TITLE:

Paternal cigarette smoke alters DNA methylation in sperm and gene expression in offspring brain

SHORT TITLE:

Paternal smoking, sperm DNA methylation and impacts in offspring

AUTHORS:

†,*Patrick J Murphy[1,3], †Jingtao Guo[2,3], Timothy G Jenkins[2], Emma R James[2,6], John R Hoidal[4], Thomas Hueckstead[4], James M Hotaling[2], Douglas T Carrell[2,5,6], *Bradley R Cairns[3], *Kenneth I Aston[2]

† These authors contributed equally to this work.

* Corresponding authors: kiaston@utah.edu, patrick_murphy@urmc.rochester.edu, brad.cairns@hci.utah.edu

1 Departments of Biomedical Genetics, Wilmot Cancer Institute, University of Rochester Medical Center, Rochester NY, USA.

2 Andrology and IVF Laboratories, Department of Surgery, University of Utah School of Medicine, Salt Lake City, Utah, USA

3 Howard Hughes Medical Institute, Department of Oncological Sciences and Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, Utah, USA

4 Department of Internal Medicine, University of Utah School of Medicine and Salt Lake VA Medical Center, Salt Lake City, Utah, USA

5 Department of Obstetrics and Gynecology, University of Utah School of Medicine, Salt Lake City, Utah, USA

6 Department of Genetics, University of Utah School of Medicine, Salt Lake City, Utah, USA

ABSTRACT:

Background: There is growing evidence that paternal pre-conception cigarette smoke (CS) exposure is associated with increased risk of behavioral disorders and cancer in offspring. The aim of the current study was to evaluate the impact of paternal pre-conception CS exposure on sperm DNA methylation and offspring phenotype.

Methodology/Principal: To characterize the effects of CS exposure on the sperm epigenome and offspring neurodevelopment, we exposed male mice to CS and bred exposed and control males to unexposed females and subsequently evaluated sperm DNA methylation in sires and frontal cortex DNA methylation and gene expression in offspring. We further investigated the role of oxidative stress on sperm epigenetic changes using a mouse model (*Nrf2*$^{-/-}$) with impaired antioxidant capacity. Lastly, we evaluated the capacity for sperm DNA methylation to recover following removal of CS for 1-5 spermatogenic cycles (28-171 days).

Conclusions/Significance: Smoking significantly impacts sperm DNA methylation as well as DNA methylation and gene expression in offspring. These changes were largely recapitulated in *Nrf2*$^{-/-}$ mice independent of smoke exposure. Recovery experiments indicated that about half of differentially methylated regions returned to normal within 28 days of removal from smoke exposure, however additional recovery following longer periods was not observed. We present strong evidence that cigarette smoke exposure induces paternally mediated, heritable epigenetic changes. Parallel studies performed in *Nrf2*$^{-/-}$ mice provide evidence for oxidative stress as the predominant underlying mechanism for smoke-induced

epigenetic changes to sperm along with the associated effects in offspring. Lastly, recovery experiments indicate that while many epigenetic changes are corrected following removal from smoke exposure, aberrant methylation persists at a significant number of regions even after five spermatogenic cycles

KEYWORDS: sperm DNA methylation, epigenetics, gene expression, frontal cortex, smoking, heritable, cigarette, oxidative stress

INTRODUCTION:

Cigarette smoke (CS) exposure is a global epidemic with significant health consequences. In a recent study it was estimated that more than one third of the world's population is regularly exposed, directly or indirectly, to tobacco smoke (1). Further, it was estimated that in 2004 involuntary exposure to tobacco smoke caused 603,000 premature deaths and a loss of 10.9 million disability adjusted life years, and an estimated six million annual deaths are attributable to tobacco smoke exposure (1). Tobacco smoke contains more than 4000 chemicals including a myriad of known carcinogens. The health consequences of smoke exposure are significant and include numerous diseases and dysfunctions of the respiratory tract, increased risk of multiple types of cancer and increased incidence of cardiovascular disease (2, 3). Tobacco smoke exposure is a global problem, the implications of which are becoming increasingly apparent. However, little is known about the impact of paternal exposure to cigarette smoke (CS) on sperm and implications of pre-conception paternal CS on offspring health (4).

83

84 Male fertility rates have steadily declined in developed countries over the past half-
85 century (5-7), and metabolic disorders have steadily increased over a similar
86 period (8, 9). These trends are due to a variety of factors, but increased exposure
87 to environmental toxins is likely a significant contributor. Tobacco smoke exposure,
88 which negatively affects semen quality (10, 11), might contribute to these trends.
89 Moreover, *in utero* exposure to maternal tobacco smoke increases the risk of
90 obesity and hypertension in offspring (12, 13). A more complete understanding of
91 the effects of preconception paternal tobacco smoke exposure on offspring is of
92 critical relevance to public health.

93

94 Although understudied, there is evidence for paternally transmitted effects of CS
95 exposure to offspring. The negative impacts of CS on semen parameters are well
96 established. Smoking is associated with an accumulation of cadmium and lead in
97 seminal plasma, reduced sperm count and motility, and increased morphological
98 abnormalities in sperm (10, 11). In addition, reduced reproductive potential has
99 been reported in tobacco smoke-exposed mice (14) and humans (15). Although
100 the impact of tobacco smoke exposure on germ cell genetics and epigenetics has
101 received surprisingly little attention, recent evidence demonstrates strong effects
102 on the adult male germline. Adult male mice exposed to sidestream tobacco smoke
103 display significant increases in sperm DNA mutations at expanded simple tandem
104 repeats (ESTRs) (16), as well as more frequent aberrations in sperm chromatin
105 structure and elevated sperm DNA damage (14). In contrast, CS-exposed male

106 mice exhibit no measurable increase in somatic cell chromosome damage,
107 indicating that germ cells may be more prone to environmentally-induced genetic
108 and/or epigenetic insults compared with somatic cells. Destabilization of ESTRs in
109 the adult male germline is well documented to be associated with
110 transgenerational effects in the mouse genome (17-21).
111
112 The International Association for Research on Cancer recently declared that
113 *paternal* smoking *prior* to pregnancy is associated with a significantly elevated risk
114 of leukemia in the offspring (22), suggesting CS-induced genetic or epigenetic
115 changes occur in sperm that are transmitted to offspring. We recently reported
116 altered sperm DNA methylation (DNAme) patterns in men who smoke (23). In
117 addition, smoking has been clearly shown to modify DNAme patterns and gene
118 expression in somatic tissues in individuals exposed to first- or second-hand
119 tobacco smoke (24-26) as well as in newborns of smoking mothers (27-29). A
120 growing body of evidence suggests that CS exposure could have negative health
121 consequences not only for the exposed fathers, but also for their offspring.
122
123 Results of these studies motivated us to investigate the impact of smoking on
124 mouse sperm DNAme as well as the potential transmission of those effects to
125 offspring. Additionally, we aimed to explore the dynamics of sperm epigenetic
126 changes after withdrawal from smoke, to determine whether DNAme changes
127 recover following removal of the insult for more than five complete cycles of
128 spermatogenesis (171 days; Figure 1).

129

130 Lastly, to investigate the mechanism underlying CS-induced alterations to the
131 sperm epigenetic profile, we used a knockout mouse model to assess the role of
132 antioxidant capacity on smoke-induced alterations to the methylome. The nuclear
133 factor (erythroid-derived 2)-like 2 (NRF2) pathway is the primary cellular defense
134 against the cytotoxic effects of oxidative stress. Thus, to investigate the mode by
135 which CS induced epigenetic changes might occur, we utilized the *Nrf2*$^{-/-}$ mouse
136 model, which has compromised antioxidant capacity.

137

138 **RESULTS:**

139 **Experimental Design and phenotypic effects:**

140 Male mice were assigned to CS- exposed or non-exposed groups (n = 10-12 per
141 group). The CS animals were exposed to the body mass-adjusted equivalent of
142 10-20 cigarettes per day, 5 days per week over a period of 60 days. CS- exposed
143 and control mice were bred to unexposed females, and offspring were analyzed
144 for phenotypic and molecular measures.

145

146 In agreement with the literature, both WT and *Nrf2*$^{-/-}$ mice that were exposed to
147 CS weighed significantly less than non-exposed control animals (TABLE 1).
148 Sperm concentration and motility were not significantly impacted by CS
149 exposure, sperm concentration was lower in *Nrf2*$^{-/-}$ than WT males, and
150 conception was significantly delayed in CS-exposed *Nrf2*$^{-/-}$mice compared with
151 unexposed *Nrf2*$^{-/-}$mice (TABLE 1). Neither growth trajectories nor sperm

152 parameters were different in F1 animals based on paternal CS exposure (TABLE
153 2).
154
155 **Smoking-induced DNAme changes in F0 sperm were either maintained,**
156 **recovered, or variable, depending on the locus.**
157 Initially, we performed reduced representation bisulfite sequencing (RRBS) to
158 explore the effects of smoking on F0 sperm DNAme patterns (sperm collected
159 within 3 days of completing a 60-day smoking treatment), and we examined
160 whether CS-induced methylation changes recover to baseline unexposed levels
161 following removal of CS exposure (28, 103 and 171 days after smoking
162 treatment; n = 10 per group).
163
164 At individual CpG sites, we found changes in DNA methylation at a large number
165 of CpGs (Figure 2A & Figure S1A) with essentially equal representation of sites
166 that lost methylation and gained methylation as a consequence of smoke
167 exposure. In addition, we found that the number of differentially methylated CpGs
168 declined only slightly after removal of CS exposure for 28-171 days (Figure 2A).
169
170 We partitioned differentially methylated CpGs into three groups: 1) shared,
171 meaning differentially methylated CpGs that were maintained between the
172 treatment group and the respective recovery group, 2) recovered, meaning CpGs
173 that were differentially methylated after initial exposure but were no longer
174 differentially methylated in the recovery group, and 3) new, meaning CpGs that

175 were not differentially methylated after initial exposure but emerged as newly
176 differentially methylated in the recovery groups. These classes were assessed
177 separately for CpGs that initially lost methylation verses those that initially gained
178 methylation relative to the control. Interestingly, for CpGs that lost methylation,
179 the three groups were relatively evenly represented across recovery times, with
180 only a slight over-representation of shared CpGs. For CpGs that gained
181 methylation, there was a slight bias toward the emergence of new differentially
182 methylated CpGs across recovery groups (Figure 2B).
183
184 Given that DNAme status of CpG regions likely has a greater capacity to confer
185 functional effects compared with individual sites, we subsequently binned
186 individual CpG sites into regions based on their proximity (see methods). We
187 then performed analyses similar to those performed for individual CpGs.
188 Interestingly, changes in DNA methylation occurred at far fewer regions
189 compared with individual CpGs (Figure S1B), and we found strong evidence for
190 recovery of smoke-induced sperm DNAme changes following a recovery period.
191 Strikingly, the majority of DNAme regions that recovered returned to control
192 levels within 28 days of removal of smoke exposure, with no evidence for
193 additional correction following longer recovery periods (Figure 2C). In contrast
194 with the individual CpG data, we found that far fewer new differentially
195 methylated regions (DMRs) emerged during the recovery period (Figure 2D).
196

197 **Regions of high and low DNAme recovered, and changes were associated**
198 **with Nrf2-mediated effects.**
199 We then sought to understand epigenetic properties that might impact the ability
200 of specific regions to recover following removal of the CS, including the initial
201 methylation status. For this analysis, we classified all DMRs into six classes,
202 based on their dynamics during recovery (shared, recovered or new) and their
203 direction of change (increase or decrease; Figure 3A). We discovered that the
204 initially hypomethylated DMRs in which methylation increased with smoke
205 exposure displayed a higher likelihood of recovery, and initially hypermethylated
206 DMRs that decreased in methylation were likewise more likely to recover (Figure
207 3A). Regions of intermediate DNAme were less likely to recover across all
208 groups. In addition, we observed that recovered DMRs that initially decreased in
209 methylation level show lower variation compared to all other groups (Figure
210 S2A), thus suggesting that regions of extreme hyper- and hypo-DNAme were
211 less likely to change after CS exposure, and when changes did occur, these
212 regions were more likely to recover.
213
214 We hypothesized that CS- induced alterations to the methylome might depend on
215 antioxidant capacity, and that reducing antioxidant capacity might mimic or
216 enhance smoke-induced changes to sperm DNAme. This led us to examine the
217 changes of sperm DNAme in *Nrf2*$^{-/-}$ mice (compared to WT mice), as well as the
218 changes in sperm DNAme in CS-exposed compared with unexposed *Nrf2*$^{-/-}$ mice.
219 We initially found that the degree of DNA methylation changes, and the number

220 of sites impacted, were similar in *Nrf2*$^{-/-}$ mice compared with WT (Figure S3A).
221 We were surprised to find that the CS-induced DNAme changes in WT sperm
222 were largely recapitulated in *Nrf2*$^{-/-}$ sperm irrespective of CS exposure (Figure
223 3B). Notably, the sperm DNAme changes observed in unexposed *Nrf2*$^{-/-}$ mice
224 were similar to changes observed in *Nrf2*$^{-/-}$ mice that were exposed to CS, with no
225 evidence for an enhanced effect attributable to CS. This correlation in the degree
226 of DNAme change was consistent when assessed genome-wide, and we did not
227 observe this correlation when comparing our data to other datasets evaluating
228 the effect of other environmental factors (vinclozolin exposure and protein
229 restricted diet) (30, 31) on sperm DNAme (Figure 3C).
230
231 **Smoking-induced DNAme changes in F1 brains were regulated by oxidative**
232 **stress**
233 CS-induced epigenetic changes in sperm might correspond with an affect in
234 offspring if they are maintained through fertilization, or if those DNAme changes
235 are associated with an alternative, unidentified inheritance mechanism. Thus, we
236 aimed to test whether paternal CS exposure could impact the next generation
237 (F1). Given that brain and nervous systems were previously reported to be
238 sensitive to preconception paternal exposures (32-35), we investigated DNAme
239 by RRBS in the prefrontal cortex of F1 mice derived from CS-exposed WT and
240 *Nrf2*$^{-/-}$ sires compared to the offspring of unexposed males (n = 8 per group).
241 Similar to observations in F0 sperm, we found that paternal smoking altered
242 DNAme patterns in offspring brains, and these DNAme changes were highly

243 similar to brain DNAme changes observed in offspring of both unexposed *Nrf2*$^{-/-}$
244 males (r=0.609) (Figure 4A) and CS-exposed *Nrf2*$^{-/-}$ males (r=0.477). In
245 agreement with F0 sperm DNAme data, exposure to CS appeared to have no
246 additional impact on brain DNAme beyond the *Nrf2*$^{-/-}$ effect (Figure 4A - right).
247 These observations suggest that CS-induced DNAme changes that occur in F1
248 brains are mediated by paternal oxidative stress, a well-established effect of
249 NRF2 depletion (36, 37).
250
251 To investigate whether F0 CS-induced effects can be passed to F2, we
252 compared the smoke-associated DNAme changes observed in F0 WT sperm
253 with that of F1 sperm and found no correlation (r=-0.001; Figure 4B). We also
254 observed that paternal CS-induced DMRs in F1 brains and those in smoke-
255 exposed F0 sperm had minimal overlap (hypergeometric p-value =1; Figure 4C),
256 indicating a lack of maintained DNAme changes from sperm through fertilization,
257 into brain development. Accordingly, regions where DNA methylation changes
258 occurred are not marked by chromatin features that are known to be maintained
259 in mature sperm (38)(FIGURE S4). Regardless of whether changes in DNAme
260 are a marker of epigenetic inheritance, or a driver of epigenetic inheritance, these
261 results suggest that the effects of CS exposure do not persist through several
262 generations.
263
264 **Smoking causes high gene expression variations in F1 brains**

265 To further investigate the potential for phenotypic effects in offspring associated
266 with paternal CS exposure, we performed RNA-seq of the F1 frontal cortex in the
267 same animals as those assessed for frontal cortex DNAme. Interestingly,
268 paternal CS exposure caused a globally elevated variation in gene expression in
269 the F1 brains, which was also observed in the *Nrf2*$^{-/-}$ offspring (Figure 5A). In
270 agreement with the F1 brain DNAme data, we found that the paternal CS-
271 induced gene expression changes in WT were highly similar to the Nrf2$^{-/-}$
272 offspring, and the effects were not further elevated in the offspring of smoked
273 Nrf2$^{-/-}$ mice (Figure 5B & S4D-E). Notably, this increased variation limited our
274 ability to identify reliably differentially expressed genes. Instead, we ranked
275 genes by changes in gene expression and performed gene ontology (GO)
276 analysis. We found that gene transcripts associated with classes including
277 neuropeptide receptors and hormone activity were over-represented in the
278 offspring of CS-exposed males (Figure 5C), while gene transcripts associated
279 with gene classes including immune response and metabolism were under-
280 represented (Figure 5D).
281
282 **DISCUSSION:**
283 Recent studies have demonstrated that paternal preconception exposures to a
284 variety of pharmacologic agents and pollutants, including nicotine, THC,
285 morphine, and benzo[a]pyrene effect offspring phenotype, and often confer
286 neurobehavioral consequences (32-35, 39-42). In some cases, the impacts are
287 inherited transgenerationally (40, 41). The direct assessment of

288 transgenerational inheritance through F2 was beyond the scope of the current
289 study, however the absence of correlation between DMRs in F1 sperm and F2
290 sperm suggests that the affects we observed in F1 frontal cortex are not likely to
291 be maintained in a second generation, although an alternate non-DNA-
292 methylation associated mechanism for transgenerational inheritance mediated
293 through noncoding RNAs or chromatin structure cannot be ruled out. Our data
294 suggest that in our experimental model, oxidative stress is the major contributor
295 to sperm DNA methylation changes and subsequent effects in the offspring. This
296 does not exclude the possibility that other factors, such as low-level nicotine
297 exposure might also contribute to DNAme effects (32, 33, 39, 42), but we
298 suspect these additional factors are only minor contributors, based on the
299 absence of transgenerational effects in our system.
300
301 CS exposure clearly impacts sperm DNAme patterns, however it is reassuring
302 that many of the DNAme effects that occurred at high CpG dense regions were
303 corrected within 28 days. Genomic regulatory regions (such as promoters and
304 enhancers) tend to have high CpG density in the mouse. Thus, DNA methylation
305 changes at these regions, which might otherwise impact gene regulation, are
306 likely to be relatively short-lived. These data are in agreement with previous
307 studies that found that CS exposure significantly impacts DNA methylation
308 patterns in whole blood, and CS-associated methylation changes are largely
309 corrected following smoking cessation in a time-dependent manner (43, 44).
310

311 We observed recovery after just twenty-eight days, which corresponds to just
312 under the duration of a full spermatogenic cycle in mice of 30 days. In man, a
313 spermatogenic cycle is 67 days, and additional research is required to
314 characterize the similarities and differences in the dynamics of sperm DNAme
315 alterations between mice and men. Additionally, it is not yet known whether the
316 corrections observed following removal of CS exposure ameliorate the affects
317 observed in offspring, or whether those effects are driven by the minority of
318 regions that persist following CS removal. Addressing these unknowns will be of
319 great interest for follow-up studies.
320
321 Interestingly, we found that regions of extreme hyper- and hypo-DNAme were
322 less likely to be impacted by CS exposure, and when changes did occur, these
323 regions were more likely to recover. This is consistent with a recent monozygotic
324 twin study that investigated differences in blood DNAme in twins discordant for
325 smoking (45). Additionally, our observation that individual CpGs did not recover
326 after smoke exposure to near the same degree as more CpG dense regions
327 suggests that CpG density might provide a degree of “buffering” against
328 environmental insults. If genomic regulatory function were held by individual
329 CpGs rather than by clusters of CpGs, minor variations or fluctuation in DNAme
330 levels might have significant impacts on physiology. Instead, our data suggests
331 that regions of higher CpG density, which are often found at gene regulator
332 regions like promoters (46), vary less in DNAme level than individual CpGs, and
333 therefore, these regions may be able to withstand a higher degree of

334 environmental insult without having lasting effects on physiology. This robustness
335 in function might explain in part why genomic regulatory regions have a high
336 degree of CpG density.
337
338 We investigated the impact of CS exposure in the oxidative stress-compromised
339 *Nrf2*$^{-/-}$ mouse strain with the expectation that we would observe more profound
340 effects of CS exposure on sperm epigenetic changes. Remarkably, while we did
341 not observe evidence for elevated susceptibility to CS-induced sperm DNAme
342 changes in *Nrf2*$^{-/-}$, we found the CS-effects observed in WT animals were
343 commensurately recapitulated in *Nrf2*$^{-/-}$ mice independent of CS exposure,
344 suggesting that elevated oxidative stress is the primary mechanism for CS-
345 mediated sperm epigenetic changes. These findings were remarkably consistent
346 with the observations of offspring frontal cortex DNAme and gene expression
347 changes as well. While CS-exposure may represent among the most extreme
348 examples of environmental insults that induce oxidative stress, the list of
349 environmentally relevant exposures that impact oxidative stress are innumerable,
350 and the results presented here indicate that all such exposures could potentially
351 impact the epigenetic status of the paternal germline and thus offspring
352 phenotype. In addition, our work provides strong evidence that *Nrf2*$^{-/-}$ mice can
353 serve as a good animal model to study CS-exposure induced effects. Further
354 studies may focus on investigating the molecular mechanisms underlying NRF-
355 mediated DNAme alterations.
356

357 Lastly, by comparing the DNAme and gene expression patterns in the frontal
358 cortex between offspring of male mice exposed to CS with those not exposed,
359 we found strong evidence for an impact of paternal smoking on offspring
360 phenotype. In keeping with data obtained from *Nrf2*$^{-/-}$ mice, we suggest that
361 oxidative stress agents might contribute to behavioral or developmental impacts
362 in offspring. Notably however, our data showed little overlap of DNAme changes
363 in the F1 frontal cortex with DNAme changes in the F0 sperm. While this was
364 somewhat expected, as DNAme state undergoes dramatic reprogramming during
365 early development and neuronal differentiation, it is important to highlight —
366 direct mitotic inheritance of DNAme state is unlikely to mechanistically contribute
367 to the oxidative stress effects we observed in F1 mice. Furthermore, we propose
368 that DNAme may be more accurately described as a marker of epigenetic
369 inheritance and not a mechanistic driver in transmitting environmental impacts
370 between generations. Additional studies are necessary to confirm this and to
371 distinguish other epigenetic features as markers or drivers, including chromatin
372 status and small RNAs, which could play a significant role in inheritance. Further,
373 the observation that CS-associated sperm DNAme changes were not identified in
374 the sperm of the F1 generation offers reassuring evidence that the changes
375 observed in F1 animals likely would not persist in the F2 generation. However,
376 direct studies to evaluate the potential for transgenerational impacts of CS
377 exposure are warranted. The implications of this for human spermatogenesis and
378 transgenerational inheritance also warrant further investigation.

379 Here we present strong evidence for: 1) a significant impact of CS-exposure on
380 the male germline, 2) relatively rapid partial correction of CS-induced sperm
381 DNAme alterations following removal of CS exposure, 3) the implication of
382 oxidative stress as the primary mechanism for CS-induced sperm DNAme
383 changes and 4) an impact of paternal CS exposure on offspring phenotype
384 based on frontal cortex DNAme and gene expression patterns. These data have
385 important implications in characterizing the potential mechanisms that underlie
386 the elevated risk observed for paternal CS exposure to offspring health. In
387 addition, the findings reported here are likely more broadly applicable to
388 understanding the risks of other environmental exposures that induce oxidative
389 stress such as air pollution and some chemical exposures.
390
391 While additional studies are necessary to fully characterize the impacts of CS
392 exposure on offspring phenotype, the mechanisms involved in heritability of pre-
393 conception paternal exposures, and the phenotypic impacts of the observed
394 epigenetic and gene expression changes in offspring, the current study
395 significantly expands our understanding of the impacts of paternal CS exposure
396 on offspring and the underlying mechanism of CS induced epigenetic changes to
397 sperm.
398
399 **MATERIALS AND METHODS:**
400 **Animals:**
401 **Animal care:**

402 All animal experiments were performed under protocols that were approved by
403 the University of Utah Institutional Animal Care and Use Committee (protocol #
404 14-11006). All animals were obtained from Jackson Laboratories (Bar Harbor,
405 ME, USA)
406
407 **Smoke exposure:**
408 All CS-exposed and control mice were age matched and smoking was initiated
409 between 6 and 7 weeks of age. Mice were exposed to CS using a Teague Model
410 TE-10 (Teague Enterprises, Woodland, CA) smoking machine, which produces a
411 combination of side-stream and mainstream CS. A pump on the machine "puffs"
412 each 3R4F University of Kentucky research cigarette for 2 seconds for a total of
413 9 puffs before ejection. The 2.5 hour daily exposure occurred for 5 consecutive
414 days per week over a period of 60 days. The smoking chamber atmosphere was
415 periodically sampled to confirm total particulate matter concentrations of
416 approximately 150 mg/m^3, the human equivalent of smoking approximately 10-20
417 cigarettes per day (47).
418
419 **Smoking and recovery experiments:**
420 To characterize the impact of smoke exposure on the sperm DNA methylome,
421 and the capacity for smoke-induced sperm DNA methylation alterations to
422 recover following removal of the insult, we exposed 40 C57BL/6J (Jackson Labs
423 Stock # 000664) to cigarette smoke for comparison against 10 age-matched,
424 non-smoked controls. Ten CS-exposed mice and the 10 non-exposed controls

425 were euthanized and tissues collected within three days of the CS exposure
426 period. Subsequent “recovery” groups of 10 CS-exposed animals were
427 euthanized 28, 103 and 171 days after the exposure period (corresponding to
428 approximately 0.8, 3 and 5 spermatogenic cycles). In addition to experiments
429 with wild-type animals, ten age-matched *Nrf2*$^{-/-}$ mice on a C57BL/6J genetic
430 background (Jackson Labs Stock # 017009) were exposed to the same doses of
431 CS for the same time period, and ten age-matched unexposed *Nrf2*$^{-/-}$ mice were
432 utilized as controls.
433
434 **Offspring transmission experiments:**
435 Founder mice for heritability experiments included wild type C57BL/6J mice
436 (Jackson Labs Stock # 000664) that were exposed and not exposed to CS (n =
437 10-12 per group). Approximately one week after the exposure period, exposed
438 and control males were introduced to 6-week old CAST/EiJ female mice
439 (Jackson Labs Stock # 000928), and pairs were kept together until F1 litters were
440 born, or for 7 weeks without conceiving, whichever came first. The motivation for
441 outcrossing males to CAST/EiJ females was to leverage polymorphic alleles to
442 enable attributing reads to a specific parent, however due to the large average
443 spacing of informative SNPs in the CAST strain and the short sequencing reads
444 inherent in Illumina sequencing we were unable to classify the large majority of
445 reads based on parent-of-origin. We therefore analyzed the data without regard
446 to parent-of-origin. F1 litters were weaned at approximately 21 days of age, and
447 pups were regularly weighed until they were euthanized. F1 animals were

448 euthanized at 14-17 weeks of age, and heart, lung, liver, kidney, brain, testis and
449 epididymal sperm were collected for molecular studies.
450
451 **Animal phenotyping:**
452 Following epididymal sperm extraction, sperm count and motility were assessed
453 in CS-exposed and control F0 animals as well as F1 offspring. In addition, time to
454 conception and litter size were compared between F0 groups. F1 offspring were
455 evaluated for growth trajectory. For statistical analysis of growth trajectories
456 between groups, animal weights were plotted against age for all pups within a
457 group (C57BL/6J or *Nrf2*$^{-/-}$). Models to fit the data were tested, and a logarithmic
458 model generally yielded the highest r^2. Theoretical weights were calculated for
459 each weight event based on the model generated, and differences between
460 theoretical and actual weight were calculated. A mean of average differences
461 within an individual across weight events was calculated for each animal, and
462 unpaired student's t-test was used to compare these differences between
463 smoked and non-smoked animals within each group. Differences in animal
464 weights and sperm parameters were evaluated using two-tailed Student's t-test,
465 and two-tailed Fisher's Exact tests were used to evaluate weekly differences in
466 conception between groups. $P < 0.05$ was considered significant.
467
468 **<u>Molecular analyses</u>:**
469 **Sperm collection and DNA extraction:**

470 Sperm was collected from the cauda epididymis and vas deferens immediately
471 after euthanasia by scoring the tissue along the length of the tubules with a 28-G
472 needle and gently pressing the tissue to expel the sperm mass. Tissues were
473 then placed in a center-well dish in equilibrated Quinn's medium supplemented
474 with FBS in a humidified CO_2 incubator for one hour. Following the swim out
475 period, sperm concentration and motility were assessed on a Makler chamber
476 and sperm were snap frozen in liquid nitrogen. Samples were subsequently
477 thawed and subjected to a stringent somatic cell lysis protocol to ensure a pure
478 population of sperm. Briefly, samples were passed through a 40 μM filter to
479 remove cell and tissue clumps followed by two 14 ml washes with ddH_2O and
480 incubation for at least 60 minutes in somatic cell lysis buffer (0.1% SDS, 0.5%
481 Triton X in ddH_2O) at 4º C. Following somatic cell lysis and visual confirmation
482 of the absence of contaminating cells, sperm DNA was extracted using the
483 Qiagen AllPrep Universal kit. Samples in cell lysis buffer were passed through a
484 28-gauge syringe multiple times to disrupt sperm membranes and liberate
485 nucleic acids prior to extraction.
486
487 **Frontal cortex dissection and nucleic acid extraction:**
488 Following euthanasia of F1 males (n = 8 per group), left brain hemispheres were
489 dissected and placed in PaxGene tissue stabilizer and after 24 hours, fixed in
490 PaxGene fixative and stored at -80º C. Samples were subsequently thawed and
491 frontal cortex dissected under a stereo microscope according to the method
492 described by Chiu et al. (48). Tissue was then disrupted using a microcentrifuge

493 pestle, and RNA and DNA were extracted using the Qiagen AllPrep Universal kit
494 according to manufacturer's protocols.
495
496 **RRBS library construction:**
497 Following DNA extraction, Bioo Scientific NEXTflex Bisulfite Library Prep Kit for
498 Illumina Sequencing was used for library preparation. To maximize coverage, we
499 employed two separate restriction digests with MspI and TaqαI. Following
500 digestion, products were pooled, and Klenow Fragment was utilized to create 3'A
501 overhangs. DNA was subsequently purified with Zymo DNA Clean and
502 Concentrate Columns followed by ligation of Methylated Illumina PE Adapters
503 and Ampure purification with SPRI beads. Purified products were Sodium
504 Bisulfite Converted using ZymoResearch EZ DNA Methylation Gold Kit, and
505 libraries were amplified over 20 cycles using Platinum Taq DNA polymerase,
506 followed by a final Ampure purification and confirmation of library size range on a
507 2% agarose gel. DNA was submitted to the Huntsman Cancer Institute High
508 Throughput genomic core for sequencing on a Hi-Seq 2500 using 50 cycle-single
509 read chemistry. Four to six samples were sequenced per lane for a minimum of
510 35-million reads per sample.
511
512 **RNAseq library construction:**
513 RNA extracted from F1 frontal cortices (n = 8 per group) was subjected to
514 Illumina TruSeq Stranded RNA kit with Ribo-Zero Gold library preparation and
515 subsequently sequenced on a Hi-Seq 2500 using 50 Cycle-Single Read

516 Sequencing v4. Eight samples were sequenced per lane for a minimum of 25-
517 million reads per sample.
518
519 **Bioinformatics analyses:**
520 For genome wide DNA methylation analysis, sequence data from RRBS libraries
521 was aligned to the mouse mm10 genome using the Bismark pipeline with special
522 attention to RRBS specific issues, as noted in the Bismark User Guide and the
523 Bismark RRBS Guide. Only CpGs where read coverage was greater than 8 for at
524 least 4 biological replicates were considered “scoreable” for downstream
525 analysis. Only CpGs with more than 5% change in methylation relative to control
526 samples were classified as differentially methylated. When considering DMRs,
527 only regions greater than 50 base-pairs in length with 3 or more scorable CpGs
528 were analyzed. Then, one third of the CpGs within each analyzed region needed
529 to be differentially methylated in order for a given region to be under
530 consideration as a DMR. Finally, qualifying regions were classified as bonafide
531 DMRs if there was more than 5% change in methylation relative to control
532 samples. For genome wide gene expression analysis, sequencing data from
533 RNASeq libraries was aligned using Novoalign. Aligned splice junction were
534 converted to genomic coordinates and low quality and non-unique reads were
535 further parsed using SamTranscriptomeParser (USeq; v8.8.8) under default
536 settings. Stranded differential expression analysis was calculated with the USeq
537 program DefinedRegionDifferentialSeq, which utilizes DESeq2 and the reference
538 mm10. Normalized read count tables were then analyzed in R, along with all

539 DNA methylation data. Integration and parsing of bed files or tables was
540 performed in R. Generation of all figures and statistical analyses was
541 accomplished using standard methods in R, with the exception of aggregate
542 histone modification profiles, which were generated using Deeptools. Gene
543 ontology analysis was performed using DAVID Functional Annotation
544 Bioinformatics Resources.

545 **REFERENCES**

546 1. Oberg M, Jaakkola MS, Woodward A, Peruga A, Pruss-Ustun A. Worldwide
547 burden of disease from exposure to second-hand smoke: a retrospective analysis of
548 data from 192 countries. Lancet. 2011;377(9760):139-46.
549 2. DiFranza JR, Aligne CA, Weitzman M. Prenatal and postnatal environmental
550 tobacco smoke exposure and children's health. Pediatrics. 2004;113(4 Suppl):1007-
551 15.
552 3. Moritsugu KP. The 2006 Report of the Surgeon General: the health
553 consequences of involuntary exposure to tobacco smoke. Am J Prev Med.
554 2007;32(6):542-3.
555 4. Laubenthal J, Zlobinskaya O, Poterlowicz K, Baumgartner A, Gdula MR,
556 Fthenou E, et al. Cigarette smoke-induced transgenerational alterations in genome
557 stability in cord blood of human F1 offspring. FASEB journal : official publication of
558 the Federation of American Societies for Experimental Biology. 2012;26(10):3946-
559 56.
560 5. Rolland M, Le Moal J, Wagner V, Royere D, De Mouzon J. Decline in semen
561 concentration and morphology in a sample of 26 609 men close to general
562 population between 1989 and 2005 in France. Human reproduction. 2012.
563 6. Priskorn L, Holmboe SA, Jacobsen R, Jensen TK, Lassen TH, Skakkebaek NE.
564 Increasing trends in childlessness in recent birth cohorts - a registry-based study of
565 the total Danish male population born from 1945 to 1980. International journal of
566 andrology. 2012;35(3):449-55.
567 7. Swan SH, Elkin EP, Fenster L. Have sperm densities declined? A reanalysis of
568 global trend data. Environmental health perspectives. 1997;105(11):1228-32.
569 8. Ford ES, Giles WH, Mokdad AH. Increasing prevalence of the metabolic
570 syndrome among u.s. Adults. Diabetes Care. 2004;27(10):2444-9.
571 9. Ford ES, Giles WH, Dietz WH. Prevalence of the metabolic syndrome among
572 US adults: findings from the third National Health and Nutrition Examination
573 Survey. JAMA : the journal of the American Medical Association. 2002;287(3):356-9.
574 10. Kiziler AR, Aydemir B, Onaran I, Alici B, Ozkara H, Gulyasar T, et al. High
575 levels of cadmium and lead in seminal fluid and blood of smoking men are

576 associated with high oxidative stress and damage in infertile subjects. Biological
577 trace element research. 2007;120(1-3):82-91.
578 11. Kulikauskas V, Blaustein D, Ablin RJ. Cigarette smoking and its possible
579 effects on sperm. Fertility and sterility. 1985;44(4):526-8.
580 12. Cupul-Uicab LA, Skjaerven R, Haug K, Melve KK, Engel SM, Longnecker MP. In
581 utero exposure to maternal tobacco smoke and subsequent obesity, hypertension,
582 and gestational diabetes among women in the MoBa cohort. Environmental health
583 perspectives. 2012;120(3):355-60.
584 13. Behl M, Rao D, Aagaard K, Davidson TL, Levin ED, Slotkin TA, et al. Evaluation
585 of the Association between Maternal Smoking, Childhood Obesity, and Metabolic
586 Disorders: A National Toxicology Program Workshop Review. Environmental health
587 perspectives. 2012.
588 14. Polyzos A, Schmid TE, Pina-Guzman B, Quintanilla-Vega B, Marchetti F.
589 Differential sensitivity of male germ cells to mainstream and sidestream tobacco
590 smoke in the mouse. Toxicology and applied pharmacology. 2009;237(3):298-305.
591 15. Fuentes A, Munoz A, Barnhart K, Arguello B, Diaz M, Pommer R. Recent
592 cigarette smoking and assisted reproductive technologies outcome. Fertility and
593 sterility. 2010;93(1):89-95.
594 16. Marchetti F, Rowan-Carroll A, Williams A, Polyzos A, Berndt-Weis ML, Yauk
595 CL. Sidestream tobacco smoke is a male germ cell mutagen. Proceedings of the
596 National Academy of Sciences of the United States of America. 2011;108(31):12811-
597 4.
598 17. Glen CD, Dubrova YE. Exposure to anticancer drugs can result in
599 transgenerational genomic instability in mice. Proceedings of the National Academy
600 of Sciences of the United States of America. 2012;109(8):2984-8.
601 18. Dubrova YE, Hickenbotham P, Glen CD, Monger K, Wong HP, Barber RC.
602 Paternal exposure to ethylnitrosourea results in transgenerational genomic
603 instability in mice. Environmental and molecular mutagenesis. 2008;49(4):308-11.
604 19. Barber RC, Dubrova YE. The offspring of irradiated parents, are they stable?
605 Mutation research. 2006;598(1-2):50-60.
606 20. Dubrova YE. Radiation-induced transgenerational instability. Oncogene.
607 2003;22(45):7087-93.
608 21. Barber R, Plumb MA, Boulton E, Roux I, Dubrova YE. Elevated mutation rates
609 in the germ line of first- and second-generation offspring of irradiated male mice.
610 Proceedings of the National Academy of Sciences of the United States of America.
611 2002;99(10):6877-82.
612 22. Secretan B, Straif K, Baan R, Grosse Y, El Ghissassi F, Bouvard V, et al. A
613 review of human carcinogens--Part E: tobacco, areca nut, alcohol, coal smoke, and
614 salted fish. Lancet Oncol. 2009;10(11):1033-4.
615 23. Jenkins TG, James ER, Alonso DF, Hoidal JR, Murphy PJ, Hotaling JM, et al.
616 Cigarette smoking significantly alters sperm DNA methylation patterns. Andrology.
617 2017.
618 24. Kohli A, Garcia MA, Miller RL, Maher C, Humblet O, Hammond SK, et al.
619 Secondhand smoke in combination with ambient air pollution exposure is
620 associated with increasedx CpG methylation and decreased expression of IFN-

621 gamma in T effector cells and Foxp3 in T regulatory cells in children. Clin
622 Epigenetics. 2012;4(1):17.
623 25. Bosse Y, Postma DS, Sin DD, Lamontagne M, Couture C, Gaudreault N, et al.
624 Molecular signature of smoking in human lung tissues. Cancer research.
625 2012;72(15):3753-63.
626 26. Word B, Lyn-Cook LE, Jr., Mwamba B, Wang H, Lyn-Cook B, Hammons G.
627 Cigarette Smoke Condensate Induces Differential Expression and Promoter
628 Methylation Profiles of Critical Genes Involved in Lung Cancer in NL-20 Lung Cells In
629 Vitro: Short-Term and Chronic Exposure. International journal of toxicology. 2012.
630 27. Breton CV, Byun HM, Wenten M, Pan F, Yang A, Gilliland FD. Prenatal tobacco
631 smoke exposure affects global and gene-specific DNA methylation. Am J Respir Crit
632 Care Med. 2009;180(5):462-7.
633 28. Perera F, Herbstman J. Prenatal environmental exposures, epigenetics, and
634 disease. Reproductive toxicology. 2011;31(3):363-73.
635 29. Joubert BR, Haberg SE, Nilsen RM, Wang X, Vollset SE, Murphy SK, et al. 450K
636 epigenome-wide scan identifies differential DNA methylation in newborns related to
637 maternal smoking during pregnancy. Environmental health perspectives.
638 2012;120(10):1425-31.
639 30. Holland ML, Lowe R, Caton PW, Gemma C, Carbajosa G, Danson AF, et al.
640 Early-life nutrition modulates the epigenetic state of specific rDNA genetic variants
641 in mice. Science. 2016;353(6298):495-8.
642 31. Brieno-Enriquez MA, Garcia-Lopez J, Cardenas DB, Guibert S, Cleroux E, Ded
643 L, et al. Exposure to endocrine disruptor induces transgenerational epigenetic
644 deregulation of microRNAs in primordial germ cells. PLoS One.
645 2015;10(4):e0124296.
646 32. Dai J, Wang Z, Xu W, Zhang M, Zhu Z, Zhao X, et al. Paternal nicotine exposure
647 defines different behavior in subsequent generation via hyper-methylation of mmu-
648 miR-15b. Sci Rep. 2017;7(1):7286.
649 33. Hawkey AB, White H, Pippen E, Greengrove E, Rezvani AH, Murphy SK, et al.
650 Paternal nicotine exposure in rats produces long-lasting neurobehavioral effects in
651 the offspring. Neurotoxicol Teratol. 2019;74:106808.
652 34. Pachenari N, Azizi H, Semnaniann S. Adolescent Morphine Exposure in Male
653 Rats Alters the Electrophysiological Properties of Locus Coeruleus Neurons of the
654 Male Offspring. Neuroscience. 2019;410:108-17.
655 35. Levin ED, Hawkey AB, Hall BJ, Cauley M, Slade S, Yazdani E, et al. Paternal
656 THC exposure in rats causes long-lasting neurobehavioral effects in the offspring.
657 Neurotoxicol Teratol. 2019;74:106806.
658 36. Venugopal R, Jaiswal AK. Nrf2 and Nrf1 in association with Jun proteins
659 regulate antioxidant response element-mediated expression and coordinated
660 induction of genes encoding detoxifying enzymes. Oncogene. 1998;17(24):3145-56.
661 37. Ishii T, Itoh K, Takahashi S, Sato H, Yanagawa T, Katoh Y, et al. Transcription
662 factor Nrf2 coordinately regulates a group of oxidative stress-inducible genes in
663 macrophages. J Biol Chem. 2000;275(21):16023-9.
664 38. Jung YH, Sauria MEG, Lyu X, Cheema MS, Ausio J, Taylor J, et al. Chromatin
665 States in Mouse Sperm Correlate with Embryonic and Adult Regulatory Landscapes.
666 Cell Rep. 2017;18(6):1366-82.

667 39. Vallaster MP, Kukreja S, Bing XY, Ngolab J, Zhao-Shea R, Gardner PD, et al.
668 Paternal nicotine exposure alters hepatic xenobiotic metabolism in offspring. Elife.
669 2017;6.
670 40. Zhang W, Yang J, Lv Y, Li S, Qiang M. Paternal benzo[a]pyrene exposure alters
671 the sperm DNA methylation levels of imprinting genes in F0 generation mice and
672 their unexposed F1-2 male offspring. Chemosphere. 2019;228:586-94.
673 41. Viluksela M, Pohjanvirta R. Multigenerational and Transgenerational Effects
674 of Dioxins. Int J Mol Sci. 2019;20(12).
675 42. McCarthy DM, Morgan TJ, Jr., Lowe SE, Williamson MJ, Spencer TJ, Biederman
676 J, et al. Nicotine exposure of male mice produces behavioral impairment in multiple
677 generations of descendants. PLoS Biol. 2018;16(10):e2006497.
678 43. Tsaprouni LG, Yang TP, Bell J, Dick KJ, Kanoni S, Nisbet J, et al. Cigarette
679 smoking reduces DNA methylation levels at multiple genomic loci but the effect is
680 partially reversible upon cessation. Epigenetics. 2014;9(10):1382-96.
681 44. Zeilinger S, Kuhnel B, Klopp N, Baurecht H, Kleinschmidt A, Gieger C, et al.
682 Tobacco smoking leads to extensive genome-wide changes in DNA methylation.
683 PLoS One. 2013;8(5):e63812.
684 45. Busche S, Shao X, Caron M, Kwan T, Allum F, Cheung WA, et al. Population
685 whole-genome bisulfite sequencing across two tissues highlights the environment
686 as the principal source of human methylome variation. Genome Biol. 2015;16:290.
687 46. Portela A, Esteller M. Epigenetic modifications and human disease. Nat
688 Biotechnol. 2010;28(10):1057-68.
689 47. Barrett EG, Wilder JA, March TH, Espindola T, Bice DE. Cigarette smoke-
690 induced airway hyperresponsiveness is not dependent on elevated immunoglobulin
691 and eosinophilic inflammation in a mouse model of allergic airway disease. Am J
692 Respir Crit Care Med. 2002;165(10):1410-8.
693 48. Chiu K, Lau WM, Lau HT, So KF, Chang RC. Micro-dissection of rat brain for
694 RNA or protein extraction from specific brain region. J Vis Exp. 2007(7):269.
695

696

697 **TABLE 1 -** Sperm characteristics, growth rates and conception data for F0 smoke-exposed

698 and control animals

	M/ml sperm	Total motility	Days to conception	# conceived w/in 1 week	# conceived w/in 2 weeks	# conceived w/in 3 weeks	# conceived w/in 4 weeks	# never conceived	Pre-smoke weight	Weight after 4 wks treatment	Weight after 8 wks treatment
CN	48.4	58.7	6.5	9.0	9.0	10.0	10.0	2.0	20.1	24.9	27.7
SN	40.2	56.5	11.8	3.0	3.0	4.0	5.0	6.0	21.4	21.1	22.7
CN vs SN p val	0.110	0.526	0.204	**0.039**	**0.039**	**0.036**	0.089	0.193	0.247	**0.001**	**<0.0001**
CWT	55.2	55.4	5.8	6.0	8.0	8.0	8.0	4.0	21.6	25.8	28.6
SWT	51.8	54.8	5.4	6.0	6.0	7.0	7.0	5.0	21.0	21.1	22.4
CWT vs SWT p val	0.618	0.868	0.908	1.000	0.608	1.000	1.000	1.000	**0.029**	**<0.0001**	**<0.0001**
NRF	44.5	57.6	8.3	12.0	12.0	14.0	15.0	8.0	20.9	23.0	25.2
WT	53.5	55.1	5.6	12.0	14.0	15.0	15.0	9.0	21.3	23.2	25.2
NRF vs WT p val	**0.039**	0.306	0.261	1.000	0.773	1.000	1.000	1.000	0.444	0.757	0.975
Smoke	46.2	55.6	8.1	9.0	9.0	11.0	12.0	11.0	21.0	21.1	22.5
No Smoke	51.8	57.0	6.2	15.0	17.0	18.0	18.0	6.0	21.4	25.6	28.4
Smoke vs No Smoke p val	0.207	0.556	0.432	0.148	**0.042**	0.075	0.135	0.135	0.221	**<0.0001**	**<0.0001**

699

700 Abbreviations: CN-control *Nrf*$^{-/-}$, SN-smoke-exposed *Nrf*$^{-/-}$, CWT-control wild type animals, SWT-

701 smoke-exposed wild type animals.

702

703

704

705 **TABLE 2 –** Sperm characteristics, litter size and growth rates for F1 offspring of smoke-

706 exposed and control animals

	Litter size	M/ml sperm	Total sperm motility	Postnatal weight (g) week 4	Postnatal weight (g) week 6	Postnatal weight (g) week 10
CN	5.3	25.7	36.1	11.6	16.4	17.4
SN	6.0	26.5	33.0	12.3	16.6	18.7
CN vs SN p val	0.379	0.851	0.409	0.351	0.727	0.100
CWT	5.0	31.3	27.3	12.1	16.2	17.4
SWT	5.2	33.6	34.0	12.7	16.8	17.8
CWT vs SWT p val	0.723	0.708	0.207	0.254	0.260	0.354
NRF	5.5	26.0	34.0	11.8	16.4	17.6
WT	5.1	32.5	30.5	12.5	16.4	17.5
NRF vs WT p val	0.356	0.063	0.264	0.194	0.979	0.828
Smoke	5.6	29.8	32.0	12.5	16.7	18.0
No Smoke	5.2	27.3	33.2	11.7	16.3	17.4
Smoke vs No Smoke p val	0.424	0.469	0.700	0.104	0.288	0.105

708 Sperm characteristics, litter size and growth rates for F1 offspring of smoke-exposed and control

709 animals. Abbreviations: CN-control *Nrf*$^{-/-}$, SN-smoke-exposed *Nrf*$^{-/-}$, CWT-control wild type

710 animals, SWT-smoke-exposed wild type animals.

711

712

713 **Figure 1 Schematic of study design.** Six to seven-week-old WT and *Nrf2*$^{-/-}$
714 mice were assigned to one of two groups: CS-exposed and non-exposed
715 controls. Following 60 days of CS exposure, mice were bred to unexposed
716 CAST/EiJ female mice. Groups of animals were euthanized and tissues collected
717 3, 28, 103, and 171 days after removal from CS exposure. Sperm DNA
718 methylation analysis was performed by RRBS on F0 exposed and control
719 animals. Offspring derived from exposed and control C57BL/6J males were
720 euthanized at 14-17 weeks of age, and tissues were collected. DNA methylation
721 analysis was performed on sperm and frontal cortex to investigate the impact of
722 paternal smoking status on methylation patterns in offspring. In addition, RNAseq
723 was performed on frontal cortex tissue to investigate the association between
724 paternal smoking and neural gene expression.
725
726 **Figure 2 Summary of the recovery, persistence or emergence of**
727 **differentially methylated loci (A and B) and regions (C and D) following a**
728 **recovery period.** A) Quantitative data indicate that the number of differentially
729 methylated loci in CS-exposed mice compared with age matched controls does
730 not diminish following a recovery period of up to 171 days. B) When considering
731 only loci that lost methylation in the CS-exposed group, about one third of
732 differentially methylated loci persisted for the entire recovery period (white), one
733 third returned to baseline levels (blue) and one third emerged as differentially
734 methylated following a recovery period (pink). Contrastingly, for loci that gained
735 methylation as a result of CS exposure, a smaller fraction of differentially

736 methylated CpGs persisted of recovered while nearly half of differentially
737 methylated loci emerged during the recovery period. C) Unlike the case of
738 recovery at the CpG level, when differentially methylated regions were analyzed,
739 a significant proportion (about half) of regions recovered within 28 days without
740 CS exposure, with no evidence for additional recovery following longer periods of
741 recovery. D) The majority of DMRs observed prior to recovery were either
742 maintained across the recovery period or returned to baseline levels, with only a
743 small fraction of DMRs emerging during the recovery period. Of note, regions
744 that gained methylation in CS-exposed animals were less likely to recover and
745 more likely to emerge as a new DMR during the recovery period compared with
746 DMRs that lost methylation. Timepoint 1 = 28-day recovery group, timepoint 2 =
747 103-day recovery group, and timepoint 3 = 171-day recovery group. All
748 comparisons were to the 3-day recovery group.
749
750 **Figure 3 DMR recovery and comparison of DMRs in WT mice with CS-**
751 **exposed and control *Nrf2*$^{-/-}$ mice as well as previously published sperm**
752 **DMRs.** A) Impact of the initial methylation status and direction of change on
753 methylation recovery. The hypomethylated DMRs (<25% DNAme) in which
754 methylation increased and the hypermethylated DMRs (> 75% DNAme) in which
755 methylation decreased with smoke exposure were more likely to recover.
756 Regions displaying methylation between 25% and 75% Regions of intermediate
757 DNAme were less likely to recover across all groups. Timepoint 1 = 28-day
758 recovery group, timepoint 2 = 103-day recovery group, and timepoint 3 = 171-day

759 recovery group. All comparisons were to the 3-day recovery group. B) Heatmap
760 illustrating the significant similarity between CS-associated DMRs identified in
761 WT mice and DMRs associated with the *Nrf2*$^{-/-}$ genotype, apparently independent
762 of CS-exposure status. C) A high correlation in DMRs was observed between
763 CS-exposed WT mice and *Nrf2*$^{-/-}$ whether or not they were exposed to CS. No
764 correlation was observed between DMRs identified in the current study
765 compared with previously published DMRs associated with vinclozolin exposure
766 (VD2) and protein restricted diet (PR).
767
768 **Figure 4 Impact of paternal CS-exposure and genotype on F1 brain and**
769 **sperm DNAme.** A) A highly significant correlation was observed in F1 brain
770 DNAme changes induced by CS exposure in WT sires (x-axis) compared with
771 DNAme changes associated *Nrf2* status irrespective of CS exposure (left 2
772 panels), however the correlation disappeared when evaluating differential
773 methylation in *Nrf2*$^{-/-}$ offspring based on CS exposure status compared with CS-
774 associated DNAme changes in offspring sired by WT mice. B) The DNAme
775 changes observed in F0 sperm were not observed in the sperm of F1 offspring,
776 suggesting the CS-associated effects likely do not confer risk transgenerationally.
777 C) Likewise, no significant overlap was observed in differentially methylated
778 regions in F0 sperm compared with differentially methylated regions in F1 frontal
779 cortex.
780

781 **Figure 5 Frontal cortex gene expression variation and correlation between**
782 **WT and *Nrf2*$^{-/-}$ offspring in relation to smoking status.** A) Variation in F1
783 frontal cortex gene expression was significantly higher in WT CS-exposed
784 animals compared with WT controls (left panel) suggesting stochastic
785 dysregulation of gene expression in paternal CS-exposed offspring and offspring
786 of mice with reduced antioxidant capacity. Likewise, variation was significantly
787 higher in *Nrf2*$^{-/-}$ controls compared with WT controls (right panel). B) A significant
788 correlation was observed between differentially expressed genes associated with
789 *Nrf2*$^{-/-}$ genotype, with (left panel) or without CS exposure (center panel) and CS-
790 exposure-induced differential gene expression. There is no association gene
791 expression impacted by smoking status in *Nrf2*$^{-/-}$ offspring (right panel). C) Gene
792 ontology analysis of significantly upregulated and D) downregulated genes
793 associated with paternal CS exposure indicated overrepresentation of several
794 gene families.
795
796 **Figure S1 Scatter plots of CpG-level and regional sperm DNAme in CS-**
797 **exposed vs control WT mice across the recovery timeline.** A) DNAme at all
798 CpGs across the genome B) Regional DNAme comparisons. Broad genome-
799 wide changes in DNAme, either at individual CpGs of within genomic regions, are
800 not observed. Further, the methylation at individual CpGs deviates much more
801 broadly than regional DNAme, supporting the role of increased CpG density in
802 buffering environmentally-induced DNAme. Smoked WT = samples collected 3

803 days after removal from CS, recovery 1 = 28-day recovery group, recovery 2 =
804 103-day recovery group, and recovery 3 = 171-day recovery group.
805
806 **Figure S2 Scatter plots of regional sperm DNAme for control and smoked**
807 **NRF versus control and smoked WT.** Like the distributions observed in figure
808 S1, these plots indicate that CS exposure and *Nrf*$^{-/-}$ do not induce broad genome-
809 wide changes in sperm DNA methylation.
810
811 **Figure S3 DNAme variation in shared, recovered and new DMRs as a**
812 **function of CpG density across recovery timepoints.** A) In every category of
813 DMR (shared, recovery or new) variation diminished as CpG density of a region
814 increased. The impact of CpG density on variation was particularly apparent for
815 regions in which DNAme decreased as a result of CS exposure and later
816 recovered to baseline. B) DMRs that displayed increased DNAme in CS-exposed
817 animals and later recovered were generally regions of lower CpG density, while
818 DMRs that gained methylation and subsequently recovered were generally at
819 regions of higher CpG density. Timepoint 1 = 28-day recovery group, timepoint 2
820 = 103-day recovery group, and timepoint 3 = 171-day recovery group. All
821 comparisons were to the 3-day recovery group.
822
823 **Figure S4 Chromatin properties at F0 sperm DMRs and variation in F1**
824 **frontal cortex gene expression.** A) Chromatin accessibility was not predictive
825 of the propensity of DMRs to recover or be maintained after removal of CS

826 exposure. B and C) Localization of H3K4me3 and H3K27me3 was likewise not
827 associated with DMR recovery or maintenance. D) Scatter plot of F1 frontal
828 cortex variation in gene expression in CS-exposed versus control *Nrf*$^{-/-}$ mice,
829 demonstrating that paternal CS exposure did not further elevate the increased
830 gene expression variation in *Nrf*$^{-/-}$ offspring beyond the effects of genotype alone
831 illustrated in Figure 5A. E) Scatter plot of F1 frontal cortex gene expression
832 variation in CS-exposed WT offspring versus control *Nrf*$^{-/-}$ demonstrating that
833 control *Nrf*$^{-/-}$ offspring exhibit a similar degree of variation as CS-exposed WT
834 animals.
835

LIST OF ABBREVIATIONS: DMR-differentially methylated regions; DNAme-DNA methylation; CS-cigarette smoke; CN-control *Nrf2*$^{-/-}$; SN-smoke-exposed *Nrf2*$^{-/-}$, CWT-control wild type animals; SWT-smoke-exposed wild type; VD2-vinclozolin exposure; PR-protein restricted diet; ESTR-expanded simple tandem repeat; NRF2- nuclear factor (erythroid-derived 2)-like 2; GO-gene ontology

Ethics approval and consent to participate: All animal experiments were performed under protocols that were approved by the University of Utah Institutional Animal Care and Use Committee (protocol # 14-11006).

Consent for publication: Not applicable

Availability of data and material: The datasets used and/or analyzed during the current study have been deposited in NCBI's Gene Expression Omnibus
are publicly available GEO Submission (GSE133742) [NCBI tracking system and are accessible through GEO Series accession number GSE133742 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE133742).
Sperm DNA methylation datasets used for comparison to data generated in the current study are likewise available through GEO: GSE7261 and GSE65784.

Competing interests: The authors declare that they have no competing interests.

Funding: This work was supported by a grant from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, USA (R01HD082062).

Authors' contributions: PJM directed RRBS experiments, analyzed RRBS and RNAseq datasets and assisted in manuscript preparation. JG analyzed RRBS and RNAseq datasets, prepared figures and assisted in manuscript preparation. TGJ assisted in study design and animal experiments as well as data analysis and interpretation. JRH assisted with initial study design, and data interpretation, TH assisted with animal experiments including CS treatments and animal phenotyping. JMH assisted with study design and manuscript preparation. DTC assisted with initial study design, data interpretation and manuscript preparation. BRC assisted with study design, data analysis and interpretation and manuscript preparation. KIA assisted in study design, animal experiments, molecular experiments and data analysis and took the primary role in manuscript preparation. All authors read and approved the final manuscript.

Acknowledgements: Not applicable

Figure 1

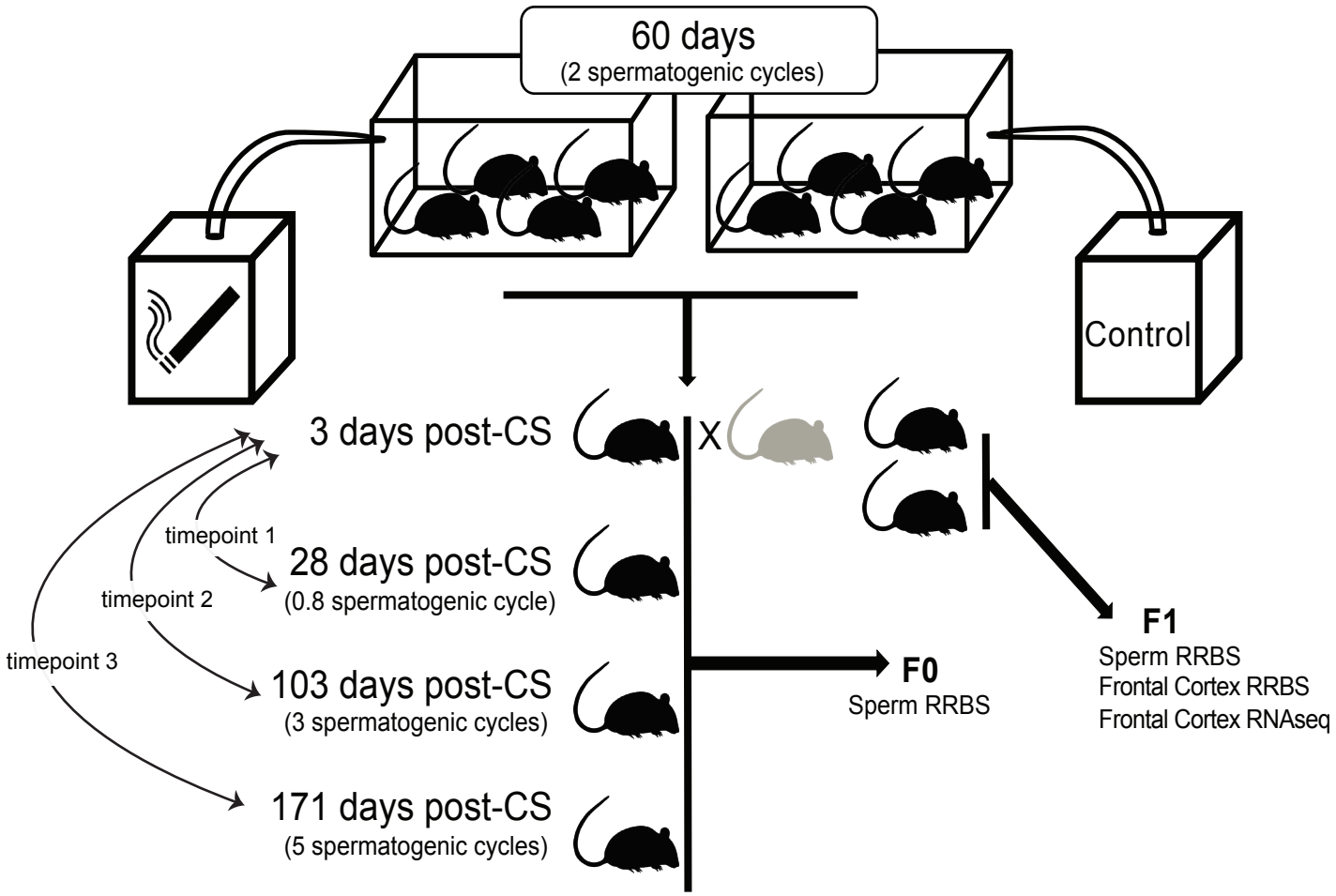

60 days
(2 spermatogenic cycles)
Control
3 days post-CS
X
timepoint 1
28 days post-CS
(0.8 spermatogenic cycle)
timepoint 2
timepoint 3
103 days post-CS
(3 spermatogenic cycles)
171 days post-CS
(5 spermatogenic cycles)
F0
Sperm RRBS
F1
Sperm RRBS
Frontal Cortex RRBS
Frontal Cortex RNAseq

Figure 2

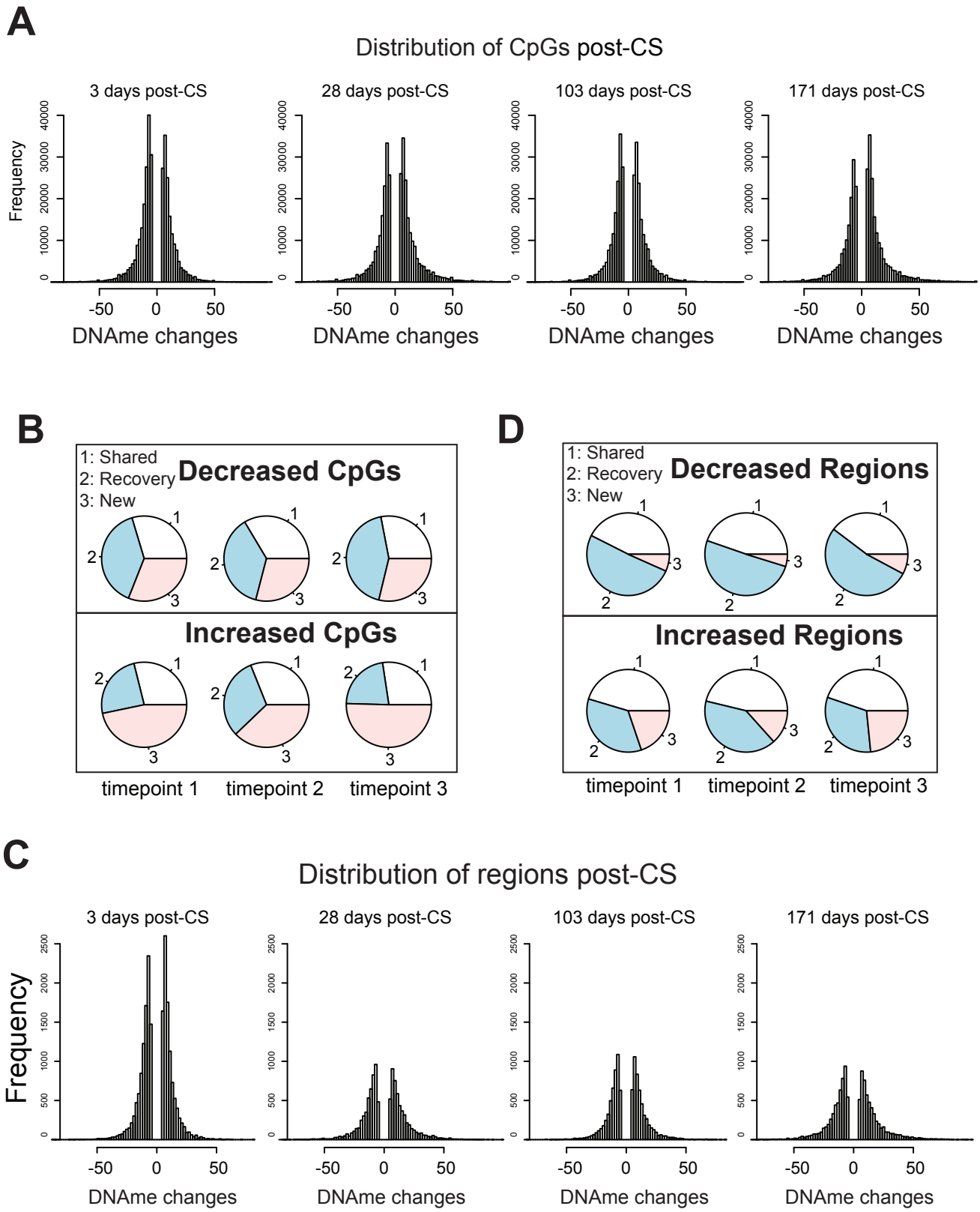

A
Distribution of CpGs post-CS
3 days post-CS
28 days post-CS
103 days post-CS
171 days post-CS
Frequency
DNAme changes
B
1: Shared
2: Recovery
3: New
Decreased CpGs
Increased CpGs
timepoint 1
timepoint 2
timepoint 3
D
Decreased Regions
Increased Regions
C
Distribution of regions post-CS

Figure 3

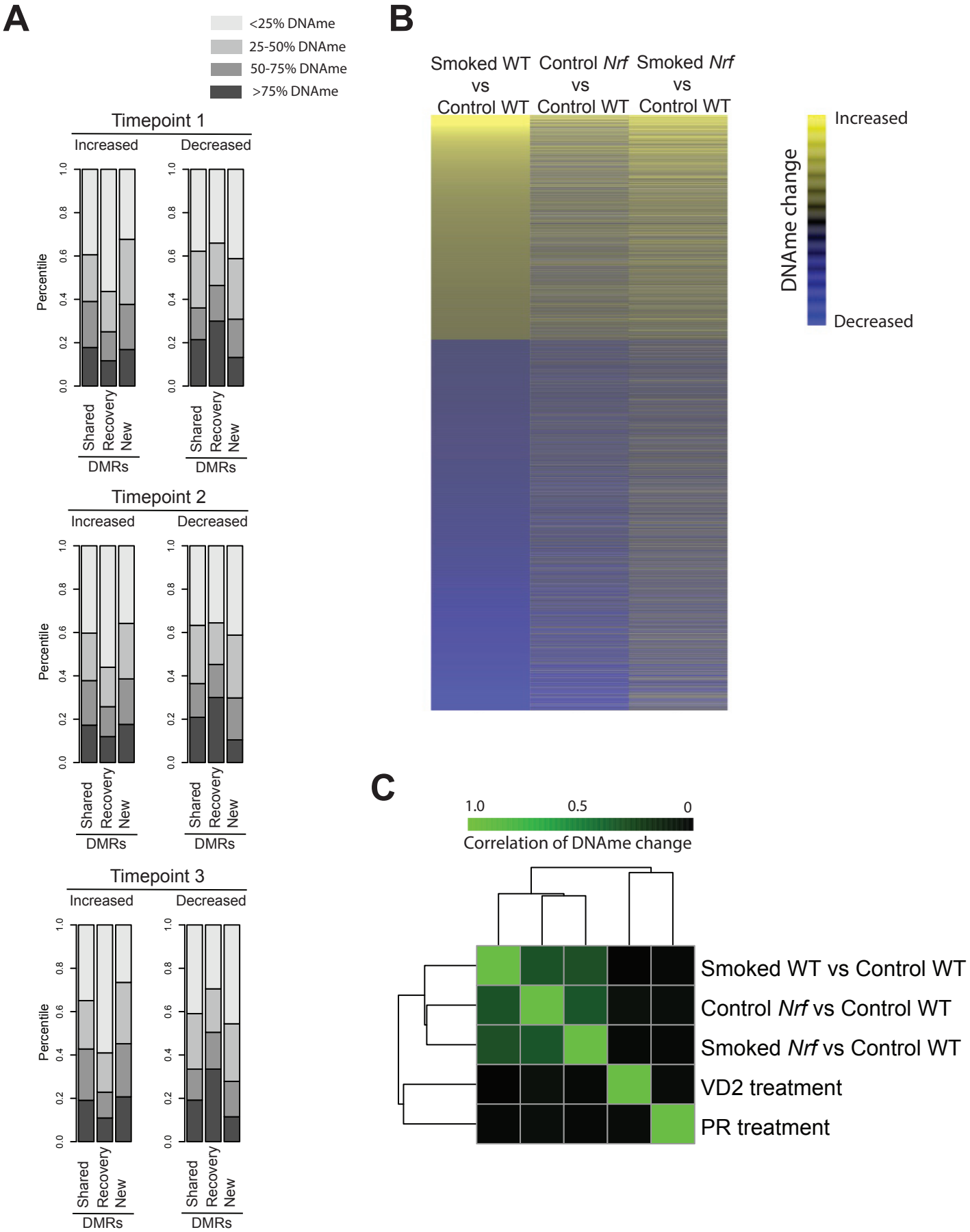

A
<25% DNAme
25-50% DNAme
50-75% DNAme
>75% DNAme
Timepoint 1
Increased
Decreased
Percentile
Shared
Recovery
New
DMRs
Timepoint 2
Timepoint 3
B
Smoked WT vs Control WT
Control Nrf vs Control WT
Smoked Nrf vs Control WT
DNAme change
Increased
Decreased
C
1.0
0.5
0
Correlation of DNAme change
Smoked WT vs Control WT
Control Nrf vs Control WT
Smoked Nrf vs Control WT
VD2 treatment
PR treatment

Figure 4

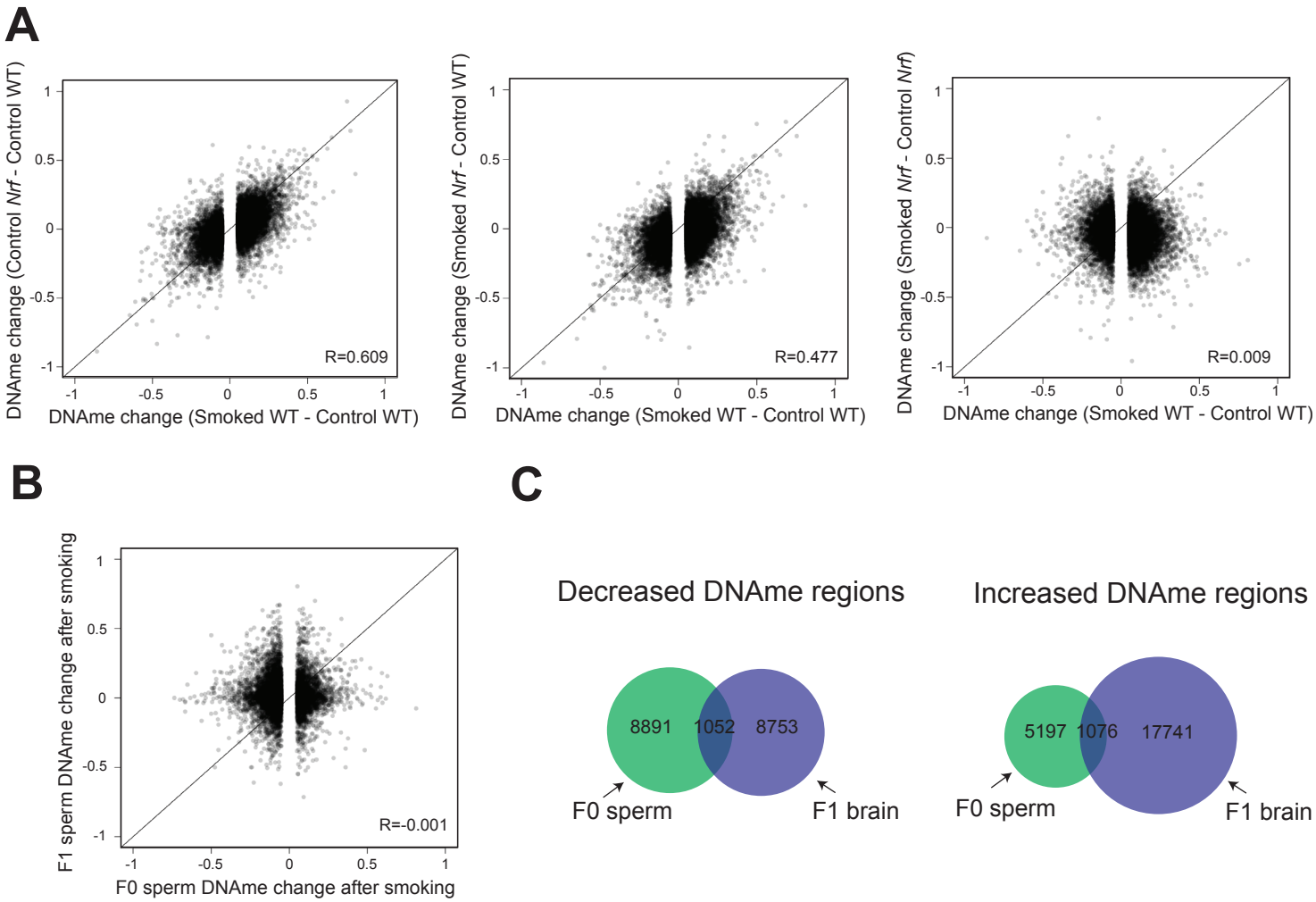

A
DNAme change (Control Nrf - Control WT)
DNAme change (Smoked WT - Control WT)
R=0.609
DNAme change (Smoked Nrf - Control WT)
DNAme change (Smoked WT - Control WT)
R=0.477
DNAme change (Smoked Nrf - Control Nrf)
DNAme change (Smoked WT - Control WT)
R=0.009
B
F1 sperm DNAme change after smoking
F0 sperm DNAme change after smoking
R=-0.001
C
Decreased DNAme regions
8891
1052
8753
F0 sperm
F1 brain
Increased DNAme regions
5197
1076
17741
F0 sperm
F1 brain

Figure 5

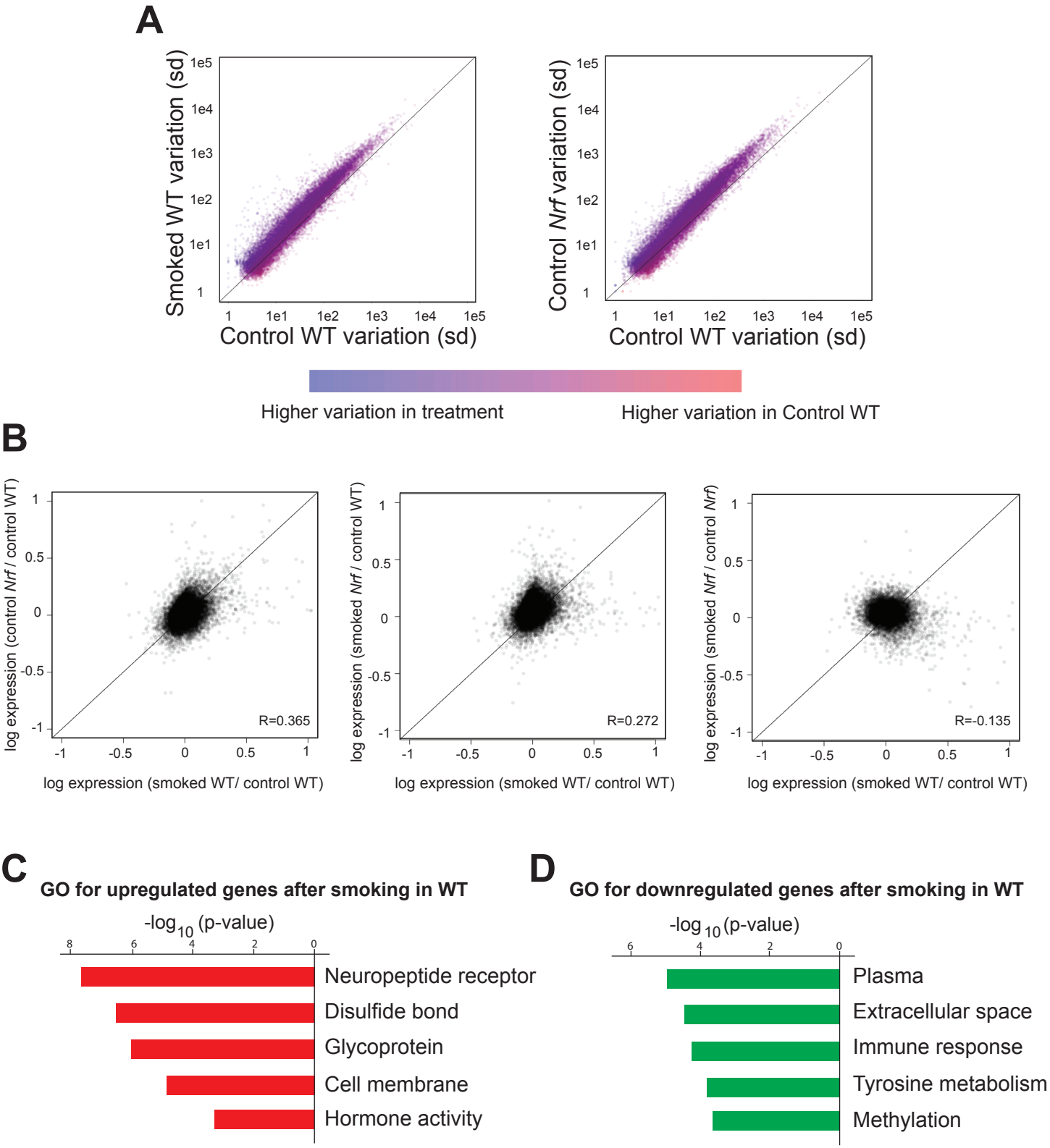

A
Smoked WT variation (sd)
Control WT variation (sd)
Control Nrf variation (sd)
Control WT variation (sd)
1
1e1
1e2
1e3
1e4
1e5
Higher variation in treatment
Higher variation in Control WT
B
log expression (control Nrf / control WT)
log expression (smoked WT/ control WT)
R=0.365
log expression (smoked Nrf / control WT)
log expression (smoked WT/ control WT)
R=0.272
log expression (smoked Nrf / control Nrf)
log expression (smoked WT/ control WT)
R=-0.135
C
GO for upregulated genes after smoking in WT
-log10 (p-value)
Neuropeptide receptor
Disulfide bond
Glycoprotein
Cell membrane
Hormone activity
D
GO for downregulated genes after smoking in WT
-log10 (p-value)
Plasma
Extracellular space
Immune response
Tyrosine metabolism
Methylation

Figure S1

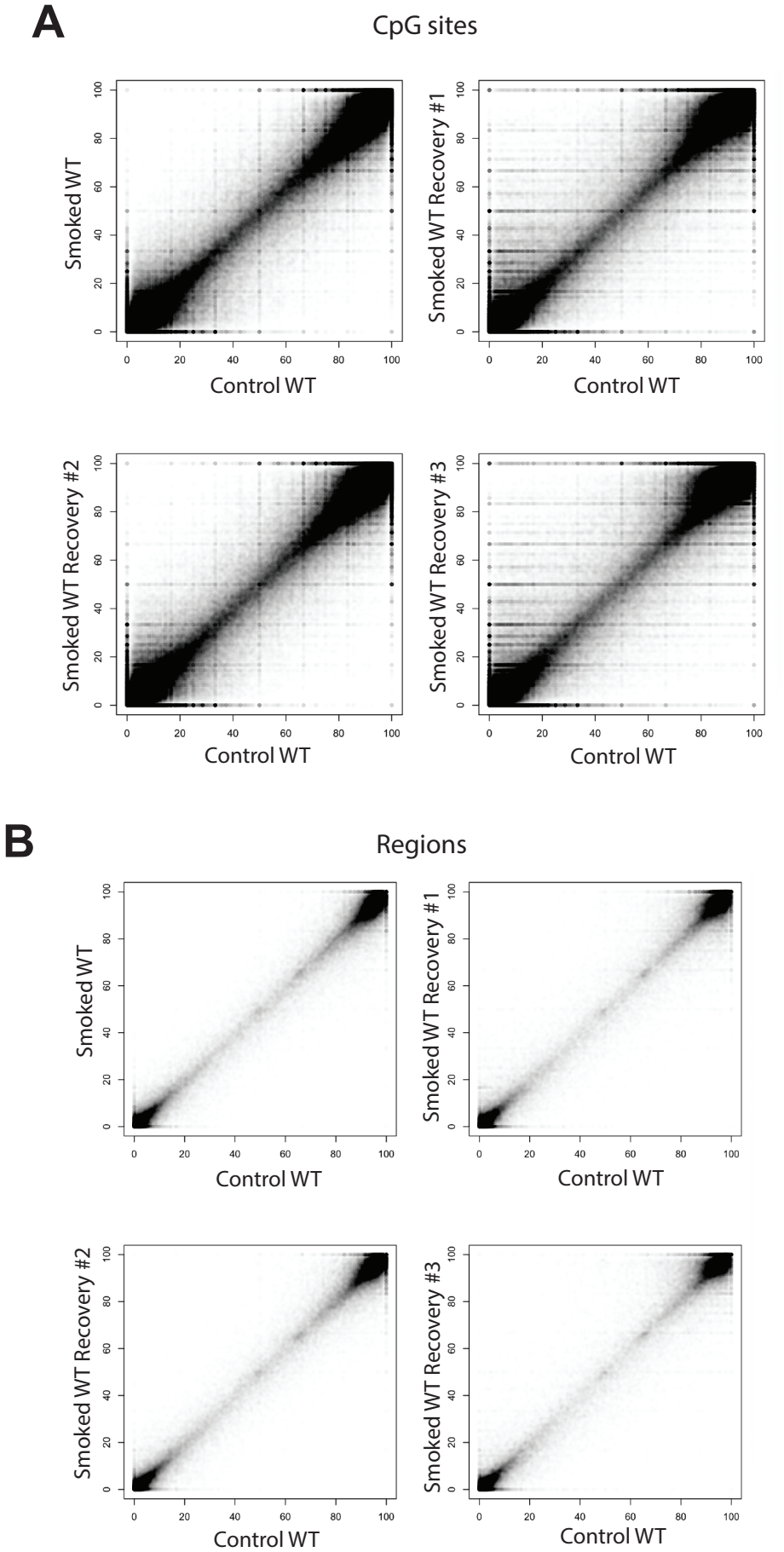

A
CpG sites
Smoked WT
Smoked WT Recovery #1
Smoked WT Recovery #2
Smoked WT Recovery #3
Control WT
B
Regions
Smoked WT
Smoked WT Recovery #1
Smoked WT Recovery #2
Smoked WT Recovery #3
Control WT

Figure S2

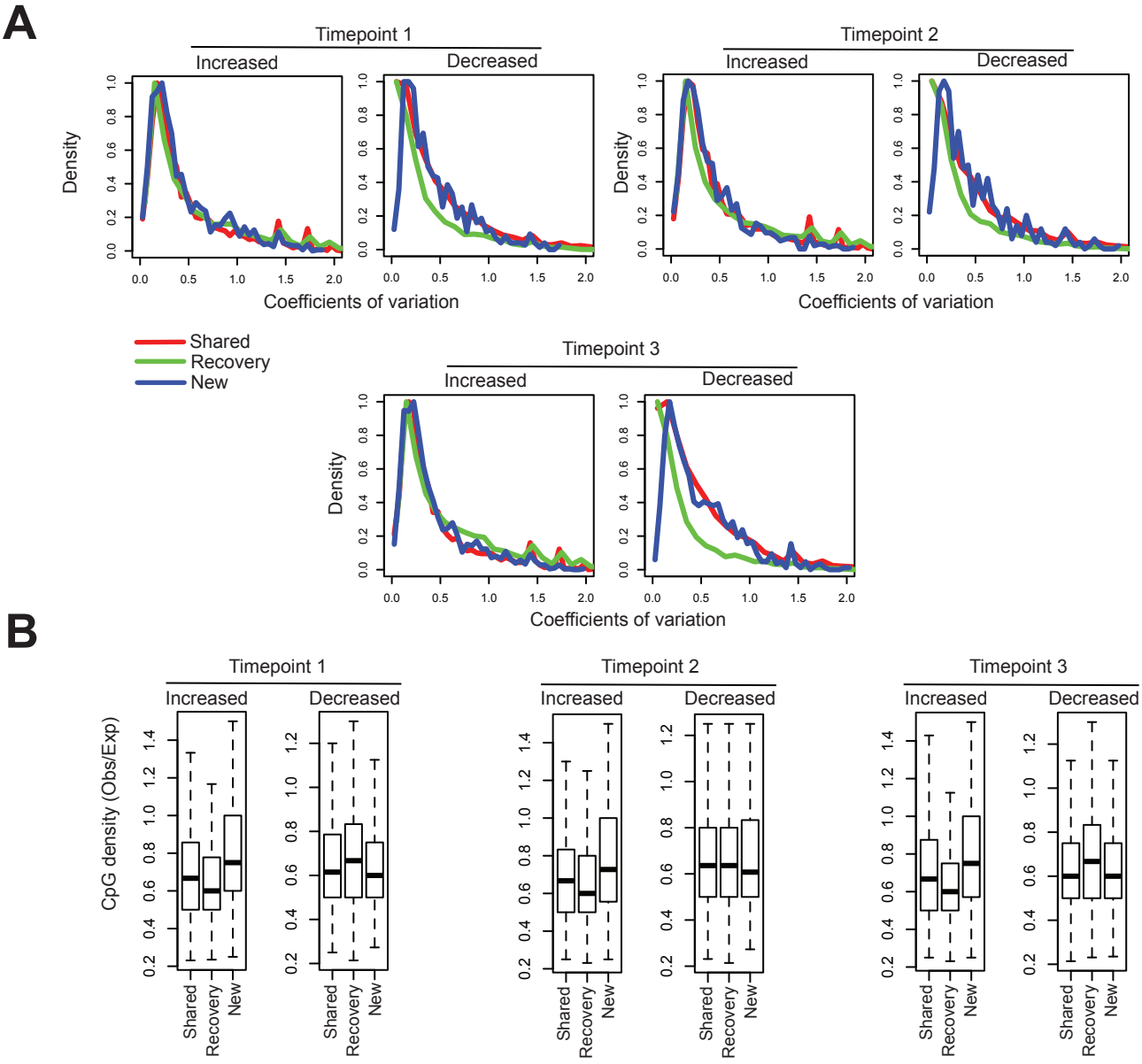

A
Timepoint 1
Increased
Decreased
Timepoint 2
Increased
Decreased
Density
Coefficients of variation
Shared
Recovery
New
Timepoint 3
Increased
Decreased
Density
Coefficients of variation
B
Timepoint 1
Increased
Decreased
Timepoint 2
Increased
Decreased
Timepoint 3
Increased
Decreased
CpG density (Obs/Exp)
Shared
Recovery
New

Figure S3

A

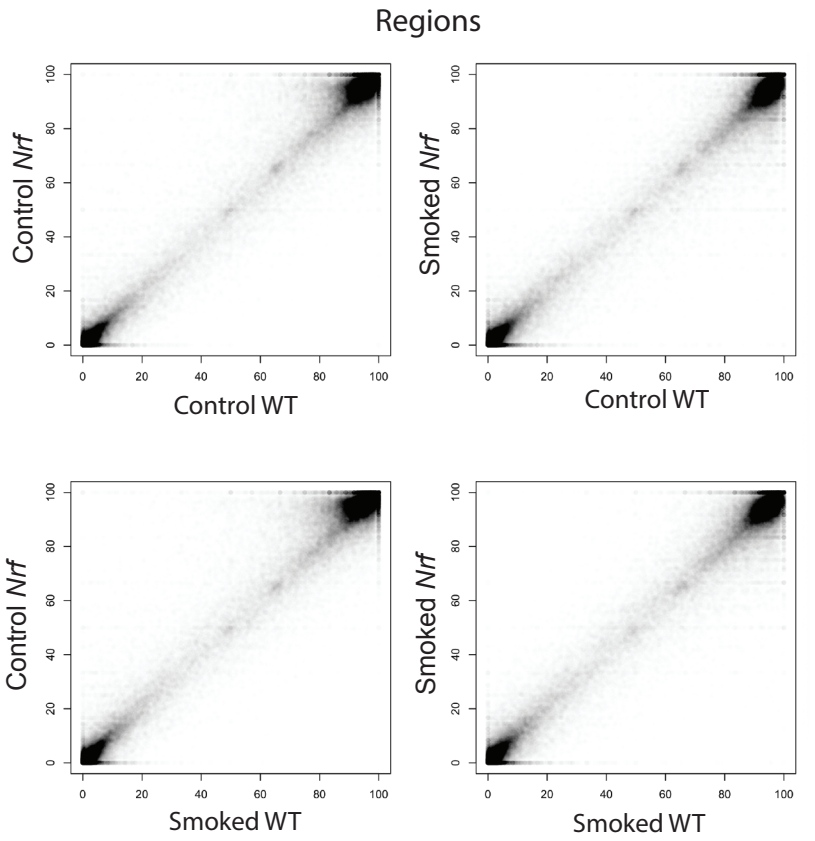
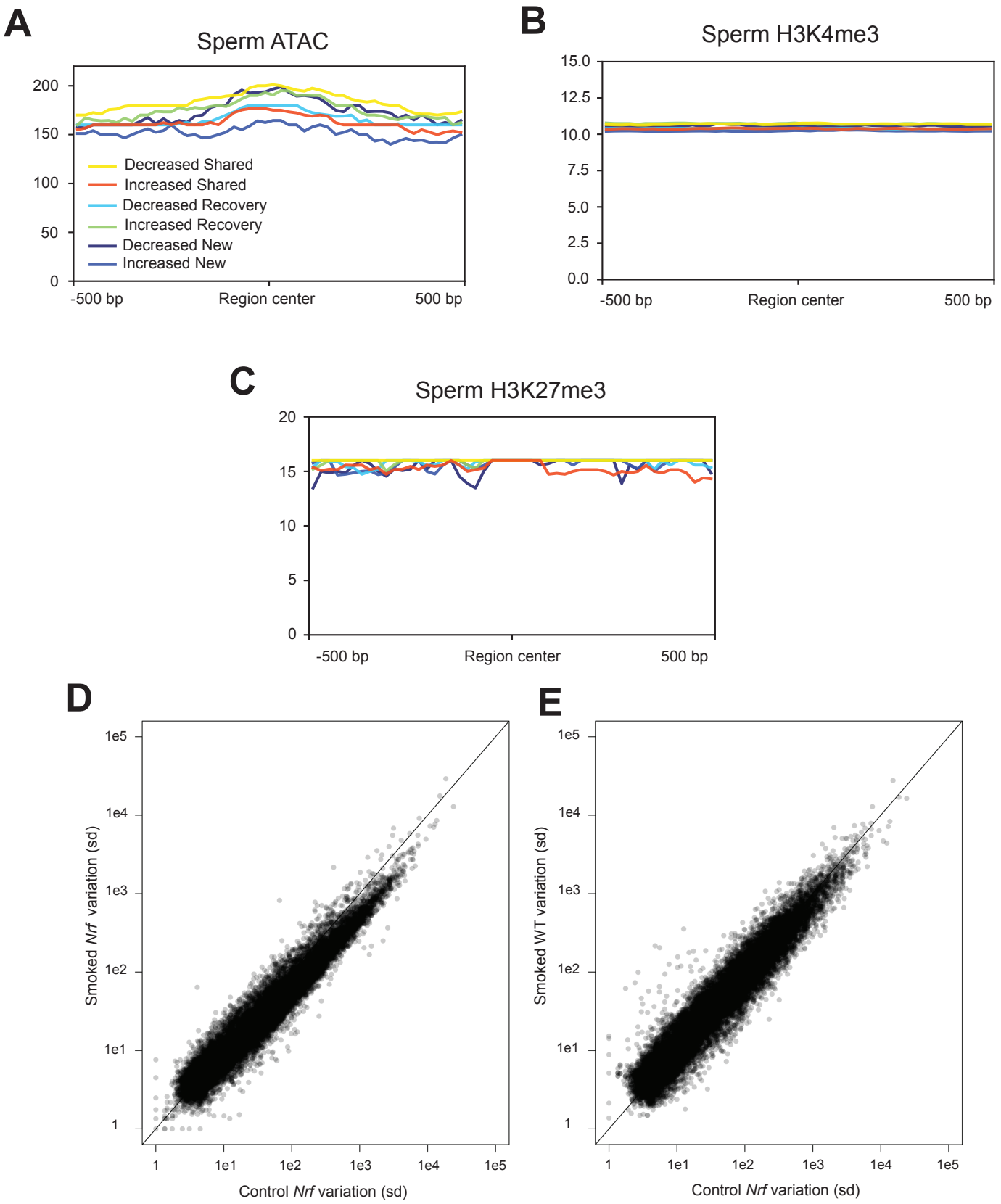

Regions
Control *Nrf*
Control WT
Smoked *Nrf*
Control WT
Control *Nrf*
Smoked WT
Smoked *Nrf*
Smoked WT
0
20
40
60
80
100

Figure S4


A
Sperm ATAC
200
150
100
50
0
Decreased Shared
Increased Shared
Decreased Recovery
Increased Recovery
Decreased New
Increased New
-500 bp
Region center
500 bp
B
Sperm H3K4me3
15.0
12.5
10.0
7.5
5.0
2.5
0.0
-500 bp
Region center
500 bp
C
Sperm H3K27me3
20
15
10
5
0
-500 bp
Region center
500 bp
D
Smoked Nrf variation (sd)
Control Nrf variation (sd)
1e5
1e4
1e3
1e2
1e1
1
E
Smoked WT variation (sd)
Control Nrf variation (sd)
1e5
1e4
1e3
1e2
1e1
1