. (**C**) Classification statistics for the ROC curves are presented in (**B**). Samples were predicted to be positive (i.e., a patient sample or a poor-quality–embryo sample) when the probability exceeded 50%. In both cases, high specificity and positive predictive value are achieved. AUC = area under the curve; FP = false positive; NPV = negative predictive value; PPV = positive predictive value; Se. = sensitivity; Sp. = specificity.

Aston. Sperm DNA methylation and infertility. Fertil Steril 2015.

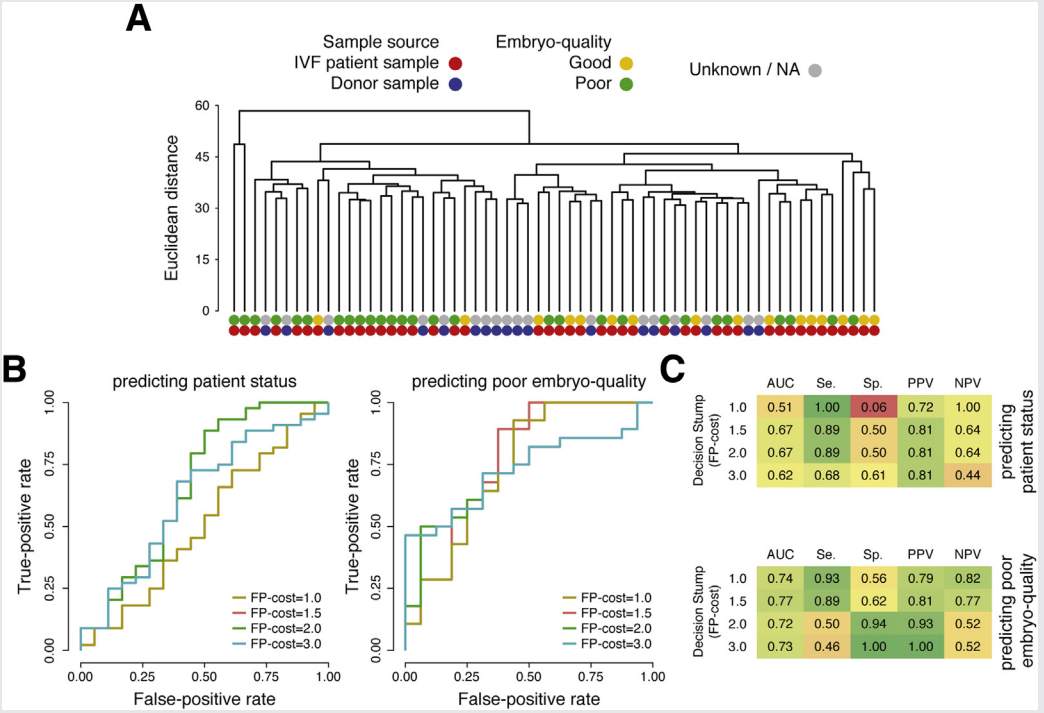
displayed in Figure 3B; it shows a sharp spike in significant *P* values for actual labels, with a relatively flat distribution for the shuffled control. The significant CpGs that were identified using actual labels showed no discernible bias toward particular genomic annotation categories (Fig. 3C).

We used the test for differential methylation at individual CpGs to perform feature selection, with the expectation that the CpGs that show the strongest indication of differential methylation would make good features for predictive models. To estimate how reproducible the selection of particular features (CpGs) would be with various datasets, we broke the samples into 10 groups of approximately equal size. For each group, we repeated the selection of DMPs with those samples that had been held out. We observed that the number of groups for which a CpG was selected in the top 100 most differentially methylated CpGs was generally low when using randomly shuffled labels, and follows a Poisson distribution, as expected. In contrast, use of actual labels produced a U-shaped distribution, with a substantially larger proportion of CpGs always appearing in the top 100 (Supplemental Table 3, available online, gives the genes that are intersected by these consistently differentially methylated CpGs).

We repeated the classifier training and evaluation described previously, this time using only the top 50, 1,000, 50,000, or 400,000 differentially methylated CpGs in the training data to compute Euclidean distances. Receiver operating characteristic curves for each are displayed in Figure 3E; these show a clear advantage to classification using fewer features. We took this process a step further by training a range of models using the β-values for the selected CpGs as features (rather than Euclidean distance between samples). Using a bagged cost-sensitive support vector machine (see Materials and Methods section), we were able to identify IVF patient samples with a positive predictive value of 99%, and a sensitivity of 82%; receiver operating characteristic curves are shown in Figure 3F.

When we applied the same approach to identifying differentially methylated CpGs for good vs. poor-quality–embryo samples (using purified samples, as these showed greater separation in our earlier clustering analysis), we found no significant CpGs after correction for multiple hypothesis testing (Supplemental Fig. 2A, available online). Further, the distribution of *P* values derived from the actual and shuffled labels did not exhibit the differences observed for IVF patient and donor samples (Supplemental Fig. 2B).

FIGURE 2



Gradient purification improves separation of good- and poor-quality sperm samples. (**A**) Hierarchic clustering of 62 gradient-purified sperm samples based on global methylation profile; 2 clear clusters are apparent, 1 of which contains almost exclusively poor-quality–embryo samples. (**B**) The ROC curve is shown (*left*), for prediction of patient status (IVF patient or fertile donor) from the 62 gradient-purified samples. Classification was performed by building a 1-level decision tree (Supplemental Methods) based on Euclidean distance between samples. Training and testing is performed using 10-fold cross-validation. The same information is shown (*right*) for the task of predicting embryo quality from the subset of 45 samples for which this information is known. Although classification of patient status is degraded from that with unpurified samples, the ability to predict poor-quality embryos is markedly improved. (**C**) Classification statistics for the ROC curves presented in (**B**). Samples were predicted as positive (i.e., a patient sample or a poor-quality–embryo sample) when the probability exceeded 50%. In the case of predicting poor-quality embryo, very high specificity and positive predictive value are achieved. AUC = area under the curve; FP = false positive; NPV = negative predictive value; PPV = positive predictive value; Se. = sensitivity; Sp. = specificity.

Aston. Sperm DNA methylation and infertility. Fertil Steril 2015.

As with the DMPs in the neat samples for fertility status, no bias occurred toward regions with any particular annotation category; however, in contrast to the comparison of IVF patients with donors, we saw no tendency for any DMPs to consistently appear in the top 100 when we repeated the feature selection with portions of the data held out, as we did previously (Supplemental Fig. 2C and D). In the training of distance-based classifiers as previously described, we observed that using more CpGs resulted in higher and more-stable specificity and positive predictive value, whereas no discernible pattern in performance was apparent with use of a reduced set of features (Supplemental Fig. 2E).

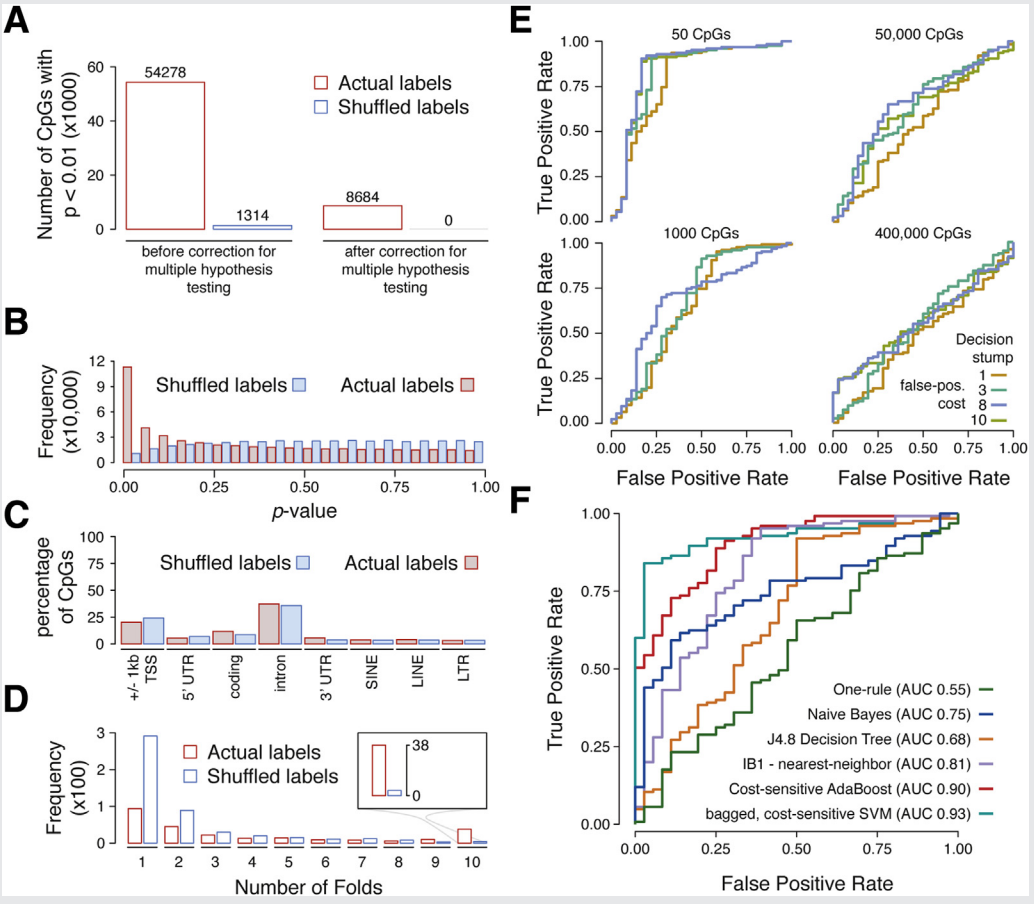
One interpretation of this result is that the poor-quality–embryo samples we are able to differentiate using Euclidean distances create genomewide changes in DNA methylation. However, because we have fewer purified samples, we additionally have reduced statistical power in this setting. Moreover, the platform we have employed here does not give a whole-genome interrogation of methylation. Further studies, including whole-genome bisulfite sequencing, are necessary to more fully explore this finding.

Do Changes in DNA Methylation in CpGs or Promoter Regions Affect Functionally Related Genes?

To better understand biological processes and pathways that are affected by the observed changes in DNA methylation, we ranked genes by the number of CpGs that were identified as differentially methylated within their promoter regions (±5 kb around the transcription start site) and performed gene ontology analysis on the top 1,000 genes. In all cases (both purified and neat samples for IVF patient vs. fertile donor samples, and good- vs. poor-quality–embryo samples) we observed significant enrichment for genes involved in cell adhesion and morphogenesis (Fig. 4A). Additionally, several gene families associated with embryogenesis and development are differentially methylated in the good vs. poor embryogenesis groups.

Following up on this finding, we combined *P* values within promoter regions for differential methylation between purified good- and poor-quality–embryo samples to generate a single *P* value for differential methylation of each gene. We found that the set of differentially methylated genes identified

FIGURE 3



Differential methylation between healthy donor and infertile IVF patient samples consistently occurs at a limited number of CpGs. (**A**) Number of differentially methylated CpGs for unpurified donor vs. IVF patient samples before and after correction for multiple hypothesis testing, using both the actual donor/patient labels and randomly shuffled donor/patient labels as a control. (**B**) Distribution of *P* values for all profiled CpGs from tests of differential methylation between donor and IVF patient samples using both actual labels and randomly shuffled donor/patient labels as a control. (**C**) Proportion of differentially methylated CpGs (before correction for multiple hypothesis testing) that fall within regions annotated as shown, both actual donor/IVF patient labels and randomly permuted labels. (**D**) Samples are split into 10 stratified, equal-size groups. A fold is formed by taking 9 of these groups and leaving 1 group out (allowing 10 separate analyses). For each fold, we use the samples in the 9 retained groups to identify differentially methylated CpGs using both actual donor/IVF patient labels and randomly permuted labels as a control. A histogram showing the number of CpGs that were contained in the top 100 most differentially methylated CpGs identified in only 1 fold, exactly 2 folds, exactly 3 folds, and so on, is displayed. The inset shows the number of CpGs contained in all 10 folds, for both actual labels and permuted labels in higher detail. (**E**) Classifiers, analogous to those in presented in Figure 2 were trained using a subset of the top 50, 1,000, 50,000, and 400,000 most significantly differentially methylated CpGs. The ROC curves are plotted for each. (**F**) A range of classifiers was trained on the top 500 most differentially methylated CpGs; shown are the ROC curves from these classifiers.

Aston. Sperm DNA methylation and infertility. Fertil Steril 2015.

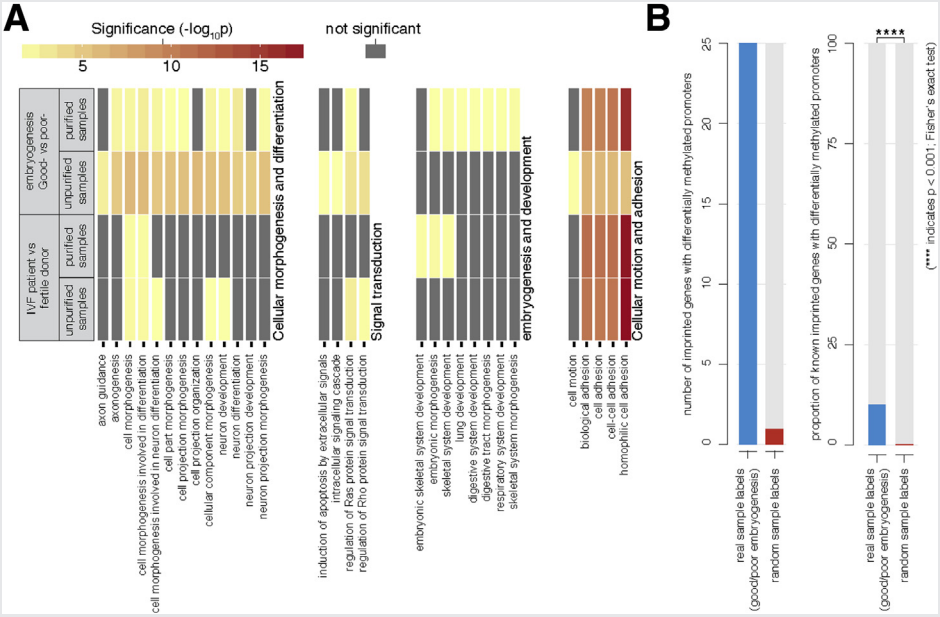
in this way contained 25 imprinted genes, accounting for close to 10% of known imprinted genes. As a control, we randomly permuted the labels on these samples and repeated the analysis; only 1 imprinted gene was identified as differentially methylated after permutation testing (Fig. 4B).

DISCUSSION

The current study evaluated genomewide sperm DNA methylation patterns in a relatively large cohort of normozoospermic, fertile men serving as a control group, and men undergoing IVF, in an effort to evaluate predictability of fertility status and IVF success based on the sperm DNA methylome. The current study has several important limitations. Replication with a larger control group of less-select fertile men, as well as a larger patient cohort, is required to confirm the findings presented here.

In particular, the comparison of methylome signatures using neat samples vs. density gradient–prepared samples, showed that methylation profiles of purified samples were more predictive of embryo quality, whereas those of neat samples were more predictive of fertility status. This finding is intriguing. Purification, in addition to selecting for more-competent sperm, may in addition eliminate immature or incompletely compacted sperm that could indicate poor fertility, possibly accounting for this difference. However,

FIGURE 4



Functional classification of differentially methylated genes. (**A**) Gene ontology analysis of the top 1,000 genes after sorting genes by number of differentially methylated CpGs. All terms that were significant (correct *P* value <.05), in comparing either IVF patients with fertile donors, or good- and poor-embryogenesis IVF, patient samples are included, with the exception of 24 terms associated with gene expression and cellular metabolism, which are omitted for clarity of visualization (full list is provided in Supplemental Table 3, available online). (**B**) The absolute number and proportion of genes with differentially methylated promoter regions (*P*<.01, Wilcoxon's signed rank test, purified samples) that are known to be imprinted.

Aston. Sperm DNA methylation and infertility. Fertil Steril 2015.

with fewer samples, power in the purified case is reduced, thereby decreasing identification of more-subtle effects, and making direct comparison of results from purified vs. neat samples difficult.

Another limitation of this study was the challenge of classifying good vs. poor embryogenesis cases and minimizing inclusion of potentially contributing female infertility factors. Embryogenesis is a complex process, and "poor embryogenesis" can be manifested in several ways, including developmental arrest or attrition at any point during embryonic development, and poor implantation. Moreover, embryogenesis is rarely homogeneous within a cohort of oocytes. Although we made every attempt to select the most-extreme cases (best and worst) while avoiding severe female infertility factors, the cases we selected for the study were heterogeneous by virtue of the inherent complexities of fertilization and embryogenesis.

As illustrated in Supplemental Table 1 (available online), although semen parameters and male and female age did not differ between groups, the poor-embryo, negative pregnancy group had, on average, fewer oocytes, metaphase II oocytes, and normally fertilized oocytes. This finding suggests that cryptic overrepresentation of female infertility factors may have occurred in this group and not in the others. Although these differences are not ideal, they are not expected to affect specificity; assuming sufficient signal for model training, the presence of female factor samples would be expected to result in an increased false-negative rate, thereby reducing sensitivity.

Although the predictive power of this approach for classifying fertile and patient samples is remarkably high in this study, the control samples were from carefully screened, normozoospermic, fertile sperm donors, which do not represent the general population, or even the fertile general population. Future studies should include more-representative samples.

To our knowledge, this study is the largest on genome-wide sperm DNA methylation performed to date. Our findings that sperm DNA methylation patterns are generally very stable across samples from different individuals and across sperm fractions from the same individual are in agreement with those of smaller genomewide sperm DNA methylation studies (44, 45, 47, 48).

This study addressed 2 fundamental questions regarding the prognostic value of sperm DNA methylation patterns in the context of male infertility. We first aimed to evaluate the power of sperm DNA methylation patterns, to distinguish normozoospermic, fertile men from infertile men. Second, we investigated the utility of sperm DNA methylation data for predicting IVF outcomes. Our findings suggest that methylation patterns are highly predictive of fertility status, and quite predictive of IVF embryo quality.

Several relevant, and probably biologically informative, differences were found in the power of methylation data to predict fertility status vs. IVF embryo quality. The primary

difference is the methylation signature for each group. Our analyses demonstrated that no individual CpG displayed significantly different methylation in good- vs. poor-quality–embryo groups, after correcting for multiple comparisons.

Predictive power was poor when we trained models based on a selected set of features (CpGs), rather than on global profiles. As we increased the number of CpGs for the predictive models, the predictive power increased (Supplemental Fig. 2, available online). This result suggests that although poor embryogenesis cannot be attributed to a few consistently altered CpGs, the genomewide methylation profile seems to be inherently different in men in the poor-quality–embryo group.

Another informative result is that with density gradient–purified samples, power to predict good vs. poor embryo quality was significantly improved. Although the reason is unclear, the purification step may have reduced background methylation heterogeneity by eliminating less-competent sperm that, although present in the sample, are unlikely to play a role in fertilization or embryogenesis. This process may thereby have the effect of improving the overall methylation signatures of the 2 groups. The methylation differences in good- vs. poor-embryogenesis samples may be subtle, and dispersed throughout the genome. Additional studies, including whole-genome assays, are necessary to further our understanding of methylation effects.

In contrast, comparison of sperm methylomes of fertile men vs. IVF patients revealed, after multiple-comparison correction, >8,500 CpGs that had significantly different methylation. We employed the same strategy for predictive modeling that was used to classify patients with good- vs. poor-quality embryos. In this case, probably driven by the large number of highly differentially methylated CpGs, we found that models were most effective at classifying samples when they were classified using only those CpGs with the most significant methylation differences (Fig. 3). In a comparison of those in the control group vs. IVF patients, whole ejaculate proved much more effective at classifying samples than did the evaluation of methylation in purified samples; however, this difference may be a result of the small number of purified samples in the control group.

Evaluation of the genomic context of the methylation alterations, as well as the gene classes affected by differential methylation between groups, indicated no enrichment for differential methylation within a specific genomic context (Figs. 3C and Supplemental Fig. 2). Gene ontology analysis indicated highly significant enrichment for several functional classes of genes (Fig. 4). Most notably, genes involved in cellular adhesion were significantly overrepresented among differentially methylated CpGs, for all comparisons. Although we can only speculate about the relevance of this finding, cell adhesion is known to be critical for both embryogenesis (58, 59) and sperm-oocyte fusion (60, 61).

Genes involved in cellular morphogenesis and differentiation were overrepresented in the good vs. poor embryogenesis comparison, and to a lesser extent, in the IVF vs. fertile donor comparison. We found that imprinted genes were significantly over-represented among the list of differentially methylated genes in a comparison of good vs. poor embryogenesis samples. This enrichment was not detected in comparison of IVF patients and fertile donors. This is primarily apparent with purified samples and may explain our observation that methylation status was more strongly related to embryo quality in purified samples; or, these differences may be subtle changes that are masked by other elements in the unpurified samples.

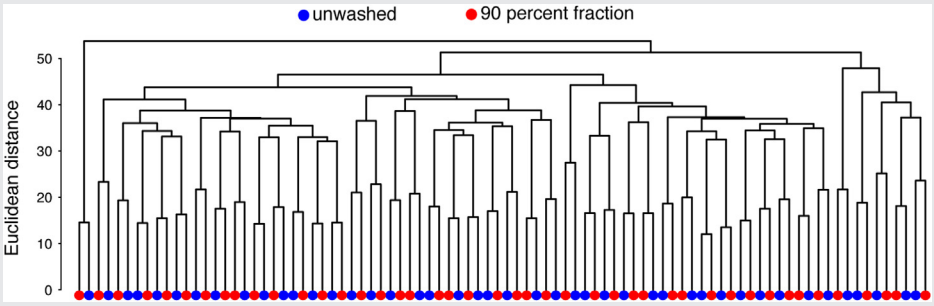
In conclusion, multiple studies have identified differences in sperm DNA methylation at single loci, most often imprinted genes, in normozoospermic vs. infertile men, and a few small studies have used array-based approaches to identify sperm DNA methylation differences between the 2 groups. However, this study is the first to exploit the observed differences to build models predictive for fertility status. The findings presented here provide an exciting and potentially clinically useful metric for assessment of male infertility. Additional studies are needed, to replicate the findings presented here, and to evaluate the generalizability of the present findings in a larger cohort of fertile and subfertile populations.

REFERENCES

1. Esteves SC, Zini A, Aziz N, Alvarez JG, Sabanegh ES Jr, Agarwal A. Critical appraisal of World Health Organization's new reference values for human semen characteristics and effect on diagnosis and treatment of subfertile men. Urology 2012;79:16–22.
2. Barratt CL, Mansell S, Beaton C, Tardif S, Oxenham SK. Diagnostic tools in male infertility—the question of sperm dysfunction. Asian J Androl 2011;13:53–8.
3. Bonde JP, Ernst E, Jensen TK, Hjollund NH, Kolstad H, Henriksen TB, et al. Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. Lancet 1998;352:1172–7.
4. Guzick DS, Overstreet JW, Factor-Litvak P, Brazil CK, Nakajima ST, Coutifaris C, et al. Sperm morphology, motility, and concentration in fertile and infertile men. N Engl J Med 2001;345:1388–93.
5. Gandini L, Menditto A, Chiodo F, Lenzi A. From the European Academy of Andrology. Italian pilot study for an external quality control scheme in semen analysis and antisperm antibiotics detection. Int J Androl 2000;23:1–3.
6. Neuwinger J, Behre HM, Nieschlag E. External quality control in the andrology laboratory: an experimental multicenter trial. Fertil Steril 1990;54:308–14.
7. Brazil C, Swan SH, Tollner CR, Treece C, Drobnis EZ, Wang C, et al. Quality control of laboratory methods for semen evaluation in a multicenter research study. J Androl 2004;25:645–56.
8. Leushuis E, van der Steeg JW, Steures P, Repping S, Bossuyt PM, Blankenstein MA, et al. Reproducibility and reliability of repeated semen analyses in male partners of subfertile couples. Fertil Steril 2010;94:2631–5.
9. World Health Organization (WHO). WHO laboratory manual for the examination and processing of human semen. 5th ed. Geneva: WHO; 2010.
10. Cooper TG, Noonan E, von Eckardstein S, Auger J, Baker HW, Behre HM, et al. World Health Organization reference values for human semen characteristics. Hum Reprod Update 2010;16:231–45.
11. Zini A, Boman JM, Belzile E, Ciampi A. Sperm DNA damage is associated with an increased risk of pregnancy loss after IVF and ICSI: systematic review and meta-analysis. Hum Reprod 2008;23:2663–8.
12. Sanchez R, Toepfer-Petersen E, Aitken RJ, Schill WB. A new method for evaluation of the acrosome reaction in viable human spermatozoa. Andrologia 1991;23:197–203.
13. Lee MA, Trucco GS, Bechtol KB, Wummer N, Kopf GS, Blasco L, et al. Capacitation and acrosome reactions in human spermatozoa monitored by a chlortetracycline fluorescence assay. Fertil Steril 1987;48:649–58.
14. Haas GG Jr, Cines DB, Schreiber AD. Immunologic infertility: identification of patients with antisperm antibody. N Engl J Med 1980;303:722–7.
15. Burkman LJ, Coddington CC, Franken DR, Krugen TF, Rosenwaks Z, Hogen GD. The hemizona assay (HZA): development of a diagnostic test for the binding of human spermatozoa to the human hemizona pellucida to predict fertilization potential. Fertil Steril 1988;49:688–97.

16. Yanagimachi R, Yanagimachi H, Rogers BJ. The use of zona-free animal ova as a test-system for the assessment of the fertilizing capacity of human spermatozoa. Biol Reprod 1976;15:471–6.
17. Bartoov B, Berkovitz A, Eltes F, Kogosowski A, Menezo Y, Barak Y. Real-time fine morphology of motile human sperm cells is associated with IVF-ICSI outcome. J Androl 2002;23:1–8.
18. Huszar G, Ozenci CC, Cayli S, Zavaczki Z, Hansch E, Vigue L. Hyaluronic acid binding by human sperm indicates cellular maturity, viability, and unreacted acrosomal status. Fertil Steril 2003;79(Suppl 3):1616–24.
19. Sigman M, Baazeem A, Zini A. Semen analysis and sperm function assays: What do they mean? Semin Reprod Med 2009;27:115–23.
20. Tiepolo L, Zuffardi O. Localization of factors controlling spermatogenesis in the nonfluorescent portion of the human Y chromosome long arm. Hum Genet 1976;34:119–24.
21. Lanfranco F, Kamischke A, Zitzmann M, Nieschlag E. Klinefelter's syndrome. Lancet 2004;364:273–83.
22. Dam AH, Koscinski I, Kremer JA, Moutou C, Jaeger AS, Oudakker AR, et al. Homozygous mutation in SPATA16 is associated with male infertility in human globozoospermia. Am J Hum Genet 2007;81:813–20.
23. Elinati E, Kuentz P, Redin C, Jaber S, Vanden Meerschaut F, Makarian J, et al. Globozoospermia is mainly due to DPY19L2 deletion via non-allelic homologous recombination involving two recombination hotspots. Hum Mol Genet 2012;21:3695–702.
24. Harbuz R, Zouari R, Pierre V, Ben Khelifa M, Kharouf M, Coutton C, et al. A recurrent deletion of DPY19L2 causes infertility in man by blocking sperm head elongation and acrosome formation. Am J Hum Genet 2011;88:351–61.
25. Koscinski I, Elinati E, Fossard C, Redin C, Muller J, Velez de la Calle J, et al. DPY19L2 deletion as a major cause of globozoospermia. Am J Hum Genet 2011;88:344–50.
26. Aoki VW, Liu L, Jones KP, Hatasaka HH, Gibson M, Peterson CM, et al. Sperm protamine 1/protamine 2 ratios are related to in vitro fertilization pregnancy rates and predictive of fertilization ability. Fertil Steril 2006;86:1408–15.
27. Aoki VW, Moskovtsev SI, Willis J, Liu L, Mullen JB, Carrell DT. DNA integrity is compromised in protamine-deficient human sperm. J Androl 2005;26:741–8.
28. Balhorn R, Reed S, Tanphaichitr N. Aberrant protamine 1/protamine 2 ratios in sperm of infertile human males. Experientia 1988;44:52–5.
29. Carrell DT, Liu L. Altered protamine 2 expression is uncommon in donors of known fertility, but common among men with poor fertilizing capacity, and may reflect other abnormalities of spermiogenesis. J Androl 2001;22:604–10.
30. Perreault SD. Chromatin remodeling in mammalian zygotes. Mutat Res 1992;296:43–55.
31. Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. Distinctive chromatin in human sperm packages genes for embryo development. Nature 2009;460:473–8.
32. Miller D, Brinkworth M, Iles D. Paternal DNA packaging in spermatozoa: more than the sum of its parts? DNA, histones, protamines and epigenetics. Reproduction 2010;139:287–301.
33. Benchaib M, Ajina M, Lornage J, Niveleau A, Durand P, Guerin JF. Quantitation by image analysis of global DNA methylation in human spermatozoa and its prognostic value in in vitro fertilization: a preliminary study. Fertil Steril 2003;80:947–53.
34. Houshdaran S, Cortessis VK, Siegmund K, Yang A, Laird PW, Sokol RZ. Widespread epigenetic abnormalities suggest a broad DNA methylation erasure defect in abnormal human sperm. PLoS One 2007;2:e1289.
35. Kobayashi H, Sato A, Otsu E, Hiura H, Tomatsu C, Utsunomiya T, et al. Aberrant DNA methylation of imprinted loci in sperm from oligospermic patients. Hum Mol Genet 2007;16:2542–51.
36. Marques CJ, Costa P, Vaz B, Carvalho F, Fernandes S, Barros A, et al. Abnormal methylation of imprinted genes in human sperm is associated with oligozoospermia. Mol Hum Reprod 2008;14:67–74.
37. Kobayashi H, Hiura H, John RM, Sato A, Otsu E, Kobayashi N, et al. DNA methylation errors at imprinted loci after assisted conception originate in the parental sperm. Eur J Hum Genet 2009;17:1582–91.
38. Chen SL, Shi XY, Zheng HY, Wu FR, Luo C. Aberrant DNA methylation of imprinted H19 gene in human preimplantation embryos. Fertil Steril 2010;94:2356–8. 2358.e1.
39. Hammoud SS, Purwar J, Pflueger C, Cairns BR, Carrell DT. Alterations in sperm DNA methylation patterns at imprinted loci in two classes of infertility. Fertil Steril 2010;94:1728–33.
40. Navarro-Costa P, Nogueira P, Carvalho M, Leal F, Cordeiro I, Calhaz-Jorge C, et al. Incorrect DNA methylation of the DAZL promoter CpG island associates with defective human sperm. Hum Reprod 2010;25:2647–54.
41. Poplinski A, Tuttelmann F, Kanber D, Horsthemke B, Gromoll J. Idiopathic male infertility is strongly associated with aberrant methylation of MEST and IGF2/H19 ICR1. Int J Androl 2010;33:642–9.
42. Sato A, Hiura H, Okae H, Miyauchi N, Abe Y, Utsunomiya T, et al. Assessing loss of imprint methylation in sperm from subfertile men using novel methylation polymerase chain reaction Luminex analysis. Fertil Steril 2011;95:129–34.
43. Nanassy L, Carrell DT. Abnormal methylation of the promoter of CREM is broadly associated with male factor infertility and poor sperm quality but is improved in sperm selected by density gradient centrifugation. Fertil Steril 2011;95:2310–4.
44. Pacheco SE, Houseman EA, Christensen BC, Marsit CJ, Kelsey KT, Sigman M, et al. Integrative DNA methylation and gene expression analyses identify DNA packaging and epigenetic regulatory genes associated with low motility sperm. PLoS One 2011;6:e20280.
45. Aston KI, Punj V, Liu L, Carrell DT. Genome-wide sperm deoxyribonucleic acid methylation is altered in some men with abnormal chromatin packaging or poor in vitro fertilization embryogenesis. Fertil Steril 2012;97:285–92.e4.
46. Klaver R, Tuttelmann F, Bleiziffer A, Haaf T, Kliesch S, Gromoll J. DNA methylation in spermatozoa as a prospective marker in andrology. Andrology 2013;1:731–40.
47. Krausz C, Sandoval J, Sayols S, Chianese C, Giachini C, Heyn H, et al. Novel insights into DNA methylation features in spermatozoa: stability and peculiarities. PloS One 2012;7:e44479.
48. Jenkins TG, Aston KI, Trost C, Farley J, Hotaling JM, Carrell DT. Intra-sample heterogeneity of sperm DNA methylation. Mol Hum Reprod 2015;21:313–9.
49. Kitamura A, Miyauchi N, Hamada H, Hiura H, Chiba H, Okae H, et al. Epigenetic alterations in sperm associated with male infertility. Congenit Anom (Kyoto) 2015;55:133–44.
50. Urdinguio RG, Bayón GF, Dmitrijeva M, Toraño EG, Bravo C, Fraga MF, et al. Aberrant DNA methylation patterns of spermatozoa in men with unexplained infertility. Hum Reprod 2015;30:1014–28.
51. Jenkins TG, Aston KI, Pflueger C, Cairns BR, Carrell DT. Age-associated sperm DNA methylation alterations: possible implications in offspring disease susceptibility. PLos Genet 2014;10:e1004458.
52. Nanassy L, Peterson CA, Wilcox AL, Peterson CM, Hammoud A, Carrell DT. Comparison of 5% and ambient oxygen during days 3-5 of in vitro culture of human embryos. Fertil Steril 2010;93:579–85.
53. 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.
54. Maksimovic J, Gordon L, Oshlack A. SWAN: subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. Genome Biol 2012;13:R44.
55. Pruitt KD, Brown GR, Hiatt SM, Thibaud-Nissen F, Astashyn A, Ermolaeva O, et al. RefSeq: an update on mammalian reference sequences. Nucleic Acids Res 2014;42:D756–63.
56. Stouffer SA, Suchman EA, DeVinney LC, Star SA, Williams RM. Adjustment during Army life. Princeton, NJ: Princeton University Press; 1949.
57. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Statis Soc, Ser B 1995;57:289–300.
58. Thiery JP, Duband JL, Rutishauser U, Edelman GM. Cell adhesion molecules in early chicken embryogenesis. Proc Natl Acad Sci U S A 1982;79:6737–41.
59. Ekblom P, Vestweber D, Kemler R. Cell-matrix interactions and cell adhesion during development. Annu Rev Cell Biol 1986;2:27–47.
60. Wassarman PM. Mammalian fertilization: molecular aspects of gamete adhesion, exocytosis, and fusion. Cell 1999;96:175–83.
61. Talbot P, Shur BD, Myles DG. Cell adhesion and fertilization: steps in oocyte transport, sperm-zona pellucida interactions, and sperm-egg fusion. Biol Reprod 2003;68:1–9.

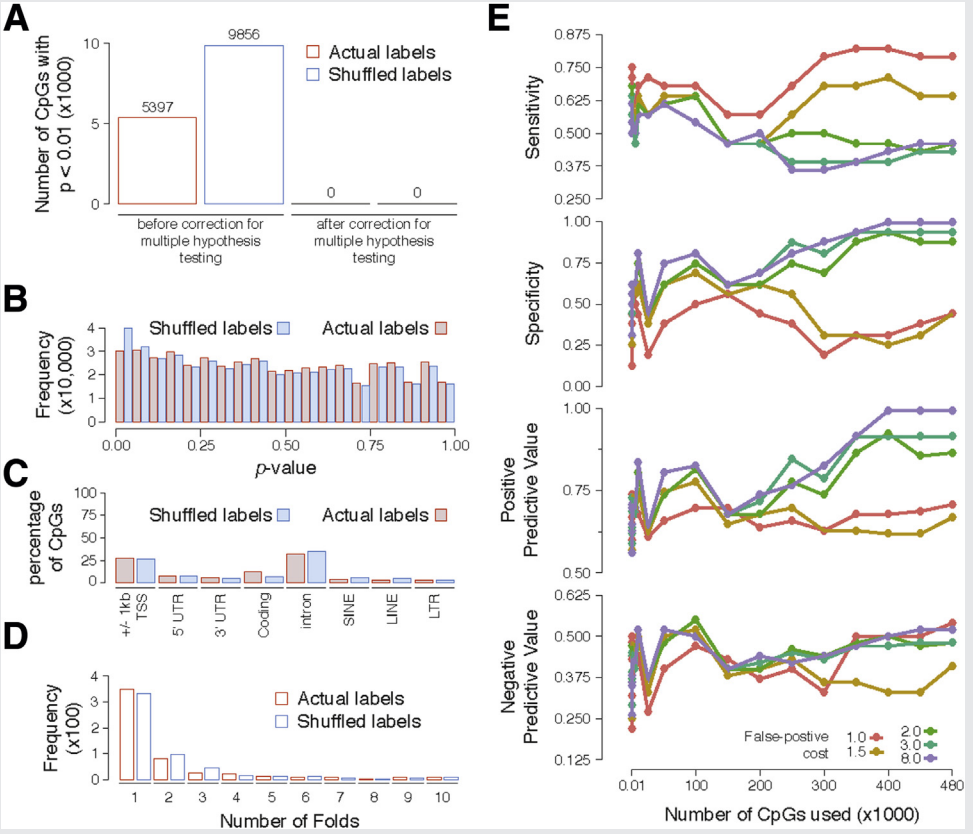
SUPPLEMENTAL FIGURE 1



Hierarchic clustering of methylation profiles for the 44 IVF patients for which both purified and unpurified samples were processed (i.e., 88 samples are plotted). For each unpurified sample, its nearest neighbor is the purified sample from the same patient.

Aston. Sperm DNA methylation and infertility. Fertil Steril 2015.

SUPPLEMENTAL FIGURE 2



Differential methylation between purified good and poor embryogenesis samples is not defined by a small group of consistent single-CpG differences. (**A**) Number of differentially methylated CpGs of purified good vs. poor embryogenesis samples before and after correction for multiple hypothesis testing, using both the actual good/poor labels and randomly shuffled good/poor labels as a control. (**B**) Distribution of *P* values for all profiled CpGs from test of differential methylation between good and poor embryogenesis, using both actual labels and randomly shuffled good/poor labels as a control. (**C**) Proportion of differentially methylated CpGs (before correction for multiple hypothesis testing) that fall within regions annotated as shown, both actual good/poor embryo quality labels and randomly permuted labels. (**D**) Samples are split into 10 stratified, equal-size groups. A fold is formed by taking 9 of these groups and leaving 1 group out (allowing 10 ways of doing this). For each fold, we use the samples in the 9 retained groups to identify differentially methylated CpGs, using both actual good/poor embryogenesis labels and randomly permuted labels as a control. A histogram showing the number of CpGs that were contained in the top 100 most differentially methylated CpGs identified in only 1 fold, exactly 2 folds, exactly 3 folds, and so on, is shown. (**E**) Classifiers, analogous to those in presented in Figure 2, were trained using a subset of the top x most differentially methylated CpGs (where x is varied along the x-axis of the plots), and the sensitivity, specificity, positive predictive value, and negative predictive value of each was evaluated as a function of how many CpGs were selected.

Aston. Sperm DNA methylation and infertility. Fertil Steril 2015.

SUPPLEMENTAL TABLE 1

General composition of the study groups, including semen parameters and IVF embryo quality.

Variable	Good embryos, pregnant		Poor embryos, not pregnant		Poor embryos, pregnant		*P* value
Male age (y)	33.08	5.97	32.98	4.29	33.15	5.03	.9913
Sperm concentration (M/ml)	58.58	62.78	69.31	77.13	76.26	71.60	.4527
Progressively motile sperm (%)	45.00	20.15	42.64	18.29	43.44	24.86	.8867
Female age (y)	30.93	4.52	32.60	5.12	30.98	4.51	.242
Eggs retrieved, n	14.07	5.91	10.47	4.13	12.98	4.72	.0103
MII eggs, n	12.09	5.15	8.87	3.89	11.29	3.83	.0073
Eggs fertilized normally, n	10.74	4.82	7.23	3.33	9.26	4.00	.0019
Embryos cryopreserved, n	4.21	3.10	0.26	1.16	0.81	1.55	$<.0001$
Fertilized eggs $\geq$level 2: 6-cell on day 3 (%)	0.73	0.23	0.37	0.30	0.45	0.30	$<.0001$
Fertilized eggs $\geq$level 2: early blast on day 5/6 (%)	0.44	0.18	0.07	0.09	0.13	0.15	$<.0001$
Embryos transferred, n	2.04	0.55	2.17	0.80	2.34	0.53	.0596

Note: P values are from ANOVA. Values for each embryo/pregnancy category are mean followed by SD.

Aston. Sperm DNA methylation and infertility. Fertil Steril 2015.

SUPPLEMENTAL TABLE 2

Frequency of male and mild female factors in 127 IVF patient-couples.

Factor	Good embryogenesis, positive pregnancy (n = 53)	Poor embryogenesis, negative pregnancy (n = 31)	Poor embryogenesis, positive pregnancy (n = 42)
Male factor only	15	12	14
Female factor only	16	13	17
Male and female factor	16	3	6
Idiopathic (unexplained)	6	3	5
Endometriosis	10	3	10
Diminished ovarian reserve	2	5	3
Polycystic ovary syndrome	8	3	3

Note: No significant differences were observed among groups.

Aston. Sperm DNA methylation and infertility. Fertil Steril 2015.

SUPPLEMENTAL TABLE 3

Genes that overlap CpGs consistently in the top 100 most differentially methylated CpGs, in comparison of unpurified donor vs. IVF patient samples.

Gene
AHDC1
ALOX5AP
BTBD17
CXXC11
EEF1A2
FBLN2
FGF18
GRM6
HIST1H4J
HIST1H4K
INPP5A
JAG2
KCNQ1
KIAA0319L
LRRC45
MIR4734
MLLT6
MTMR6
MXRA7
NCDN
NDUFS6
NDUFS8
OGFOD2
PSTPIP1
RAP1GAP2
SERPINF2
STRA13
SYT8
TCIRG1
TNNI2
USP24

Aston. Sperm DNA methylation and infertility. Fertil Steril 2015.