Sperm morphology differences associated with pig fertility

Mandawala AA[1], Skinner BM[2], Walling GA[3], Fowler KE[1*] and Harvey SC[1*]

[1] School of Human and Life Sciences, Canterbury Christ Church University, Canterbury, UK

[2] University of Cambridge, Cambridge, UK

[3] JSR Genetics, Southburn, Driffield, East Yorkshire, UK

[*] Fowler KE and Harvey SC are joint last authors

Corresponding author:

Dr. Simon Harvey,

School of Human and Life Sciences,

Canterbury Christ Church University,

Canterbury,

CT1 1QU,

UK.

Email: simon.harvey@canterbury.ac.uk

Telephone: +44 1227 782327

Short title:

Pig sperm morphology analysis

Abstract

Artificial insemination is routinely used in commercial pig breeding, for which the use of high quality semen samples is imperative. Currently, semen quality is determined manually by morphological assessment. This method leads to high inter-operator variability due to its subjective nature. The development of a semi-automated software-based approach to assess sperm morphology would enable faster identification of morphological defects and permit identification of subtle differences that may affect fertilisation success. Here we have used a novel method to comprehensively analyse pig sperm nuclear morphology in greater detail than was previously possible. Semen samples from 50 fertile and 50 sub-fertile samples that had been previously manually categorised as fertile or sub-fertile were analysed using this new method, with at least 200 fixed and DAPI (4',6-diamidino-2-phenylindole) stained sperm heads imaged per sample. Differences in sperm nuclear morphology were observed between fertile and sub-fertile samples; specifically, fertile samples were associated with higher mean nuclear area, a consequence of a greater head width and a lower variability between sperm heads. This novel, unbiased and fast analysis method demonstrates a significant difference in sperm head morphology between fertile and sub-fertile animals, and has the potential to be further developed and used as a tool for sperm morphology assessment in the pig breeding industry.

Keywords: Morphology, morphometrics, nucleus, pig, sperm

Introduction

Male fertility is a consequence of both the number and quality of sperm (Cooper *et al.*, 2009). In humans, many issues with male infertility are addressed using one of a range of assisted reproductive technology (ART) procedures, such as *in vitro* fertilisation (IVF) and intracytoplasmic sperm injection (ICSI). In agriculturally significant species (for example, pigs, cattle and sheep), where a key goal is to maximise the production of meat at a low cost, male fertility is also a challenge (Tardif *et al.*, 1999). To this end, improving reproductive traits is of paramount importance. In such species, the critical aim is often to identify sub-fertile animals quickly and cheaply so they can be removed from breeding schemes; boars from a nucleus herd with fertility problems have the potential to reduce litter sizes throughout the breeding population (O'Connor et al., 2017).

Artificial insemination (AI), is the oldest (Roca *et al.*, 2006) and most routinely used technique in commercial animal breeding, especially in livestock species (Dziuk and Henshaw, 1958; Polge *et al.*, 1968; Johnson *et al.*, 1981; Singleton, 2001; Gerrits *et al.*, 2005; Roca *et al.*, 2006; Feitsma, 2009). Over the past three decades, the use of AI has benefited the pig breeding industry, particularly in Europe where over 80% of sows are bred through AI (Roca *et al.*, 2006). In North America the technique is also widely implemented, especially in large farming units (Gerrits *et al.*, 2005). The principal objective of AI in the pig breeding industry is to permit the dissemination of genetics from high genetic merit boars to as many sows as possible. Without AI more boars would be needed and hence animals of lower genetic merit would be required in breeding programmes. Moreover, the technique enables the opportunity to introduce superior genetic traits into sows whilst reducing the incidence of disease transmission, an advantage that does not exist with natural mating (Maes *et al.*, 2008). AI is achieved by depositing spermatozoa into the female genital tract using artificial devices and processes. The standardised method of insemination is the intra-cervical insemination technique, with the semen dose deposited in the posterior region of cervical canal (Roca *et al.*, 2006).

In humans, semen analysis is widely used to evaluate male fertility in infertile couples (Rowe *et al.*, 1993) and may also be used for the determination of reproductive toxicity in therapeutic and environmental agents (Apostoli *et al.*, 1998; U.S. Environmental Protection Agency, 1996). Various physical characteristics of semen are assessed and whilst parameters such as volume, appearance, pH and viscosity are considered important (Maree *et al.*, 2010), several studies have shown that sperm morphology is critical when determining semen quality and hence quantifying male fertility (MacLeod and Gold, 1951; Hartman *et al.*, 1964; Eliasson, 1971; Menkveld and Kruger, 1996; Coetzee *et al.*, 1998; Auger *et al.*, 2016). Generally, the cut-off values of what is considered 'normal' vary and are dependent on the fertility clinic. However, the following benchmarks were published in the World Health Organisation's 5th edition of "normal semen analysis": morphology (≥4% normal forms), total motility (≥40%), vitality (≥58% live), sperm concentration (≥15,000,000 per mL) and volume (≥1.5 mL) (Rowe *et al.*, 1993; WHO, 2010). To date, a number of studies have been performed to analyse semen composition (Huggins *et al.*, 1942; Owen, 2005) and to establish the relationship between sperm quality and fertility in men (Paz *et al.*, 1977; Overstreet and Katz, 1987; Martin and Rademaker, 1988; Perreault *et al.*, 2003; Jung and Schuppe, 2007; Agarwal *et al.*, 2008). One such study by Guzick and colleagues used several comparative semen analyses of fertile and infertile men, to determine the most appropriate measurements that could be used in the determination of fertility potential in men (Guzick *et al.*, 2001). Here, it was established that whilst threshold values for sperm motility, concentration and morphology could be used in the classification of males into fertile, indeterminate fertility, or sub-fertile categories, these measures cannot be used independently for the diagnosis of male infertility (Guzick *et al.*, 2001).

In livestock species, a key contribution to successful fertilisation following AI is also the use of high quality semen during insemination. As such, routine assessment of semen quality is a standard process in the animal breeding industry (López Rodríguez *et al.*, 2013). A issue is however that the mammalian ejaculate does not contain a homogenous population of spermatozoa (Holt and Van Look, 2004); in

general analysis of boar semen, semen quality is considered "normal" if the frequency of abnormal sperm heads does not exceed 10%. "Normality" can also be assumed if the frequency of abnormalities in acrosomes, mid-pieces, tails or proximal cytoplasmic droplets is less than either 5% each or 15% when combined (Saravia *et al.*, 2007). Whilst some abnormalities such as the existence of distal cytoplasmic droplets are considered irrelevant to fertility assessment, other abnormalities are regarded as serious defects due to their ability to hinder fertilisation and cause infertility (Saravia *et al.*, 2007). Some of these include decapitated spermatozoa, acrosomal plicae (knobbed defect), nuclear vacuoles (diadem defect), short tails (tail stump), coiled tails (dag defect) and corkscrewed mid-pieces (Saravia *et al.*, 2007).

As well as studying the morphology of whole spermatozoa, a limited number of studies have focused on analysis of sperm nuclear morphometry in agricultural animals (Vicente-Fiel *et al.*, 2013a, b). A number of measurements of sperm nuclei have been compared in cattle (*Bos taurus taurus)*, sheep (*Ovis orientalis aries*), goats (*Capra aegagrus hircus)* and pigs (*Sus scrofa domesticus)* using computer-assisted sperm morphometry analysis-F (CASMA-F) (Vicente-Fiel *et al.*, 2013a, b). Whilst mean numerical values for area, perimeter, length and width were identified for each of the four species studied, the key findings of this study were that drying and fixation only has a minimal effect on sperm nuclear morphometry and that variations between morphometric parameters do exist between the sperm nuclei of each species (Vicente-Fiel *et al.*, 2013a). A second study (Vicente-Fiel *et al.*, 2013b) used a combination of the CASMA-F method and multivariate cluster analysis to identify subpopulations of spermatozoa within the same four species. Based on these nuclear morphometrics, three subpopulations, namely, large, small-elongated and small-round were identified. Whilst it had previously been shown that sperm shape differed between high and low fertility bulls (Ostermeier *et al.*, 2001), such observations have not been made in pigs. Currently, manual morphological assessment of a semen sample requires the observation of at least 500 sperm heads per sample (Saravia *et al.*, 2007) which is laborious and may lead to high inter-operator variability due to the

124 subjective nature of this assessment method. Here, we demonstrate the use of a semi-automated
125 software based approach to assess sperm head morphometrics in both fertile and sub-fertile pigs.
126

127 ## Materials and methods

128 ### *Semen collection*

129 Fresh ejaculated sperm samples from boars of various breeds including Large White, Landrace, White
130 Duroc, Hampshire and Pietrain were collected using the 'gloved hand method' (King and Macpherson,
131 1973), by trained staff at JSR Genetics Ltd.. Samples were stored in Duragen extender, supplemented
132 with no less than: 500 IU per ml streptomycin; 500 IU per ml penicillin; 150 mg per ml lincomycin; and
133 300 mg per ml spectinomycin, diluted to 2.3 billion sperm per dose. Samples were stored at 17°C and
134 were prepared within two days following collection.
135

136 ### *Sample preparation*

137 Prior to preparation of samples for this study, semen samples were identified as either fertile or sub-
138 fertile using a combination of computer assisted sperm analysis (CASA), followed by manual
139 assessment. Specifically, samples that had a normal morphology score of above 70% (obtained from
140 CASA) and a motility score of above 4 (motility was graded from 1 to 5, 1 being dead and 5 being
141 excellent) (subjective manual assessment) were graded as fertile and those falling below these criteria
142 were graded as sub-fertile.
143

144 50 fertile and 50 sub-fertile samples were used in this study. 2mL of each semen sample was
145 centrifuged at 300g for 5 minutes at 17°C. The supernatant was discarded, the pellet was re-
146 suspended in 1.5mL of fixative solution (100% methanol and 100% acetic acid, added dropwise at a
147 3:1 ratio) and centrifuged at 300g for 5 minutes at 17°C. The supernatant was discarded and the pellet
148 was re-suspended in 1.5mL of fixative solution. 10µL of each sample was then dropped onto the centre
149 of the surface of a labelled (sample ID, date), steam-warmed slide, immediately followed by 10µL of

fixative solution. Subsequently, slides were air-dried for 2 minutes before 1 drop of fluorescent DAPI (4',6-diamidino-2-phenylindole) was added to the centre of the slide. Prepared slides were air-dried in the dark, for at least 20 minutes prior to microscopy.

Image acquisition

An Olympus IX83 inverted fluorescence microscope equipped with CellSens Dimension version 1.9 (expandable imaging software for Life Science microscopy) was used for image capturing. A minimum of 200 nuclei were imaged (at 1000x magnification) per sample.

Data analysis

Images were analysed using the ImageJ plugin 'Nuclear Morphology Analysis' (see Skinner *et al.,* 2018, freely available under the GNU General Public License version 3 from https://bitbucket.org/bmskinner/nuclear_morphology/wiki/Home). The analysis software enables automated recognition of round or asymmetric nuclei within an image of interest, and subsequent morphological analysis of these nuclei. Initially developed for the analysis of mouse sperm (Skinner *et al.,* 2018), we adapted the feature recognition to analyse pig sperm. The software generates a range of measures; in this study, sperm heads were measured for: Area; Perimeter; Bounding Height and Bounding Width, the dimensions of the rectangle perfectly enclosing the nucleus when vertically oriented; Circularity, a measure between 0 and 1 indicating how circular the heads are, with 1 indicating a perfect circle; Minimum Width across the centre of mass; Aspect, the ratio produced from height divided by width; and Variability, a per-nucleus measure calculated as the square root of the sum-of-squares difference at each index between the nuclear profile and the dataset median profile, normalised to the length of the median profile. Analysis was carried out using software version 1.13.5. Initial correlation analysis was used to identify redundant variables. Non-redundant variables were then further analysed. For initial comparisons between fertile and sub-fertile animals, means of each

variable were compared by ANOVA in R version 3.4.4 (R Core Team, 2018), with the normality of residual distribution assessed using scatter and quantile-quantile plots.

For the cluster analysis, samples were randomly allocated into two groups; a training group of 30 fertile and 30 sub-fertile animals and a test group of 20 fertile and 20 sub-fertile animals. Sperm heads from the training group were used to determine if there were sperm of detectably different types. This was done by cluster analysis, with Ward linkage, squared Euclidean distance and standardised variables in Mintab v17. The representation of sperm heads from fertile and sub-fertile animals within in each identified cluster was then compared to that expected by chance. The training group was then used to investigate if identified clusters could be used to predict fertility status. This was done by using the mean cluster value from the training set to define the starting partitions for a K-means clustering.

Results

Pig sperm are (mostly) symmetrical about the anterior-posterior axis. Without a distinctive hook, as in mouse sperm, to act as a reference point, the tail attachment point was chosen to anchor the angle profiles generated and orient nuclei. The tail attachment point is characterised by a ‘dimple’ in the nucleus (Figure 1A, point 1). For consistent alignment of the nuclei, we placed the tail attachment region directly below the centre of mass of the nucleus. The software output prior to statistical analyses is summarised in Figure 1. This indicates how we generate an angle profile for a given sample and how a consensus sperm head shape for the population is produced. These refolded sperm heads are then used both to trace any identified abnormalities on a particular segment of the refolded nucleus back to the angle profile, and to calculate a range of sperm head parameters for analysis.

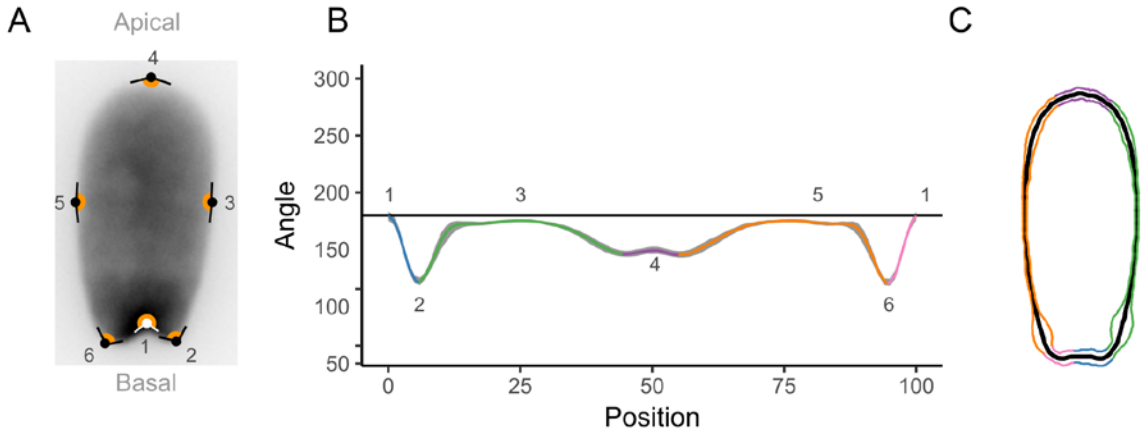
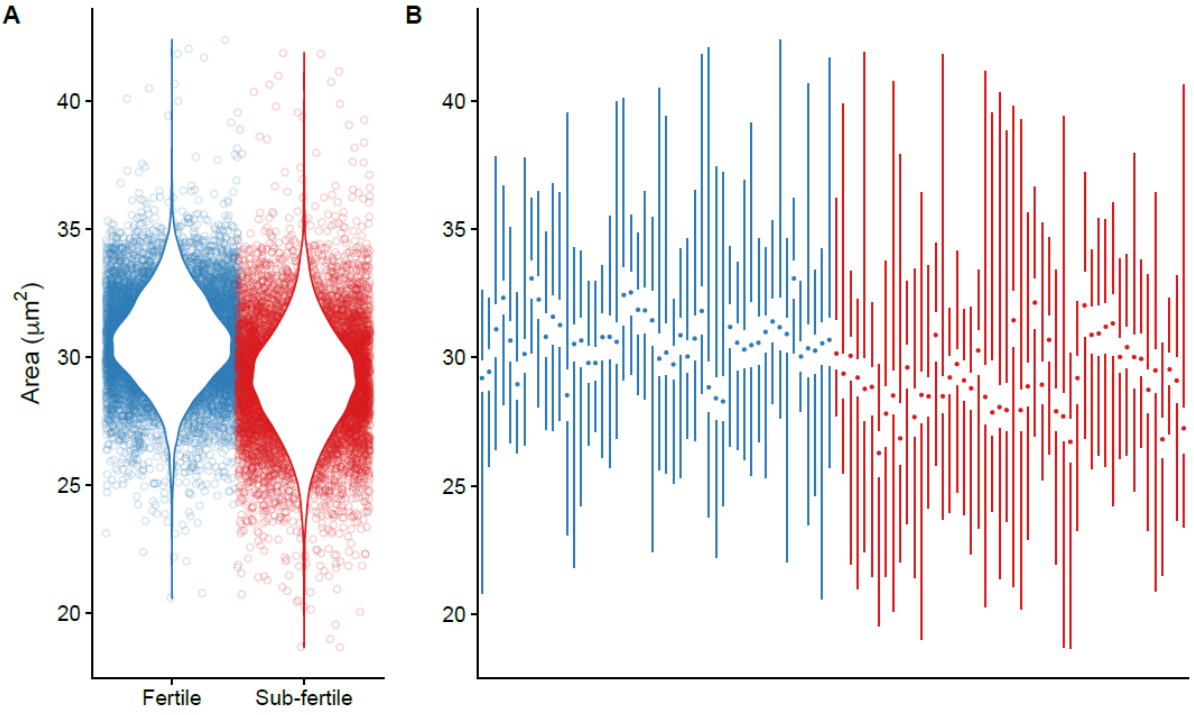


Figure 1: A representative summary of analysis workflow. (A) DAPI stained nucleus from a fertile boar sample, captured using an Olympus IX83 fluorescence microscope with pre-installed CellSens software at 1000x magnification. The software measures the interior angles along the perimeter of the nucleus as represented by point 1-6. (B) Schematic showing how these angles generate a profile. The figure shows the median and interquartile rages for one fertile sperm sample, and has been segmented at local maxima and minima. (C) The consensus fertile pig sperm head shape, showing positions of profile segments in the nucleus.

Analysis of nuclear morphology from 50 fertile and 50 sub-fertile animals yielded measures from 11,534 and 11,326 nuclei, respectively. Correlation analysis of measured sperm head characters indicated that many of the measures were highly correlated (Table S1). Therefore, further analysis was undertaken using measures of Area, Circularity, Variability, Bounding Height and Bounding Width. Given the orientation of the sperm heads, Bounding Height and Width are subsequently referred to as Maximum Height and Width.

Comparisons between fertile and sub-fertile animals indicate that sperm heads differ according to fertility status, but that there is a large amount of variation between individuals (Figure 2A and B, and Figures S1-4). Analysis of mean trait values from each animal indicates that sperm heads from fertile animals have a higher overall area (Area, $F_{1,98} = 34.55$, $p < 0.001$), are wider (Width, $F_{1,98} = 11.58$, $p =$

218 0.001), taller (Height, $F_{1,98} = 21.68$, $p < 0.001$) and are less variable in shape (Variability, $F_{1,98} = 24.98$, p
219 < 0.001) than those from sub-fertile animals. At this level of analysis, no difference between the sperm
220 heads of fertile and sub-fertile animals in Circularity is detected (Circularity, $F_{1,98} = 1.80$, $p = 0.18$).
221



222

223 **Figure 2:** Comparison of sperm head area between 50 fertile (blue) and 50 sub-fertile (red) boars. (A)
224 Area of 11,534 sperm heads from fertile animals (blue) and 11,326 sperm heads from sub-fertile
225 animals (red). (B) Individual Tufte boxplots of sperm head area for 50 fertile (blue) and 50 sub-fertile
226 (red) boars.
227

228 We hypothesised that individual sperm samples would contain different nuclei types and that this may
229 allow the identification of sub-fertile animals. For example, a certain sperm head type might only be
230 found in sub-fertile animals or might be overrepresented within such animals. Such differences might
231 represent abnormalities within these sub-fertile animals, which could also be used to investigate why
232 their fertility was impaired, *i.e.* this could be informative about the biology. Additionally, the presence
233 of differences of this type might allow the fertility of an animal to be predicted, *i.e.* this could allow
234 the automated assessment of fertility. We therefore, randomly allocated samples into two groups: a

training group of 30 fertile and 30 sub-fertile animals in which we investigated the presence, or not, of different nuclei types; and a test group of 20 fertile and 20 sub-fertile animals in which we investigate the ability of this type of analysis to predict fertility status.

In the training group, cluster analysis (Ward linkage, squared Euclidean distance and standardised variables) of 6,924 sperm heads from fertile animals and 6,684 sperm heads from sub-fertile animals supported the existence of three clusters (Figure 3). The identified clusters showed different membership for the fertile and sub-fertile sperm heads, with sperm heads from fertile animals overrepresented in the largest cluster and underrepresented in the smallest cluster.

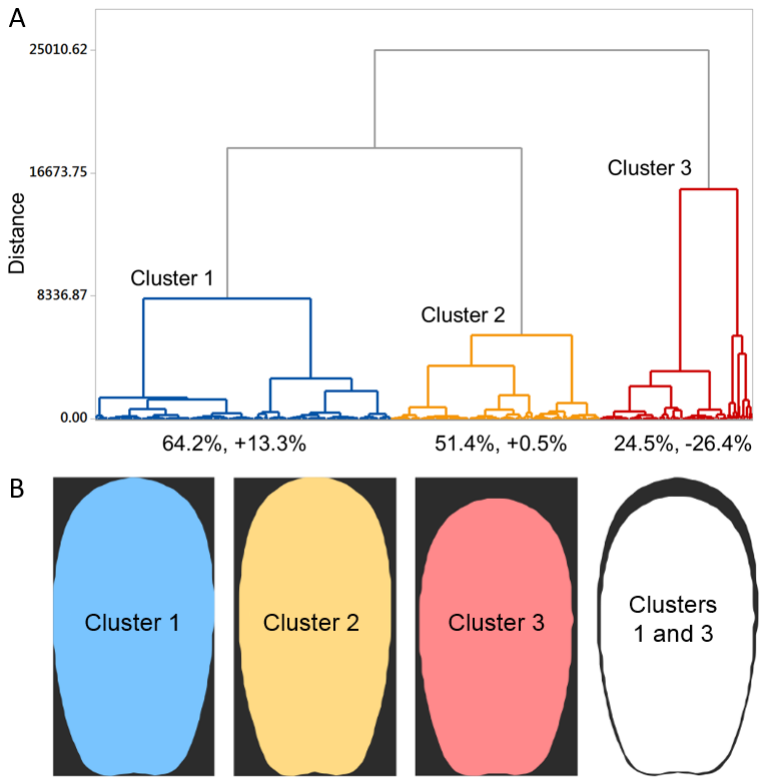


Figure 3: Analysis of the training group identifies distinct morphological clusters. (A) Cluster analysis of sperm from 30 fertile and 30 sub-fertile individuals using measures of Area, Circularity, Variability, Height and Width by Ward linkage using squared Euclidean distance and standardised variables. This identifies three distinct clusters. Shown below each cluster is the percentage of sperm heads within

that cluster that are from fertile animals and the difference in representation from that expected by chance (*i.e.* a positive value indicates that fertile nuclei are overrepresented in a cluster and *vice versa* for a negative value). (B) The consensus shapes of the three clusters demonstrate the separation of nuclei on size and shape. Cluster 1 contains the majority of the fertile sperm; cluster 2 has equal representation of fertile and sub-fertile sperm; cluster 3 predominantly contains sub-fertile sperm. The increasing compaction of the nuclei is readily apparent overlaying consensus nuclei from clusters 1 and 3.

Comparison of these clusters indicates that the cluster in which sperm heads from fertile animals are overrepresented is characterised by sperm heads with a low variability and a high area (Figure 4). In contrast, the cluster in which sub-fertile animals are overrepresented is characterised by sperm heads with a lower area and a high variability (Figure 4).

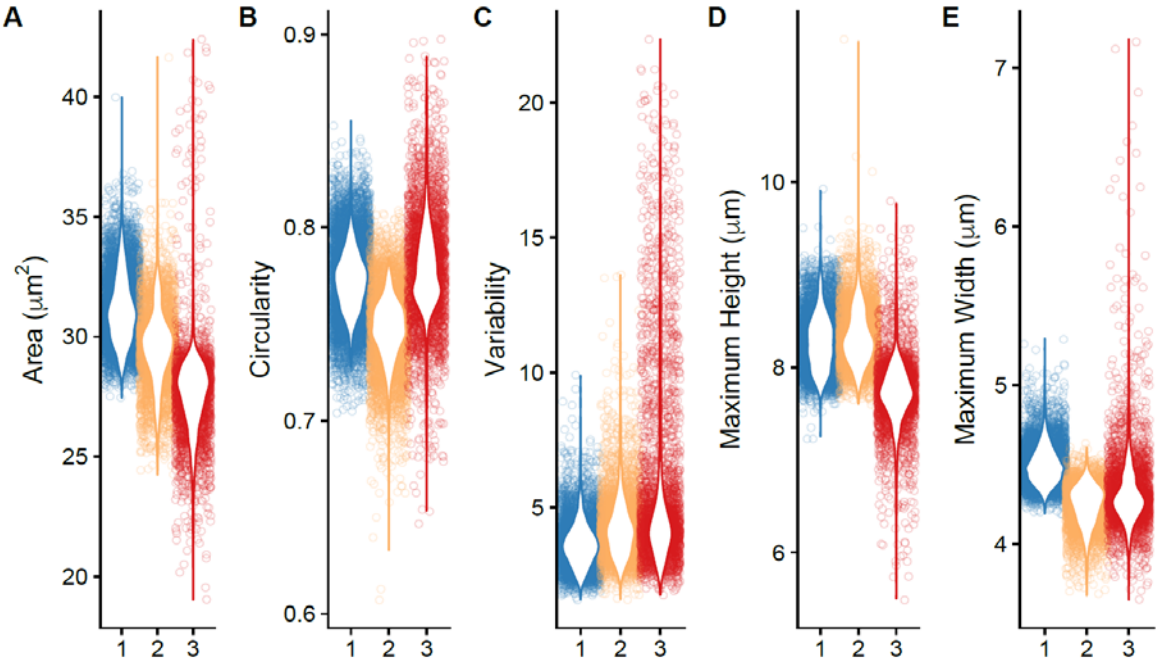


Figure 4: Sperm heads from the three clusters detected are morphologically distinct. Comparison of (A) Area, with 1>2>3 ($p < 0.001$ by pairwise *post hoc* test), (B) Circularity, with 3>1>2, (C) Variability,

267 with 3>2>1, (D) Maximum Height, with 2>1>3, and (E) Maximum width, with 1>3>2, for the three
268 clusters.
269
270 The three clusters identified in the training group were also recovered in the test group (Figure S5),
271 with analysis of sperm head morphology showing the same differences between clusters (data not
272 shown). As in the training group, sperm heads from fertile animals were overrepresented in cluster 1
273 and underrepresented in cluster 3 (Figure S5). This supports the idea that the frequency of certain
274 sperm head morphology types can be used to predict fertility. Several approaches were investigated
275 to test this. Firstly, we used the mean trait values for the three clusters identified in the training group
276 to define the initial partitions for a K-means clustering of the sperm heads from the test group of 20
277 fertile and 20 sub-fertile animals. The proportion of nuclei from each animal that was allocated to
278 each cluster was then determined and compared to the proportions determined from the training set
279 (Figure 5). This indicates that this method does accurately predict fertility in some animals, but that
280 no scheme would correctly identify status for all animals. That is, if the aim was to exclude any animal
281 where sub-fertility was suggested, then 8 of the 20 fertile animals and 6 of the 20 sub-fertile animals
282 would be retained (Table S2). Giving better results, retaining all those animals where cluster
283 membership suggested fertility would leave 13 of the 20 fertile animals and 6 of the 20 sub-fertile
284 animals (Table S2). Similarly, attempts to predict fertility using other approaches – for example
285 defining thresholds based on rates of variability within samples – also resulted in the inclusion of sub-
286 fertile animals or the exclusion of fertile animals.
287

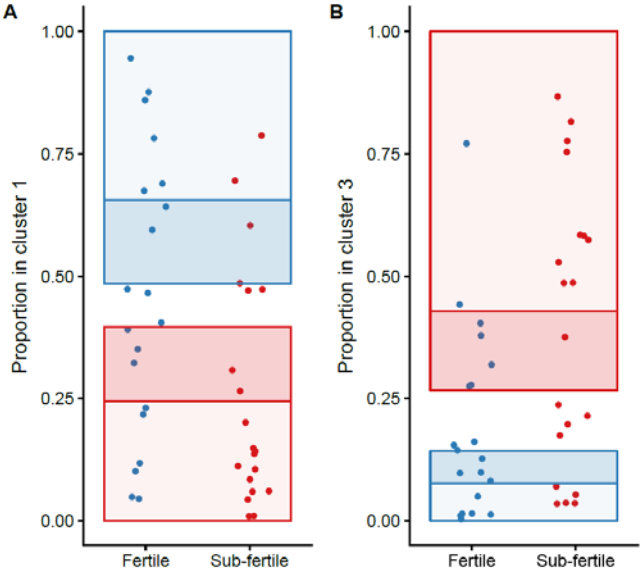


Figure 5: Cluster membership can predict fertility for some individuals. Shown are the proportion of sperm heads from each individual in the test group that are grouped into (A) cluster 1 and (B) cluster 3. Dark shaded rectangles highlight the 95% confidence intervals from cluster representation in the training set. Light shaded rectangles highlight areas between these confidence intervals and either 0 or 1. Red points falling into the region defined by the blue rectangles therefore, represent sub-fertile animals that would be classified as fertile and *vice versa* for blue points falling inside the regions defined by red rectangles.

Discussion

Over the last decade, AI has become commonplace in the pig breeding industry; it is therefore, an economical imperative to identify boars with prime fertility. To this end, various computer technologies have been developed, or adapted, to allow quantitative analysis of boar sperm characteristics. Prime examples are the CASA systems that have been developed and improved over almost four decades; these are now widely used in semen handling centres and spermatology laboratories. Currently, the Sperm Class Analyser (SCA) is considered the gold standard in automated sperm morphology analysis (ASMA) systems (Horst, 2015). SCA has been shown to provide accurate measurements of head, mid-piece and tail morphometry in several mammalian species including

306 humans (Soler *et al.*, 2003), horses (Hidalgo *et al.*, 2005) and goats (Hidalgo *et al.*, 2006). The absence
307 of standardised morphometric parameters and corresponding values which could be used to identify
308 both normal and abnormal sperm however, is a limitation which is common to most species (Horst,
309 2015).
310
311 We previously developed the Nuclear Morphology Analysis software for rapid and accurate
312 assessment of nuclear morphology in mouse lines (Skinner *et al.,* 2018). In contrast to other
313 morphometric approaches, such as elliptic Fourier analysis (Ostermeier *et al.*, 2001), or Procrustes-
314 based geometrics (Varea Sánchez *et al.*, 2013), our analysis can be run rapidly on many thousands of
315 nuclei, using automatic detection of landmarks and semi-landmarks within the nucleus, and the results
316 map cleanly back to the biological structure of the nucleus. In this study we have extended the
317 capability of the software to recognise pig sperm. The ability to distinguish subtle morphological
318 differences in pig sperm demonstrate the utility of this approach for other spatulate nuclei, as
319 commonly found in other mammalian species, including humans (Skinner and Johnson, 2017).
320
321 Here we have used this software to analyse sperm head morphology in pigs assessed as either fertile
322 or sub-fertile based on CASA data and manual assessment of morphology. These analyses identify
323 differences between the sperm heads of fertile and sub-fertile animals, with sperm heads from fertile
324 animals having a higher overall area – a consequence of greater width and height – and being less
325 variable in shape than those from sub-fertile animals (Figures 2 and S1-4). This observation does not
326 agree with previous work that, using a limited sample of 12 Pietrain boars, had suggested that high-
327 fertility boars had less elongated sperm heads that were significantly smaller than those of lower
328 fertility (non-return rate lower than 86%) (Hirai *et al.*, 2001). This study used the relationship between
329 fertility – based on non-return rate, which was expressed as a percentage of sows who were not
330 chosen for a second insemination between 60 and 90 days following the first insemination – and the
331 results for motility and morphology as assessed using the CASA system for was evaluated (Hirai *et al.*,

2001). This may represent a difference between breeds as it is known that sperm head dimensions differ between breeds; for example, the average sperm head area of Landrace and Large White sperm is 34.4 and 34.7μm^2, respectively (Saravia *et al.*, 2007).

Our analysis also identifies three clusters of morphology types (Figure 3). These clusters group sperm heads that have a low variability and a high area (cluster 1), that are tall and narrow (cluster 2), and those that have a low area and a high variability (cluster 3) (Figure 4). This mirrors the clusters identified by previous work on sperm head morphology in pigs that used measures of area, height (or length) and width (Hirai *et al.*, 2001; Vicente-Fiel *et al.*, 2013a, b). Here however, our analysis also indicates that the variability differs between these groups (Figure 4). We hypothesised that a certain type of sperm head might only be present, or be overrepresented, in semen from sub-fertile animals. Whilst it is expected that such differences in sperm head morphology would have an impact on the spermatozoa's fertilising potential (Curry, 2000), there is limited research on which precise morphological parameters can in fact impact fertility. Comparison of cluster membership indicated that sperm heads from fertile animals are overrepresented in cluster 1 and underrepresented in cluster 3, *i.e.* sub-fertile animals have a high incidence of sperm heads a low area and a high variability (Figure 4).

Given that chromatin packaging in the sperm nucleus has been shown to be vital for a successful pregnancy, the morphological differences may be associated with poor chromatin packaging, due to potential DNA damage and failure of sperm decondensation, which could result in fertilisation failure (Bianchi *et al.*, 1996; Sakkas *et al.*, 1996). Sperm subpopulations have also been identified based on biochemical parameters (Calamera *et al.*, 2003; Buffone *et al.*, 2004) and it would be of interest to determine the extent to which such groupings are coincident with groupings identified by assessment of morphology.

Given that automated approaches to identify sub-fertile individuals would be of value in pig production, we sought to determine if our measures of morphology could be used to predict fertility. It did not prove possible to completely separate the fertile and sub-fertile animals in our test group (see Figure 5 for an example of one approach). There are several possible reasons for this. Firstly, our analyses indicate that the fertility status does not explain all of the variation between individuals (see Figure 2B), but do not allow us to determine if this represents biological or technical variation as only one preparation was made of a single semen sample from each individual. Given that the samples used in this study were obtained from several different breeds of pig, Large White, Landrace, White Duroc, Hampshire and Pietrain, it would therefore, be of interest to determine how sperm head morphology differs between breeds. Secondly, the assessment of animals as either fertile or sub-fertile means that more quantitative differences between samples cannot be factored into the analysis. For example, a sample with a morphology score of 70% and a motility score of 4 would be classified as fertile whilst one with a morphology score of 80% and a motility score of 3 would be classified as sub-fertile.

In conclusion, here we have shown that high-throughput morphometric analysis of pig sperm reveals morphologically distinct populations and that there are differences in sperm head morphology between animals assessed as fertile and those assessed as sub-fertile. That variability exists between individual animals is, given the biological nature of the question asked here, unsurprising, but does suggest several ways that this work can be extended to look at the variation within individuals over time and between both individuals and breeds. As some studies have also suggested that routinely assessed sperm parameters (morphology, motility and concentration) are not entirely indicative of fertility or prolificacy (Gadea, 2005; O'Connor *et al.*, 2017), is becoming apparent that more detailed investigation of chromatin organisation is necessary. We aim to extend our analyses to determine whether chromatin compaction or chromosome position in pig sperm varies between fertile and sub-fertile samples.

384

385 **Declaration of interest**

386 None of the authors of this paper has a financial or personal relationship with other people or
387 organisations that could inappropriately influence or bias the content of the paper.

388

389 **Funding**

390 BMS was supported by the Leverhulme Trust (grant RPG337) and the Biotechnology and Biological
391 Sciences Research Council (BBSRC, grant BB/N000129/1).

392

393 **Acknowledgments**

394 We thank Canterbury Christ Church University for supporting AM, SH and KF. We would also like to
395 thank JSR Genetics Ltd. for supplying samples used in this study.

396

397 **Authors' contributions**

398 Conceptualisation, BMS and KF; Methodology, BMS, KF, SH; Software and Validation, BMS;
399 Investigation, AM; Data Curation and Formal Analysis, AM, SH; Visualisation, SH, BMS; Supervision and
400 Project Administration, KF, SH; Writing - Original Draft, AM, KF; Writing - Review and Editing, BMS, KF,
401 AM, SH; Resources, KF, GW. All authors gave final approval for publication.

402

403

References

Agarwal A, Deepinder F, Sharma RK, Ranga G and Li J (2008) Effect of cell phone usage on semen analysis in men attending infertility clinic: an observational study. *Fertility and Sterility* **89** 124–128.

Apostoli P, Kiss P, Porru S, Bonde JP and Vanhoorne M (1998) Male reproductive toxicity of lead in animals and humans. *Occup Environ Med* **55** 364–374.

Auger J, Jouannet P and Eustache F (2016) Another look at human sperm morphology. *Human Reproduction* **31** 10–23.

Bianchi PG, Manicardi GC, Urner F, Campana A and Sakkas D (1996) Chromatin packaging and morphology in ejaculated human spermatozoa: evidence of hidden anomalies in normal spermatozoa. *Molecular Human Reproduction* **2** 139–144.

Buffone MG, Doncel GF, Briggiler CIM, Vazquez-Levin MH and Calamera JC (2004) Human sperm subpopulations: relationship between functional quality and protein tyrosine phosphorylation. *Human Reproduction* **19** 139–146.

Calamera J, Buffone M, Ollero M, Alvarez J and Doncel GF (2003) Superoxide dismutase content and fatty acid composition in subsets of human spermatozoa from normozoospermic, asthenozoospermic, and polyzoospermic semen samples. *Molecular Reproduction and Development* **66** 422–430.

Coetzee K, Kruge TF and Lombard CJ (1998) Predictive value of normal sperm morphology: A structured literature review. *Human Reproduction Update* **4** 73–82.

Cooper TG, Noonan E, von Eckardstein S, Auger J, Baker HWG, Behre HM, Haugen TB, Kruger T, Wang C, Mbizvo MT *et al.* (2009) World Health Organization reference values for human semen characteristics. *Human Reproduction Update* **16** 231–245.

Curry M (2000) Cryopreservation of semen from domestic livestock. *Reviews of Reproduction* **5** 46–52.

Dziuk PJ and Henshaw G (1958) Fertility of boar semen artificially inseminated following *in vitro*

Storage1. *Journal of Animal Science* **17** 554–558.

Eliasson R (1971) Standards for investigation of human semen. *Andrologia* **3** 49–64.

Feitsma H (2009) Artificial insemination in pigs, research and developments in The Netherlands, a review. *Acta Scientiae Veterinariae* **37** 61–72.

Gadea J (2005) Sperm factors related to *in vitro* and *in vivo* porcine fertility. *Theriogenology* **63** 431–444.

Gerrits RJ, Lunney JK, Johnson LA, Pursel VG, Kraeling RR, Rohrer GA and Dobrinsky JR (2005) Perspectives for artificial insemination and genomics to improve global swine populations. *Theriogenology* **63** 283–299.

Guzick D, Overstreet J, Factor-Litvak P, Brazil C, Nakajima S, Coutifaris C, Carson S, Cisneros P, Stenkampf M, Hill J *et al.* (2001) Sperm morphology, motility, and concentration in fertile and infertile men. *The New England Journal of Medicine* **345** 1388–1393.

Hartman CG, Schoenfeld C and Copeland E (1964) Individualism in the Semen Picture of Infertile Men. *Fertility and Sterility* **15** 231–253.

Hidalgo M, Rodriguez I, Dorado J, Sanz J and Soler C (2005) Effect of sample size and staining methods on stallion sperm morphometry by the Sperm Class Analyzer. *Veterinarni Medicina* **50** 24–32.

Hidalgo M, Rodríguez I and Dorado J (2006) Influence of staining and sampling procedures on goat sperm morphometry using the Sperm Class Analyzer. *Theriogenology* **66** 996–1003.

Hirai M, Boersma A, Hoeflich A, Wolf E, Foll J, Aumuller R and Braun J (2001) Objectively measured sperm motility and sperm head morphometry in boars (*Sus scrofa*): Relation to fertility and seminal plasma growth factors. *Journal of Andrology* **22** 104–110.

Holt W V. and Van Look KJW (2004) Concepts in sperm heterogeneity, sperm selection and sperm competition as biological foundations for laboratory test of semen quality. *Reproduction* **127** 527–535.

Horst G Van Der (2015) Automated sperm morphology analysis. *Proceedings of the 25th Technical Conference on Artificial Insemination & Reproduction* 125–128.

456 **Huggins C, Scott WW and Heinen JH** (1942) Chemical composition of human semen and of the
457 secretions of the prostate and seminal vesicles. *American Journal of Physiology - Legacy Content*
458 **136** 467–473.
459 **Johnson LA, Aalbers JG, Willems CMT and Sybesma W** (1981) Use of boar spermatozoa for artificial
460 insemination and fertilizing capacity of fresh and frozen spermatozoa in sows on 36 farms.
461 *Journal of Animal Science* **52** 1130–1136.
462 **Jung A and Schuppe HC** (2007) Influence of genital heat stress on semen quality in humans. *Andrologia*
463 **39** 203–215.
464 **King GJ and Macpherson JW** (1973) A Comparison of Two Methods for Boar Semen Collection. *Journal*
465 *of Animal Science* **36** 563–565.
466 **López Rodríguez A, Rijsselaere T, Beek J, Vyt P, Van Soom A and Maes D** (2013) Boar seminal plasma
467 components and their relation with semen quality. *Systems Biology in Reproductive Medicine* **59**
468 5–12.
469 **MacLeod J and Gold RZ** (1951) The male Factor in fertility and infertility. *Fertility and Sterility* **2** 394–
470 414.
471 **Maes D, Nauwynck H, Rijsselaere T, Mateusen B, Vyt P, de Kruif A and Van Soom A** (2008) Diseases
472 in swine transmitted by artificial insemination: An overview. *Theriogenology* **70** 1337–1345.
473 **Maree L, Du Plessis SS, Menkveld R and Van Der Horst G** (2010) Morphometric dimensions of the
474 human sperm head depend on the staining method used. *Human Reproduction* **25** 1369–1382.
475 **Martin RH and Rademaker A** (1988) The relationship between sperm chromosomal abnormalities and
476 sperm morphology in humans. *Mutation Research Letters* **207** 159–164.
477 **Menkveld R and Kruger TF** (1996) Evaluation of sperm morphology by light microscopy. *Human*
478 *Spermatozoa in Assisted Reproduction* 89–107.
479 **O'Connor RE, Fonseka G, Frodsham R, Archibald AL, Lawrie M, Walling GA and Griffin DK** (2017)
480 Isolation of subtelomeric sequences of porcine chromosomes for translocation screening reveals
481 errors in the pig genome assembly. *Animal Genetics* **48** 395–403.

Ostermeier GC, Sargeant GA, Yandell BS, Evenson DP and Parrish JJ (2001) Relationship of bull fertility to sperm nuclear shape. *Journal of Andrology* **22** 595–603.

Overstreet JW and Katz DF (1987) Semen analysis. *The Urologic Clinics of North America* **14** 441–449.

Owen DH (2005) A Review of the physical and chemical properties of human semen and the formulation of a semen simulant. *Journal of Andrology* **26** 459–469.

Paz GF, Sofer A, Homonnai ZT and Kraicer PF (1977) Human semen analysis: seminal plasma and prostatic fluid compositions and their interlations with sperm quality. *International Journal of Fertility* **22** 140–147.

Perreault SD, Aitken RJ, Baker HWG, Evenson DP, Huszar G, Irvine DS, Morris ID, Morris RA, Robbins WA, Sakkas D *et al.* (2003) Integrating new tests of sperm genetic integrity into semen analysis: breakout group discussion. *Advances in Male Mediated Developmental Toxicity* 253–268.

Polge C, Day B and Groves T (1968) Synchronisation of ovulation and artificial insemination in pigs. *Veterinary Record* **83** 136–142.

R Core Team (2018) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL, https://www.R-project.org/

Roca J, Vázquez JM, Gil MA, Cuello C, Parrilla I and Martínez EA (2006) Challenges in pig artificial insemination. *Reproduction in Domestic Animals* **41** 43–53.

Rowe P, Comhaire F, Hargreave T and Mellows H (1993) WHO manual for the standardized investigation and diagnosis of the infertile couple. *Cambridge, England: Cambridge University Press*.

Sakkas D, Urner F, Bianchi PG, Bizzaro D, Wagner I, Jaquenoud N, Manicardi G and Campana A (1996) Sperm chromatin anomalies can influence decondensation after intracytoplasmic sperm injection. *Human Reproduction (Oxford, England)* **11** 837–843.

Saravia F, Núñez-Martínez I, Morán J, Soler C, Murielb A, Rodríguez-Martíneza H and Peña F (2007) Differences in boar sperm head shape and dimensions recorded by computer-assisted sperm morphometry are not related to chromatin integrity. *Theriogenology* **68** 196–203.

508 **Skinner BM and Johnson EE** (2017) Nuclear morphologies: their diversity and functional relevance.
509 *Chromosoma* **126** 195-212.
510 **Skinner BM, Rathje CC, Bacon J, Johnson EEP, Larson EL, Kopania EEK, Good JM, Yousafzai G, Affara**
511 **NA and Ellis PJI** (2018) A high-throughput method for unbiased quantitation and categorisation
512 of nuclear morphology. *bioRxiv* 312470, doi: 10.1101/312470.
513 **Singleton WL** (2001) State of the art in artificial insemination of pigs in the United States.
514 *Theriogenology* **56** 1305–1310.
515 **Soler C, De Monserrat JJ, Gutiérrez R, Nuñez J, Nuñez M, Sancho M, Pérez-Sánchez F and Cooper TG**
516 (2003) Use of the sperm-class analyser® for objective assessment of human sperm morphology.
517 *International Journal of Andrology* **26** 262–270.
518 **Tardif S, Laforest JP, Cormier N and Bailey JL** (1999) The importance of porcine sperm parameters on
519 fertility in vivo. *Theriogenology* **52** 447–459.
520 **U.S. Environmental Protection Agency .** (1996) Guidelines for reproductive toxicity risk assessment.
521 *Environmental Protection Agency* **61** 56274–56322.
522 **Varea Sánchez M, Bastir M and Roldan ERS** (2013) Geometric morphometrics of rodent sperm head
523 shape. *PLOS ONE* **8** 1–10.
524 **Vicente-Fiel S, Palacín I, Santolaria P, Hidalgo CO, Silvestre MA, Arrebola F and Yániz JL** (2013a) A
525 comparative study of the sperm nuclear morphometry in cattle, goat, sheep, and pigs using a
526 new computer-assisted method (CASMA-F). *Theriogenology* **79** 436–442.
527 **Vicente-Fiel S, Palacín I, Santolaria P and Yániz JL** (2013b) A comparative study of sperm
528 morphometric subpopulations in cattle, goat, sheep and pigs using a computer-assisted
529 fluorescence method (CASMA-F). *Animal Reproduction Science* **139** 182–189.
530 **WHO** (2010) Examination and processing of human semen. URL,
531 http://whqlibdoc.who.int/publications/2010/9789241547789_eng.pdf
532
533

Figure legends

Figure 1: A representative summary of analysis workflow. (A) DAPI stained nucleus from a fertile boar sample, captured using an Olympus IX83 fluorescence microscope with pre-installed CellSens software at 1000x magnification. The software measures the interior angles along the perimeter of the nucleus as represented by point 1-6. (B) Schematic showing how these angles generate a profile. The figure shows the median and interquartile rages for one fertile sperm sample, and has been segmented at local maxima and minima. (C) The consensus fertile pig sperm head shape, showing positions of profile segments in the nucleus.

Figure 2: Comparison of sperm head area between 50 fertile (blue) and 50 sub-fertile (red) boars. (A) Area of 11,534 sperm heads from fertile animals (blue) and 11,326 sperm heads from sub-fertile animals (red). (B) Individual Tufte boxplots of sperm head area for 50 fertile (blue) and 50 sub-fertile (red) boars.

Figure 3: Analysis of the training group identifies distinct morphological clusters. (A) Cluster analysis of sperm from 30 fertile and 30 sub-fertile individuals using measures of Area, Circularity, Variability, Height and Width by Ward linkage using squared Euclidean distance and standardised variables. This identifies three distinct clusters. Shown below each cluster is the percentage of sperm heads within that cluster that are from fertile animals and the difference in representation from that expected by chance (*i.e.* a positive value indicates that fertile nuclei are overrepresented in a cluster and *vice versa* for a negative value). (B) The consensus shapes of the three clusters demonstrate the separation of nuclei on size and shape. Cluster 1 contains the majority of the fertile sperm; cluster 2 has equal representation of fertile and sub-fertile sperm; cluster 3 predominantly contains sub-fertile sperm. The increasing compaction of the nuclei is readily apparent overlaying consensus nuclei from clusters 1 and 3.

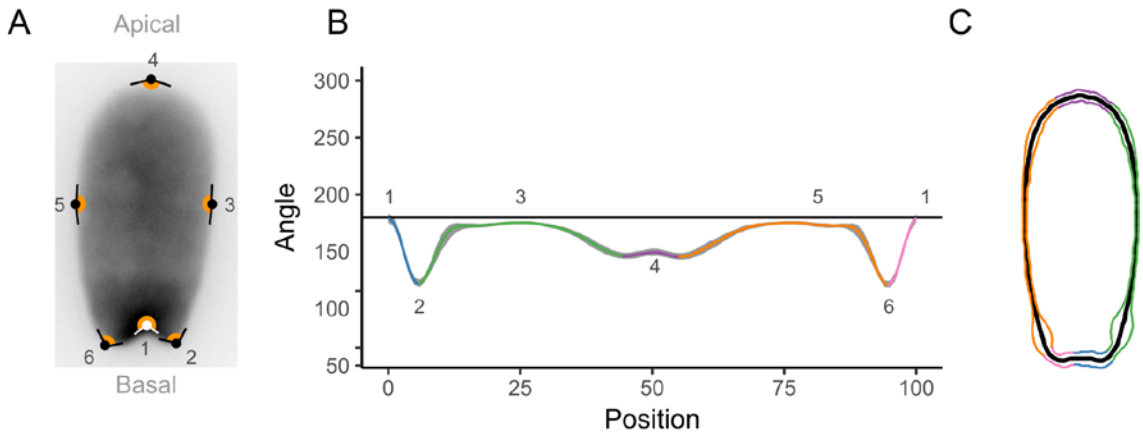
560 **Figure 4:** Sperm heads from the three clusters detected are morphologically distinct. Comparison of
561 (A) Area, with 1>2>3 ($p < 0.001$ by pairwise *post hoc* test), (B) Circularity, with 3>1>2, (C) Variability,
562 with 3>2>1, (D) Maximum Height, with 2>1>3, and (E) Maximum width, with 1>3>2, for the three
563 clusters.

564

565 **Figure 5:** Cluster membership can predict fertility for some individuals. Shown are the proportion of
566 sperm heads from each individual in the test group that are grouped into (A) cluster 1 and (B) cluster
567 3. Dark shaded rectangles highlight the 95% confidence intervals from cluster representation in the
568 training set. Light shaded rectangles highlight areas between these confidence intervals and either 0
569 or 1. Red points falling into the region defined by the blue rectangles therefore, represent sub-fertile
570 animals that would be classified as fertile and *vice versa* for blue points falling inside the regions
571 defined by red rectangles.

572

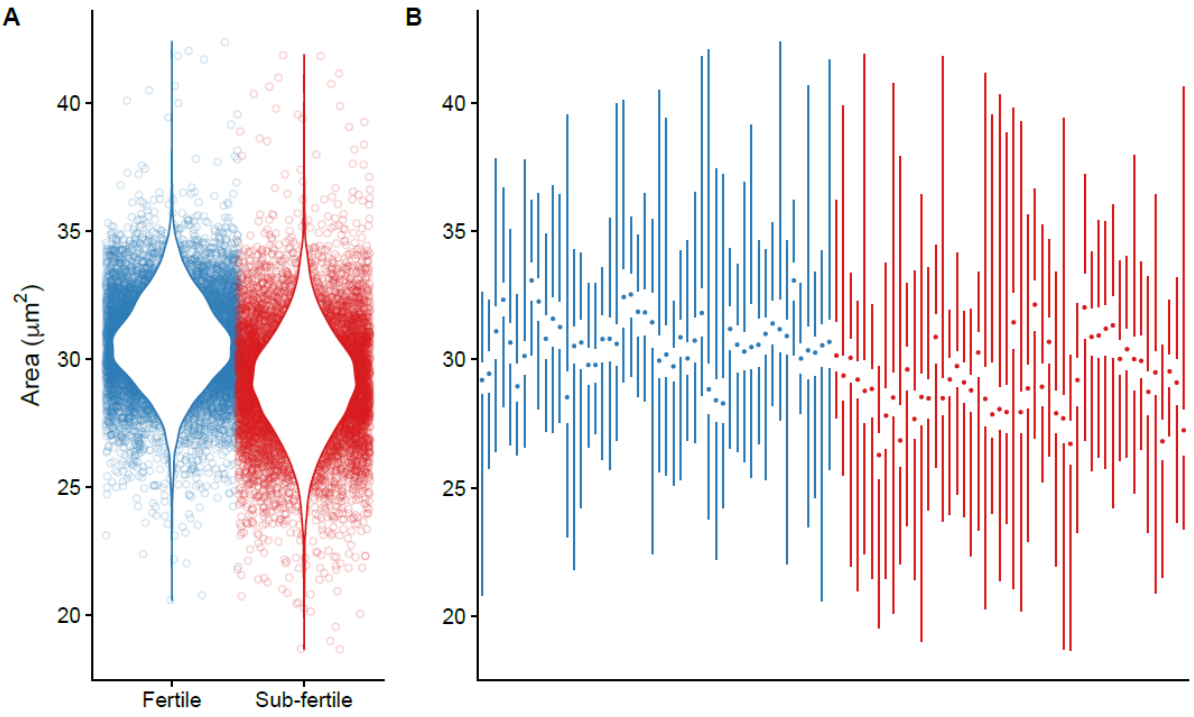
573 **Figures**



574

575 Figure 1

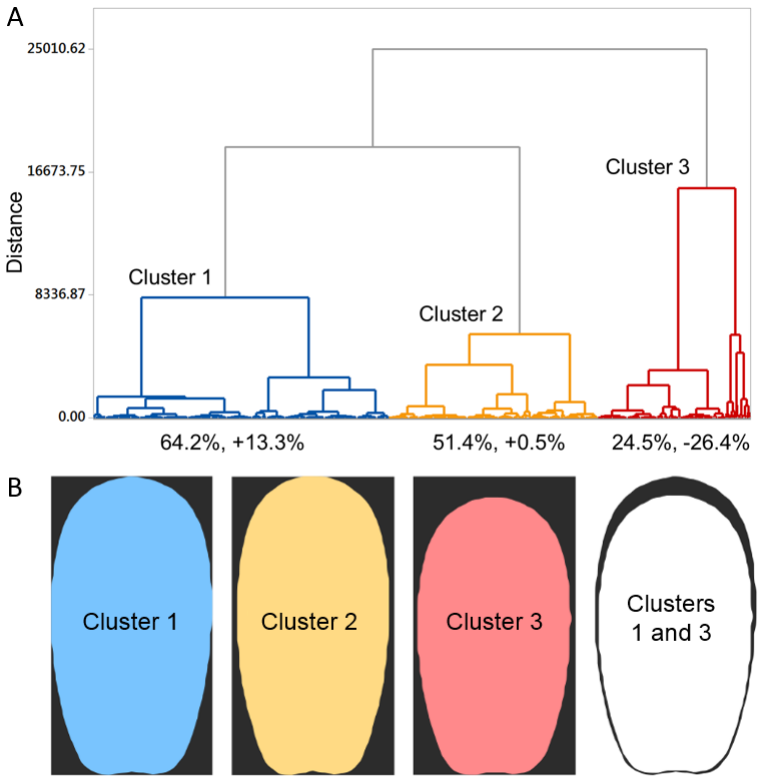
576



577

578 Figure 2

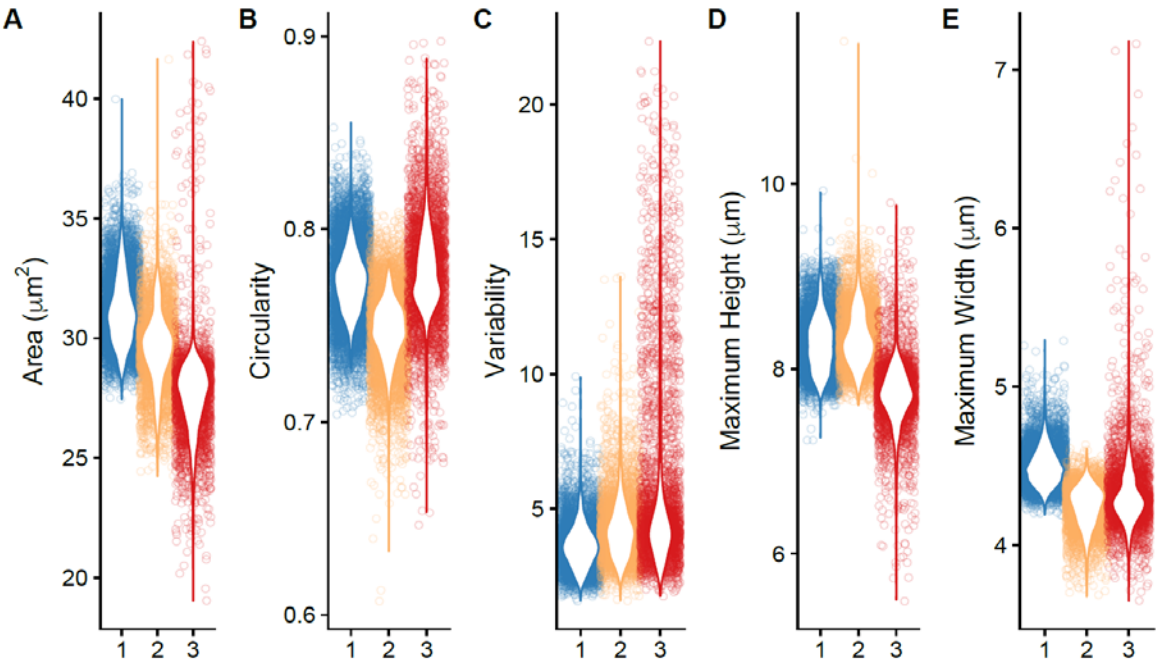
579



580

581 Figure 3

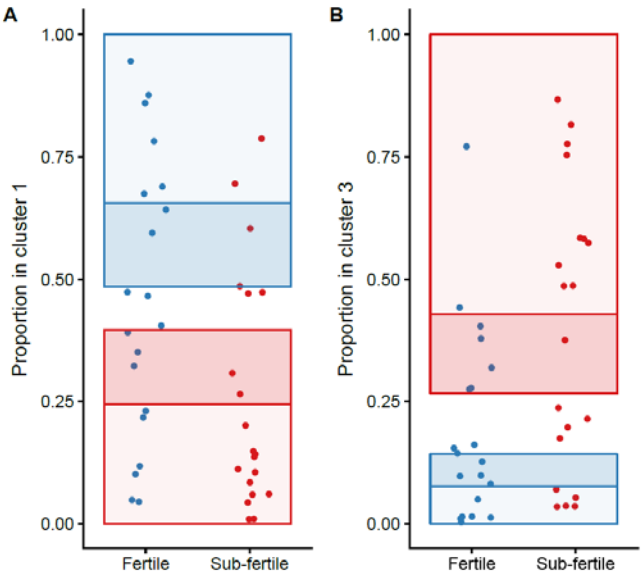
582



583

584 Figure 4

585



586

587 Figure 5

588