

## RESEARCH ARTICLE

# Validation of a laboratory-developed test of human sperm capacitation

Melissa A. Moody<sup>1</sup> | Cristina Cardona<sup>1</sup> | Alana J. Simpson<sup>1</sup> | T. Timothy Smith<sup>2</sup> | Alexander J. Travis<sup>1,3</sup>  | G. Charles Ostermeier<sup>1</sup><sup>1</sup> Androvia LifeSciences, Mountainside, New Jersey<sup>2</sup> ARTWorks Consulting, Inc., New York, New York<sup>3</sup> Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Ithaca, New York**Correspondence**

G. Charles Ostermeier, Sr. Scientist, Androvia LifeSciences, 187 Mill Lane, Mountainside, NJ 07092.

Email: costermeier@androvia.com

**Funding information**

Androvia LifeSciences

Sperm must undergo capacitation to become fertilization competent. Here we validated that monosialotetrahexosylganglioside ( $G_{M1}$ ) localization patterns, which were assessed in the Cap-Score™ Sperm Function Test, reflect a capacitated state in human sperm. First, we defined patterns representing sperm that do or do not respond to stimuli for capacitation. Sperm with “capacitated” patterns had exposed acrosomal carbohydrates and underwent acrosome exocytosis in response to calcium ionophore (A23187). Precision was evaluated by percent change of the Cap-Score measured for 50, 100, 150, and 200 sperm. Changes of 11%, 6%, and 5% were observed ( $n \geq 23$ ); therefore, we counted  $\geq 150$  sperm per condition. Variance within and between readers was evaluated using 20 stitched image files generated from unique ejaculates. Two trained readers randomly resampled each image 20 times, reporting an average standard deviation of 3 Cap-Score units and coefficient of variation of 13% when rescoring samples, with no difference between readers. Semen liquefaction times  $\leq 2$  hr and mechanical liquefaction with Pasteur or wide-orifice transfer pipettes did not alter Cap-Score values. However, liquefaction with chymotrypsin ( $p = 0.002$ ) and bromelain ( $p = 0.049$ ) reduced response to capacitating stimuli and induced membrane damage, while counterintuitively improving sperm motility. Together, these data validate the Cap-Score assay for the intended purpose of providing information on sperm capacitation and male fertility. In addition to its clinical utility as a diagnostic tool, this test of sperm function can reveal the impact of common practices of semen handling on the ability of sperm to respond to capacitation stimuli.

**KEYWORDS**bromelain, Cap-Score, chymotrypsin,  $G_{M1}$ , liquefaction methods, sperm function

## 1 | INTRODUCTION

Current literature estimates the number of infertile couples to be 73 million globally, with over 40% attributed to a male factor (Kumar & Singh, 2015). Standard semen analysis—assessing sperm count,

motility, and morphology—diagnoses less than 50% of all male infertility. Most infertile men are instead believed to have defects in sperm function, which are only diagnosed by repeated failed cycles of intrauterine insemination (Aboulghar et al., 2001; Practice Committee of the American Society for Reproductive Medicine, 2006; Tournaye,

**Abbreviations:** [Non-]Cap, incubated with [without] capacitation stimuli; Cap-Score, Cap-Score™ Sperm Function Test; CTB, cholera toxin beta subunit;  $G_{M1}$ , monosialotetrahexosylganglioside;  $G_{M1}/NCP$ , pattern of  $G_{M1}$  localization found in non-capacitated sperm;  $G_{M1}/CP$ , pattern of  $G_{M1}$  localization found in capacitated sperm; PNA, peanut agglutinin.

Melissa A. Moody and G. Charles Ostermeier contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2017 The Authors. *Molecular Reproduction and Development* Published by Wiley Periodicals, Inc.

2012). Traditional semen analysis does not test for sperm function and, as such, cannot report on the ability of sperm in that semen sample to fertilize (Lamb, 2010; Wang & Swerdloff, 2014). Tests of sperm function would therefore have enormous clinical impact in helping physicians counsel couples to the most appropriate form of assisted reproduction; indeed, the development of assays of sperm fertilizing ability are a recognized priority (Lamb, 2010; Oehninger, Franken, & Ombelet, 2014; Wang & Swerdloff, 2014).

Although freshly ejaculated spermatozoa appear morphologically mature and motile, they are fertilization incompetent until they undergo a maturational process known as “capacitation” (Austin, 1952; Chang, 1951). In most species, capacitation is dependent on the removal of sterols from the sperm plasma membrane (sterol efflux) and the influx of bicarbonate and calcium ions (Baldi et al., 1991; Bedu-Addo, Lefièvre, Moseley, Barratt, & Publicover, 2005; Cohen et al., 2014; Gadella & Van Gestel, 2004). The efflux of sterols that occurs during sperm capacitation changes membrane fluidity, allowing for the redistribution of specific membrane components (Cohen et al., 2014; Cross, 1998; Selvaraj et al., 2007, 2009).

The Cap-Score™ Sperm Function Test (Cap-Score) is an *in vitro*, laboratory-developed test designed to assess sperm function, particularly regarding capacitation. This assay detects and analyzes the localization patterns of the ganglioside  $G_{M1}$  to evaluate the fertilizing ability of sperm. Conducting a Cap-Score test involves the incubation of sperm in non-capacitating (Non-Cap) medium and medium containing capacitating stimuli (Cap). The sperm that respond to the capacitation stimuli are identified by specific  $G_{M1}$  localization patterns. The final readout—the “Cap-Score”—reports the proportion of sperm within a sample that display the localization patterns that correspond with capacitation.

Validation of diagnostic assays involves multiple steps, typically measuring assay precision, reproducibility, and accuracy (Cleophas & Zwinderman, 2012). “Accuracy” is used in different ways, including for demonstration of “fit for purpose,” demonstration that the assay quantifies its intended target, and demonstration of statistical accuracy or quantification within an acceptable range of uncertainty. The process of validating an assay without precedent (such as the Cap-Score) is more complex than for an assay that simply alters the methodology of an established “gold standard” test: an assay of a novel biomarker must be reproducible and precise within samples and between readers. Other endpoints or markers of capacitation must also be used to support the underlying assertion—in this case, that the “capacitated”  $G_{M1}$  localization pattern is actually identifying a capacitated sperm. Finally, clinical and population data are required to show that the capacitation status provides information on male fertility (“fit for purpose”).

Relevant to the question of whether or not the Cap-Score assay provides information on male fertility, it is important to note that  $G_{M1}$  localization patterns have been tested in an independent, *post hoc* clinical trial using a combination of retrospective and prospective medical histories analyzed after performing the assay (Cardona et al., 2017). In that study, Cap-Scores over a certain threshold were highly correlated with successful fertilization by natural conception or within three or fewer cycles of intrauterine insemination, and Cap-Scores

below that threshold were correlated with low fertilization success (Cardona et al., 2017). Because that trial was performed with a highly skewed patient population consistent with a tertiary care fertility clinic, cohort comparison data were also collected, comparing the distribution of Cap-Scores in a population of men with known fertility versus those questioning their fertility. Significantly more men questioning their fertility had Cap-Scores at or below one standard deviation below the normal mean (Cardona et al., 2017). Together, the data from those studies demonstrate the fitness of the test for its intended purpose. Indeed, capacitation status is reflected by the Cap-Score using the ultimate endpoint or criterion of capacitation—the acquisition of fertilizing ability. However, one needs to interrogate those specific cells with other methods to look at capacitation status to prove that the individual sperm with “capacitated” patterns are actually capacitated.

During capacitation, the plasma membrane and outer acrosomal membranes are reported to communicate, thereby exposing some molecules associated with the acrosome (Asano, Nelson-Harrington, & Travis, 2013; Cohen, Mukai, & Travis, 2016; Jin et al., 2011; Kim & Gerton, 2003; Kim, Foster, Kvasnicka, & Gerton, 2011). Capacitated sperm are also able to undergo acrosome exocytosis. Here we utilized the established tools of peanut agglutinin (PNA), a lectin that binds specific carbohydrate residues associated with the acrosome (Mortimer, Curtis, & Miller, 1987; Vázquez et al., 1996), and calcium ionophore (A23187), which causes calcium influx into sperm and induces acrosome exocytosis. Use of these or similar reagents can provide mutually reinforcing evidence of the capacitation status of sperm, and allow us to correlate these states to specific patterns of  $G_{M1}$  localization. For example, we previously showed that murine sperm exhibiting a specific pattern of  $G_{M1}$  localization were the same population that underwent acrosome exocytosis, confirming that the pattern in question reflected the capacitation status of those cells (Selvaraj et al., 2007).

Assay reproducibility and precision pose special problems when dealing with the highly heterogeneous populations of human sperm. The World Health Organization has published multiple guidelines to reduce subjectivity in evaluations of male fertility and standardize routine semen analysis (World Health Organization, 1999, 2010). Nonetheless, differences still exist within and among individual andrologists (Auger et al., 2000), raising the question of whether observed differences in semen quality are real or simply reflect modifications in processing or measurement methods (Knuth, Neuwinger, & Nieschlag, 1989). To improve the identification and treatment of male fertility, semen quality evaluations should demonstrate minimal variation both within single samples and when read by different individuals. This attribute is known as “precision,” which refers to how well a given measurement can be reproduced when applied repeatedly to multiple aliquots of a single homogeneous sample (Burd, 2010). Changes in precision (also referred to as random analytical error) result entirely from factors that may vary during normal operation of an assay. Larger sample sizes often lead to increased precision, particularly when estimating unknown parameters. For example, to estimate the proportion of sperm having intact acrosomes within an ejaculate, a more precise

estimate is obtained when 100 rather than 50 cells are sampled. However, at some point sampling additional cells results in diminishing returns and constitutes an unnecessary investment of time.

Handling techniques, such as semen liquefaction, must also be considered to ensure that differences in cell preparation that could cause changes in the assay readout are avoided. According to the World Health Organization guidelines, most semen samples should liquefy on their own within 15 min at room temperature. However, samples can be liquefied for up to 60 min at 37°C. And, if samples are not liquefied after 60 min, liquefaction may be induced by: (i) diluting the sample (1:1) with medium; (ii) repeatedly passing the sample through an 18 or 19 gauge needle or pipette (mechanical liquefaction); or (iii) treating the samples with enzymes such as chymotrypsin or bromelain (Mortimer, 1994; World Health Organization, 2010). Interestingly, certain liquefaction treatments affect seminal plasma biochemistry and semen parameters (World Health Organization, 2010). For example, increasing liquefaction time can decrease progressive motility in sperm (Mortimer, Swan, & Mortimer, 1998; Shao et al., 2010) and increase DNA fragmentation (Balasuriya, Serhal, Doshi, & Harper, 2012). Increases in sperm DNA fragmentation were also observed when mechanical liquefaction is performed with 18-gauge needles (Kussler et al., 2014).

Enzymatic liquefaction of semen is an alternative to extended time or mechanical methods. Chymotrypsin is a proteolytic enzyme that cleaves amide bonds alongside hydrophobic amino acids (tyrosine, tryptophan, phenylalanine), and is commonly used to treat hyperviscous semen samples (Mortimer, 1994). Non-viscous and viscous specimens exposed to chymotrypsin have shown little if any change in total motility, forward progression motility, concentration, pH, or volume, but do exhibit a slight change in sperm motility patterns (Chen, Lu, Xu, Huang, & Lu, 2006). Increased levels of  $\alpha$ -glucosidase activity, a biomarker that is positively correlated with semen quality (Said, Galeraud-Denis, Carreau, & Saad, 2009; Zöpfgen et al., 2000), are also observed following liquefaction with chymotrypsin (Xu et al., 2006), further supporting the minimal impact of chymotrypsin on semen quality. Bromelain, a thiol proteolytic enzyme found in the stem of pineapples, has also been used to treat viscous seminal plasma from dromedary camels, and did not affect sperm motility (Monaco et al., 2016). This enzyme also has anti-inflammatory, anti-thrombotic, and anti-metastatic properties (Chobotova, Vernallis, & Majid, 2010; Maurer, 2001; Metzger, Grabowska, Eckert, Rehse, & Maurer, 1999; Taussig & Batkin, 1988). Yet, very little research has focused on bromelain and its effect on human seminal parameters, such as morphology, sperm count, or sperm function, although it is being investigated/used anecdotally to treat hyperviscosity.

The present study provides a series of experiments that contribute to assay validation for the Cap-Score. First, cell biological experiments were performed to demonstrate assay accuracy. Large image files were generated, and two different readers were trained to determine Cap-Score. The data generated were analyzed to evaluate Cap-Score precision and its variation when determined by the same and different operators. In addition, the Cap-Score was

utilized to determine if the length of time of liquefaction, mechanical liquefaction, and enzymatic liquefaction affect capacitation.

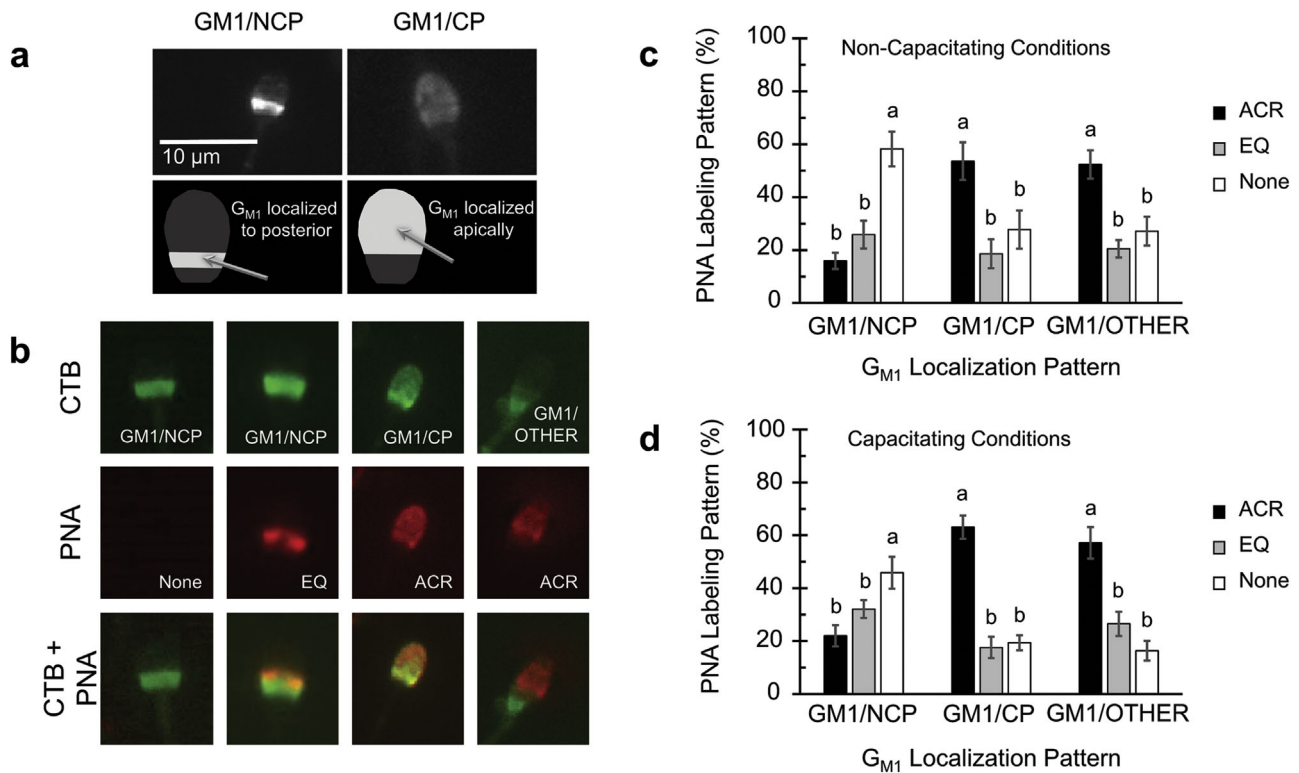
## 2 | RESULTS

### 2.1 | Accuracy

The Cap-Score detects and analyzes localization patterns of the ganglioside  $G_{M1}$ , which differ in sperm that respond to stimuli for capacitation. Typical  $G_{M1}$  localization patterns in sperm that either have responded to stimuli for capacitation ( $G_{M1}/CP$ ), or have either not been exposed to—or have not responded to—such stimuli ( $G_{M1}/NCP$ ) are shown (Figure 1a,b). Because human sperm are so heterogeneous, with many abnormal cells, a great many patterns can be seen beyond these two; one such example is shown in Figure 1b. However, we focused on the two patterns that are the keys for scoring.

Although the  $G_{M1}/CP$  pattern reproducibly appeared in sperm from different fertile men following exposure to stimuli for capacitation, and was correlated with clinical fertility (Cardona et al., 2017), those findings did not prove that the sperm having the  $G_{M1}/CP$  pattern were indeed capacitated. We therefore incubated sperm under non-capacitating (Non-Cap) and capacitating (Cap) conditions, and performed dual labeling with cholera toxin beta subunit (CTB) and PNA to investigate the relationship between capacitation and patterns of  $G_{M1}$  localization at the level of single cells (Figure 1b). Under Non-Cap conditions, the typical  $G_{M1}/NCP$  pattern predominated ( $68.3 \pm 3\%$ ;  $n = 10$ ). Most of the cells with the  $G_{M1}/NCP$  pattern did not label with PNA, regardless of the treatment condition (Figure 1c,d); this is consistent with an intact plasma membrane. The percentage of sperm having the  $G_{M1}/CP$  pattern increased from  $16.3 \pm 2.5\%$  ( $n = 10$ ) under Non-Cap conditions to  $28.0 \pm 2.6\%$  ( $n = 10$ ) under Cap conditions. Of great interest, the majority of cells with the  $G_{M1}/CP$  pattern did label with PNA over the acrosome (Figure 1c,d).

Acquiring the ability to exocytose the acrosome is a hallmark of capacitated sperm. If the Cap-Score did in fact reflect the sperm that were capacitated, we hypothesized that the Cap-Score should decrease following treatment with a calcium ionophore, which can trigger acrosome exocytosis in sperm that have undergone capacitation. A decrease in Cap-Score (vs. sperm from that same sample incubated under Cap conditions, but no ionophore) would be the predicted consequence of exocytosis, which disrupts the plasma membrane overlying the acrosome, preventing labeling or resulting in atypical, “other”  $G_{M1}$  patterns. Incubating sperm under capacitation conditions with the calcium ionophore A23187, we found the Cap-Score indeed decreased (Figure 2a,b). This finding is consistent with the model that cells undergoing acrosome exocytosis originated from the sub-population with a  $G_{M1}/CP$  pattern, and fit with earlier studies in the mouse showing that only the sub-population of sperm possessing the  $G_{M1}/CP$  pattern was capable of undergoing acrosome exocytosis (Selvaraj et al., 2007). Sperm incubated in basal (Non-Cap) medium only, and then treated with A23187, showed no change in Cap-Score (Figure 2c,d), supporting the notion that only capacitated cells are capable of acrosome exocytosis.



**FIGURE 1** Relationship between  $G_{M1}$  localization patterns, capacitation status, and exposure of acrosomal carbohydrates as detected with PNA. (a) Fluorescence microscopy images (top row) and diagrams (bottom row) of typical human sperm that have not been exposed to, or not responded to, stimuli for capacitation ( $G_{M1}/NCP$ ) (left column) or the  $G_{M1}$  distribution typical for cells that have responded to stimuli for capacitation ( $G_{M1}/CP$ ) (right column). (b) Representative images of  $G_{M1}/NCP$  patterns predominantly found in sperm incubated under non-capacitating conditions;  $G_{M1}/CP$  pattern that increased in response to incubation with stimuli for capacitation; and a pattern of  $G_{M1}$  localization that appeared infrequently (OTHER). Three patterns of PNA labeling were commonly seen in cells labeled with CTB: no labeling of PNA (None), labeling over the equatorial region (EQ); and labeling over the acrosome (ACR). (c and d)  $G_{M1}$  localization was compared to PNA labeling in sperm incubated under Non-Cap ( $n = 10$ ) and Cap ( $n = 10$ ) conditions. Black, gray, and white bars represent the proportion of cells with a given CTB labeling pattern having ACR, EQ, and None PNA label, respectively. Means within a given pattern of  $G_{M1}$  localization were compared. Those means with different superscripts were found to be different using Fisher's LSD ( $p < 0.05$ ). In general, the majority of sperm with the  $G_{M1}/NCP$  pattern did not label with PNA in both treatment conditions. In contrast, sperm with the  $G_{M1}/CP$  pattern predominantly showed PNA labeling over the acrosome

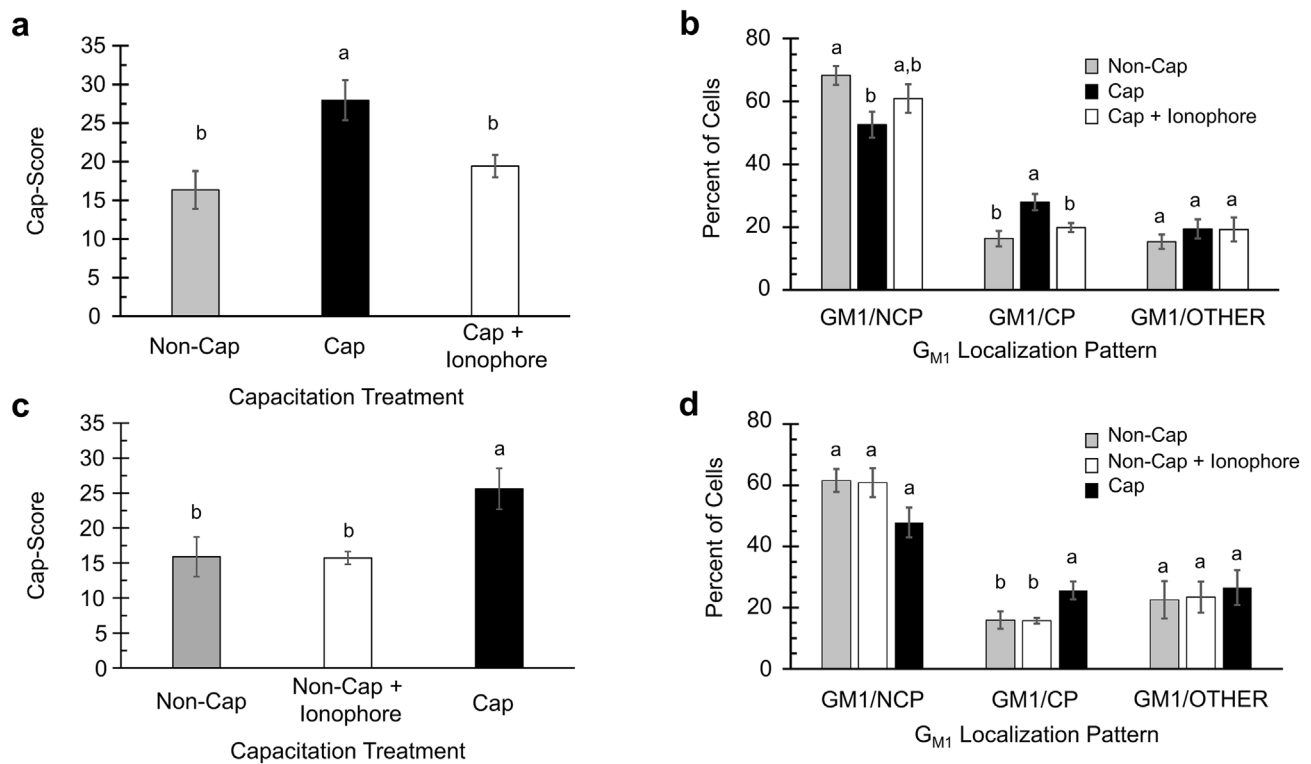
## 2.2 | Precision

Precision is defined as the repeatability or reproducibility of a measurement performed on the same sample (JCGM/WG2, 2008; Taylor & Cohen, 1998). The first step in determining Cap-Score precision was to define the number of cells to count per sample. In general, as the number of cells counted increases, there is an increase in precision, up to a point when the Cap-Score will not change appreciably with additional observations. We defined this threshold by measuring the percent change in Cap-Score when 50, 100, 150, and 200 sperm were evaluated. The percent change was large when counting 50 versus 100 sperm compared to counting 100 versus 150 and 150 versus 200 sperm (Table 1). Thus, Cap-Score precision was only modestly improved by counting more than 100 sperm; however, the 95% confidence intervals for the percent change when counting 50 versus 100 and 150 versus 200 did not overlap, suggesting a significant reduction in percent change when at least 150 cells were counted. A conservative Cap-Score value was therefore determined by counting the  $G_{M1}$  localization patterns of at least 150 cells.

## 2.3 | Statistical accuracy and reproducibility within a sample

Statistical accuracy can be defined as the proximity of measurements to the true value. The true value of an unknown population can be estimated by its central tendency, or the mean. One can judge whether a dataset has a strong or a weak central tendency based on its dispersion, or the inverse of precision (JCGM/WG2, 2008). The standard deviation and coefficient of variation (coefficient of variation = standard deviation/mean) measure the amount of dispersion within a sample.

Prior to evaluating Cap-Score accuracy by the same reader, we estimated the number of images for each reader to sample. Two semen donor groups were defined based on a cut-off of 1 standard deviation below the mean Cap-Score for a population of men with presumed fertility (pregnant wife or child less than 3 years old). The mean Cap-Score for the group with "lower Cap-Scores" was 27 and the "presumed fertile" group was 40. The standard deviations for each group were 5.2 and 4.9, respectively. A power analysis using a two-tailed test was done at the  $p < 0.05$  and  $p < 0.01$  levels, with a probability



**FIGURE 2** Effect of calcium ionophore A23187 on Cap-Score. Sperm were incubated with basal, Non-Cap medium ( $n = 10$ ), with stimuli for capacitation (Cap) ( $n = 10$ ), or with stimuli for capacitation and with calcium ionophore (A23187; Cap + ionophore) ( $n = 7$ ). (a) An increase in Cap-Score was observed from the Non-Cap to Cap treatment. A reduction in Cap-Score was observed in the Cap + ionophore treatment, showing that the sperm having undergone acrosome exocytosis originated from the subpopulation of cells having the  $G_{M1}/CP$  pattern. (b) Changes in  $G_{M1}$  localization patterns that led to the change in Cap-Score. The following capacitation treatments were assessed: Non-Cap (gray bar), Cap (black bar), and Cap + ionophore (white bar). (c) In a second set of studies, sperm were incubated with basal, Non-Cap medium ( $n = 5$ ), with non-capacitating medium and then incubated with A23187 (Non-Cap + ionophore;  $n = 5$ ), or with stimuli for capacitation (Cap) ( $n = 4$ ). No difference in Cap-Score was observed between the Non-Cap and Non-Cap + ionophore treatments, substantiating a lack of exocytosis in unstimulated sperm. (d) The  $G_{M1}$  localization pattern versus the percent of cells having that pattern for each of the following capacitation treatments: Non-Cap (gray bar), Non-Cap + ionophore (white bar), and Cap (black bar). The mean Cap-Scores were compared among capacitation treatments in (a) and (c), whereas the proportion of cells within the  $G_{M1}$  localization pattern were compared in (b) and (d). Those means with different superscripts were found to be different using Fisher's LSD ( $p < 0.05$ )

of detecting a difference this large, if it exists, of 90% (1-beta = 0.90). These analyses indicated that, respectively, 10 and 14 images should be sampled (five and seven per group). We therefore generated 10 images each in the "lower Cap-Score" and "presumed fertile" groups to ensure that each was sufficiently interrogated to identify any differences in reproducibility that might occur because of either low- or high-value Cap-Scores. Two different readers determined Cap-Scores by randomly resampling the 10 images each from the "lower Cap-Score" and "presumed fertile" groups 20 times. The average

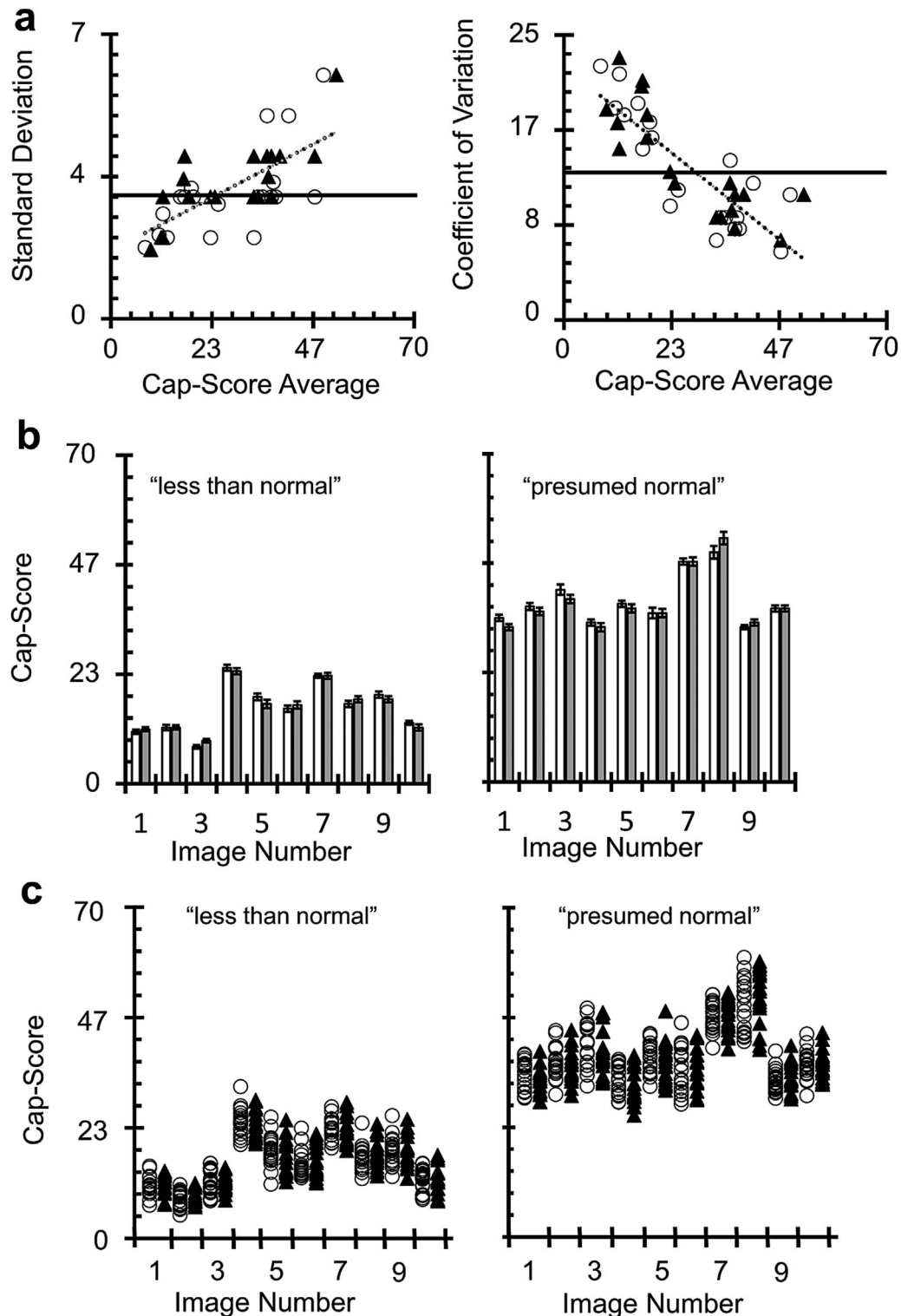
standard deviation across images and readers was 3 while the average coefficient of variation was 13% (Figure 3a). Both the standard deviation and coefficient of variation showed a linear relationship to Cap-Score, as determined by linear regression. Thus, while there was greater dispersion associated with reading higher Cap-Scores, it appeared to result from a greater Cap-Score magnitude. These data were consistent with a high degree of statistical accuracy because Cap-Score values were clustered tightly about the true value when the same sperm population was randomly resampled by the same or a different reader.

**TABLE 1** Percent change in Cap-Score with increasing number of counted sperm

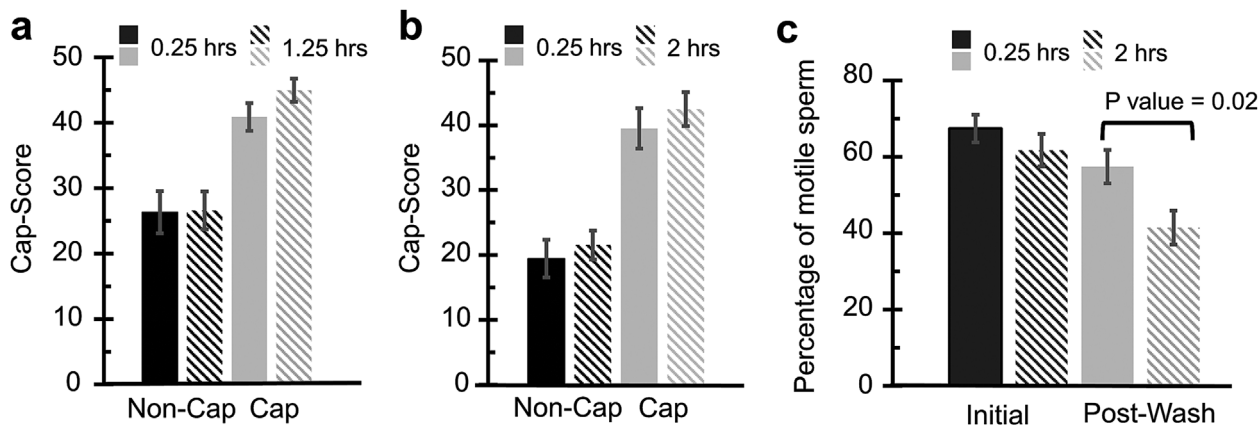
Number of sperm counted	Mean % change <sup>a</sup>	Number of observations	95% Confidence interval	
			Lower limit (%)	Upper limit (%)
50 versus 100	11 ± 9%	23	7	14
100 versus 150	6 ± 5%	26	4	8
150 versus 200	5 ± 3%	26	4	6

$\% \Delta = (y_2 - y_1) / y_2$ , where  $y_2$  and  $y_1$  are the larger and smaller Cap-Scores, respectively, with either the upper or lower number of sperm counted.

<sup>a</sup> ± standard deviation shown.



**FIGURE 3** Cap-Score readings are tightly clustered around the true value, and did not differ between readers. (a) Statistical accuracy of the Cap-Score assay. The average Cap-Score was plotted against the corresponding standard deviation or coefficient of variation (=standard deviation/mean). The average standard deviation and coefficient of variation for all images were 3 and 13, respectively, and are shown by the solid horizontal lines. The dotted lines show the linear dependence of the standard deviation ( $y = 0.06x + 0.02$ ;  $r = 0.69$ ;  $p < 0.0001$ ) and coefficient of variation ( $y = -0.32x + 0.22$ ;  $r = -0.84$ ;  $p < 0.0001$ ) to the Cap-Score. (b) Mean Cap-Scores were not different between readers for any image file. The  $p$ -values from two sample  $t$ -tests, comparing the average Cap-Score between readers, ranged from 0.02 to 0.99. (c) The variances in Cap-Score readings were not different between readers for any image file. The  $p$ -values for Bartlett's test of homoscedasticity ranged from 0.11 to 0.94. Reader 1, open circles and bars; reader 2, closed triangles and bars



**FIGURE 4** Impact of liquefaction time. Ejaculates were split into two portions: one was liquefied for 0.25 hr (solids bars) and the other was liquefied for either (a) 1.25 ( $n = 10$ ; dashed bars) or (b) 2 hr ( $n = 9$ ; dashed bars). After washing, both fractions were incubated under Non-Cap (black bars) and Cap (gray bars) conditions. Incubation treatment versus Cap-Score is presented. The average Cap-Scores in (a) were  $26 \pm 3\%$  (Non-Cap, 0.25 hr),  $27 \pm 3\%$  (Non-Cap, 1.25 hr),  $41 \pm 2\%$  (Cap, 0.25 hr), and  $45 \pm 2\%$  (Cap, 1.25 hr). The average Cap-Scores in (b) were  $19 \pm 3\%$  (Non-Cap, 0.25 hr),  $22 \pm 2\%$  (Non-Cap, 2 hr),  $40 \pm 3\%$  (Cap, 0.25 hr), and  $43 \pm 3\%$  (Cap, 2 hr). No difference was observed for Cap-Score between the liquefaction times for either capacitating treatment. (c) Motilities for nine samples were obtained following 0.25 hr (solid bars) and 2 hr (dashed bars) liquefaction times, immediately after liquefaction (initial) and after washing (post-wash). Average initial motilities for the 0.25- and 2-hr liquefaction times were  $67 \pm 4\%$  and  $62 \pm 4\%$ , respectively. Average post-wash motilities for the liquefaction times were  $57 \pm 4\%$  (0.25 hr) and  $41 \pm 4\%$  (2 hr). The 2-hr liquefaction treatment resulted in a significant drop in post-wash motility compared to the 0.25-hr liquefaction treatment. Means  $\pm$  standard errors are shown in each graph

## 2.4 | Evaluation of reproducibility between readers

Ten stitched images were obtained each for the “lower Cap-Score” and “presumed fertile” groups. Cap-Scores were determined by two different readers who randomly resampled each image 20 times; this allowed us to compare reproducibility in scoring between readers. Also, since each image file contained several magnitudes more sperm than were being sampled, each random resampling represented a distinct subsample of cells from within an individual ejaculate. There was no difference in mean Cap-Score (Figure 3b) or variance (Figure 3c) between readers for any image file. Cap-Score was therefore reproducible between readers because independent readers obtained similar Cap-Score distributions when resampling the same population of sperm. These data also further established assay precision and accuracy when sampling 150 sperm, as there was little change in Cap-Score when different subsamples of the same population were compared between readers.

## 2.5 | Impact of liquefaction time

Cap-Score was compared between the control of 0.25 hr and test samples of 1.25 ( $n = 10$ ) and 2 hr ( $n = 9$ ) to determine if an increase in liquefaction time would alter the proportion of sperm capable of responding to capacitation stimuli. When compared to the control, Cap-Scores of samples liquefied for 1.25 or 2 hr were not different for Non-Cap (Figure 4a; 0.25 vs. 1.25 hr  $p = 0.82$ ; Figure 4b; 0.25 vs. 2 hr  $p = 0.44$ ) or Cap samples (Figure 4a; 0.25 vs. 1.25 hr  $p = 0.16$ ; Figure 4b; 0.25 vs. 2 hr  $p = 0.47$ ).

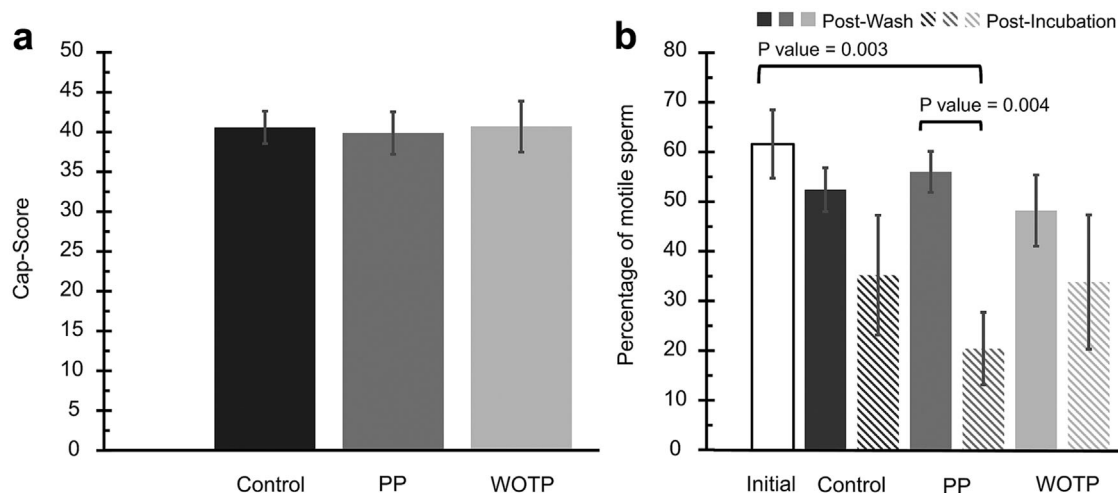
The percentage of motile sperm was measured immediately following liquefaction (initial) and after washing (post-wash) for the 0.25- and 2-hr liquefaction times (Figure 4c). Motilities for the 1.25-hr liquefaction time were only obtained after sample washing (data not shown). Initial motilities were similar for both 0.25- and 2-hr

liquefaction treatments ( $p = 0.39$ ), although a significant drop was observed in post-wash motility for the 2-hr liquefaction treatment ( $p = 0.02$ ). Post-wash motilities were similar for both the 0.25- and 1.25-hr liquefaction treatments ( $p = 0.87$ ; data not shown). Although capacitation was not affected, motility decreased for the 2-hr liquefaction subsequent to washing.

## 2.6 | Impact of mechanical liquefaction

Pilot studies revealed that passage through a hypodermic needle negatively affected the percentage of motile sperm and membrane integrity, so this method of viscosity reduction was not included in this study. Cap-Score was, however, obtained for five semen samples processed with a Pasteur pipette or a wide-orifice transfer pipette, and then compared to samples liquefied for 0.25 hr (control) to determine the impact of mechanical liquefaction on capacitation ability. Cap-Scores for Pasteur pipette and wide-orifice transfer pipette were not altered when compared to the control (Pasteur pipette,  $p = 0.73$ ; wide-orifice transfer pipette  $p > 0.99$ ) (Figure 5a).

Initial, post-wash, and post-incubation motilities were obtained for the control, Pasteur pipette, and wide-orifice-transfer-pipette treatments; no difference was noted in post-wash ( $p = 0.19$ ) or post-incubation ( $p = 0.11$ ) motilities for the wide-orifice transfer pipette treatment when compared to initial populations (Figure 5b). In contrast, motility decreased for samples processed with a Pasteur pipette when comparing initial to post-incubation treatments ( $p = 0.003$ ) as well as post-wash to post-incubation treatments ( $p = 0.004$ ). Although mechanical liquefaction did not appear to affect the ability to capacitate, as measured by the Cap-Score, processing samples with a Pasteur pipette can have a negative impact on motility subsequent to washing and extended incubation.



**FIGURE 5** The use of Pasteur pipettes and wide-orifice transfer pipettes for mechanical liquefaction. Ejaculates of five individuals were split into three portions: mechanical liquefaction (control; black bar); 1:1 dilution with mHTF medium and mechanical liquefaction with Pasteur pipettes (PP; dark gray bar); and 1:1 dilution with mHTF medium and mechanical liquefaction with wide orifice transfer pipettes (WOTP; light gray bar). All three portions were incubated under Cap conditions, and then processed for Cap-Score. (a) Cap-Score following each liquefaction treatment. No difference in the average Cap-Scores was observed. (b) Analysis of the same samples for the percentage of motile sperm at the following times: initial (immediately after liquefaction; open bar), post-wash (solid bars), and following incubation with stimuli for capacitation (dashed bars). No change in motility was observed between the post-wash and post-incubation samples for the control or for those liquefied with WOTP. In contrast, a drop in motility for the Pasteur pipette post-incubation treatment was observed when compared both to the initial and post-wash treatments. Means  $\pm$  standard errors are shown in each graph

## 2.7 | Impact of enzymatic liquefaction

### 2.7.1 | Chymotrypsin

The impact of enzymatic liquefaction on capacitation was tested via Cap-Score, first by using samples liquefied with chymotrypsin (which was then removed during the washing process) or incubated for 3 hr in the presence of chymotrypsin and then compared to samples liquefied only for 0.25 hr (control;  $N = 10$ ). Samples liquefied with chymotrypsin and incubated under Non-Cap conditions were not affected ( $p = 0.66$ ). In contrast, samples liquefied with chymotrypsin and incubated under Cap conditions exhibited a drop in Cap-Score when compared to the control ( $p = 0.002$ ) (Figure 6a) due to a decrease in the number of capacitated cells combined with an increase in the number of abnormal cells. Cap-Scores could not be obtained for samples incubated in the presence of chymotrypsin due to high concentrations of membrane damage and what appeared to be cell swelling consistent with volume regulatory defects (Figure 6b).

Post-wash and post-incubation Non-Cap and Cap motilities were calculated for all three chymotrypsin treatments. A difference was observed for samples incubated with chymotrypsin in both the Non-Cap and Cap treatments when comparing the average motilities of the control to the chymotrypsin liquefaction treatments ( $p = 0.001, 0.004$ ) (Figure 6c).

### 2.7.2 | Bromelain

A pilot study was performed using the enzyme bromelain to determine if the effect of enzymatic liquefaction with chymotrypsin on capacitation was specific to chymotrypsin or could be induced by other proteases used to reduce viscosity. Five samples were liquefied for 0.25 hr with bromelain, and compared to control samples. Samples

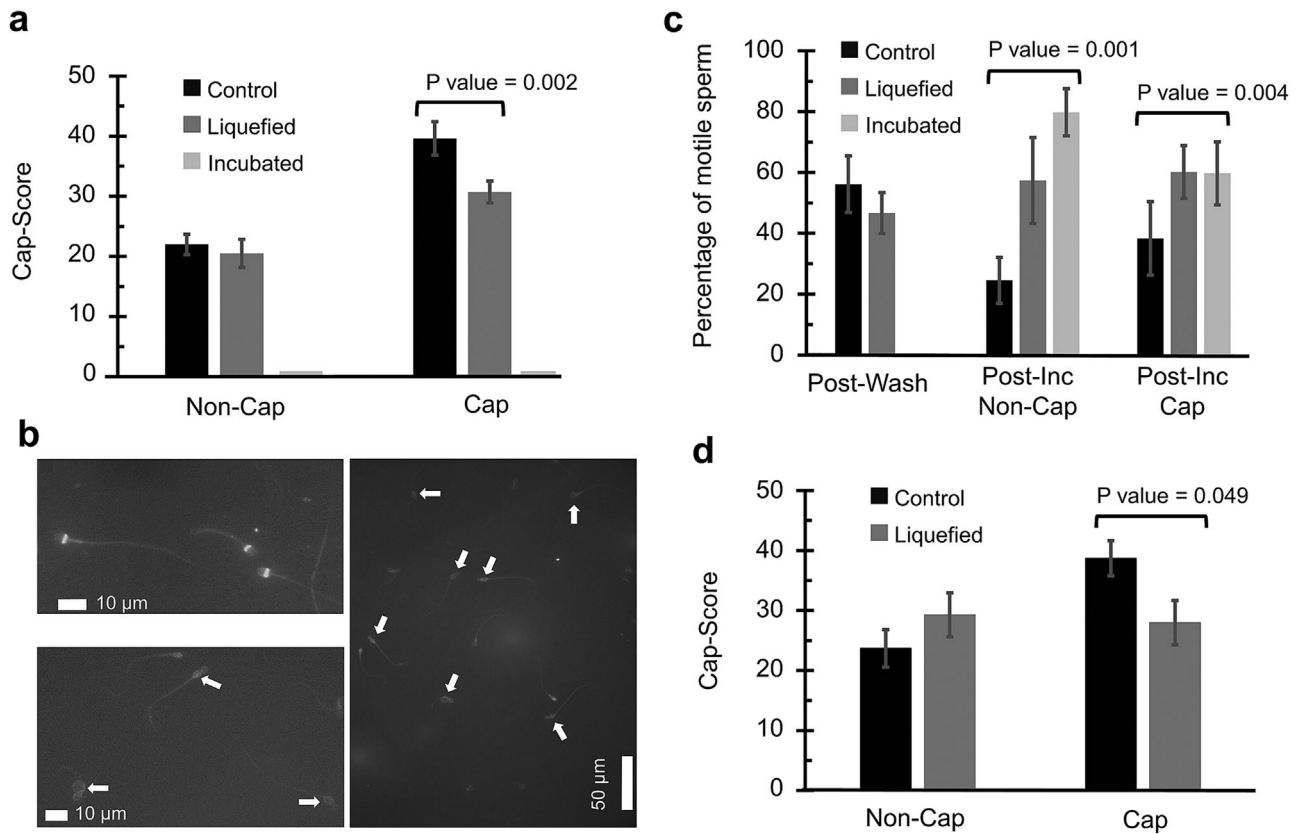
liquefied with bromelain and incubated under Cap conditions exhibited a drop in Cap-Score compared to the control due to a decrease in capacitated cells and an increase in abnormal cells ( $p = 0.049$ ) (Figure 6d). Similar to our findings with chymotrypsin, samples liquefied with bromelain and incubated under Non-Cap conditions were not affected ( $p = 0.32$ ). Cap-Scores were not obtained for samples incubated with bromelain due to the large concentration of damaged cells found in the incubated chymotrypsin treatment.

## 3 | DISCUSSION

The results presented here validate the accuracy, repeatability, and precision of the Cap-Score Sperm Function Test; they also provide important information on the influence of semen handling/preparation methods on assay outcomes. This report does not provide the routine protocols for validation of the laboratory methodology to perform the assay, which is beyond the scope of this work.

We investigated if the assay accurately identified sperm that responded to stimuli for capacitation from those that either were not exposed, or did not respond. We used two different approaches to establish that the  $G_{M1}/CP$  pattern did in fact identify capacitated sperm: (i) dual-labeling experiments with PNA supported the notion that sperm with the  $G_{M1}/CP$  pattern also demonstrated communication between the plasma membrane and outer acrosomal membrane. This observation is consistent with the current working model of plasma membrane dynamics during capacitation, in which point-fusions occur between the plasma membrane and underlying outer acrosomal membrane, exposing acrosomal matrix contents





**FIGURE 6** Use of chymotrypsin or bromelain for liquefaction negatively affected Cap-Score. (a) Ten ejaculates were split into three portions. The first portion was liquefied for 0.25 hr (control; black bars), the second portion was liquefied with 5 mg chymotrypsin per ejaculate for 0.25 hr (dark gray bars), and the third portion was incubated with 3 mg/ml chymotrypsin for 3 hr (light gray bars). A significant drop in Cap-Score for Cap samples was observed between chymotrypsin conditions compared to control. Non-Cap and Cap samples incubated with chymotrypsin could not be scored due to high levels of membrane damage. (b) Samples incubated under control conditions and labeled for localization of  $G_{M1}$  appear normal (top left). However, samples incubated with chymotrypsin had high concentrations of sperm lacking recognizable  $G_{M1}$  localization patterns and exhibiting membrane damage (bottom left). Many of the sperm had enlarged cytoplasmic droplets (arrows), which are consistent with defects in volume regulation. A wider field of view demonstrates the abundance of cells having enlarged cytoplasmic droplets (arrows; image on right). The samples from (a) were further analyzed in (c). Post-wash and post-incubation motilities were obtained for Non-Cap and Cap samples liquefied for 0.25 hr (control; black bars), samples liquefied with chymotrypsin for 0.25 hr (dark gray bars), and samples incubated with chymotrypsin for 3 hr (light gray bars). A significant increase in the percentage of motile sperm for samples incubated with chymotrypsin was observed compared to the control in the post-incubation Non-Cap and Cap treatments. (d) Five other ejaculates were split into two portions, and the first portion was liquefied for 0.25 hr (control; black bars) while the second portion was liquefied for 0.25 hr with bromelain (dark gray bars). All portions were washed and incubated for 3 hr at 37°C in either Non-Cap or Cap conditions, and processed for Cap-Score. A significant drop in Cap-Score was observed for Cap samples with bromelain compared to control. Means  $\pm$  standard errors are shown in each graph

and thus the ligand for PNA. (ii) A portion of the sperm with the  $G_{M1}$ /CP pattern completed acrosome exocytosis in response to calcium ionophore, more definitively linking Cap-Score and capacitation at the level of individual gametes. Because the  $G_{M1}$ /CP pattern is localized over the rostral sperm head, we hypothesized that acrosome exocytosis would result in loss of the membranes contributing to that pattern, and hence either no labeling with CTB or the appearance of cells having an “other” pattern that is consistent with membrane perturbation; both outcomes would lower the Cap-Score, which was observed.

Next, we showed that the precision of this assay was only modestly improved by counting more than 100 sperm, which provides a threshold to ensure that the number of sperm evaluated for each assay was representative of the ejaculate’s distribution. Despite the

relatively low number required, we chose to score the  $G_{M1}$  localization pattern of at least 150 sperm to be conservative.

The consistency of assay readout among different andrologists was addressed by training two individuals to read  $G_{M1}$  patterns, and then assessing the distribution of their Cap-Score readings. Twenty large-image files that contained up to 5,000 sperm each were created and resampled 20 times by each reader. An average standard deviation and average coefficient of variation revealed that the assay variance and/or dispersion were small and stable, which is indicative of a high degree of Cap-Score reproducibility per reader. Comparison of the scoring between the two readers showed an average difference of 1 in mean Cap-Score. When the Bonferroni correction was applied, no discernable differences were observed. Similarly, Cap-Score variances were not different

between readers. Collectively, these comparisons demonstrate that the Cap-Score is highly reproducible and reliable within and between trained readers, which are key considerations when attempting to evaluate male reproductive fitness.

Another source of variance that could affect assay reliability lies in semen handling/preparation and processing techniques. We focused on three main liquefaction approaches, as sperm are sensitive to this step of handling and there is significant variation among andrologists in liquefaction methods. The time allowed for liquefaction in assisted reproduction clinics typically ranges from 0.25 to 1 hr, at 37°C in air or CO<sub>2</sub> incubators. In the present study, liquefaction durations of up to 2 hr did not change Cap-Score, whereas motility was reduced in post-wash samples liquefied for 2 hr. This effect might be a consequence of extended sperm exposure to proteins found in seminal plasma that inhibit motility, such as semenogelin I (Lilja, Abrahamsson, & Lundwall, 1989; Robert & Gagnon, 1999) or its binding partner EPPIN (Silva, Hamil, & O'Rand, 2013). Prolonged exposure to motility-inhibiting proteins might also affect the ability to separate such inhibitory proteins from sperm by washing.

Mechanical liquefaction with glass Pasteur pipettes or plastic wide-orifice transfer pipettes did not affect capacitation, although use of Pasteur pipettes did result in a decline in post-wash motility. This effect could be a result of differences in expulsion properties between glass and plastic (e.g., stickiness) or in orifice diameter, which influences physical stress (e.g., shear forces). Physical stressors, such as centrifugation and repeated pipetting, were reported to damage rodent and human sperm and to interfere with their motility (Alvarez et al., 1993; Katkov & Mazur, 1998; Varisli, Uguz, Agca, & Agca, 2009). Interestingly, some physical stressors lead to sublethal damage in human sperm that is only observed during extended incubations (Alvarez et al., 1993). Thus it is possible that damage induced by Pasteur pipettes rendered these sperm less able to withstand sample washing or to survive the subsequent incubation. Nonetheless, our data support the view that the functional impacts were minimal, as even after extended incubations there was no impact of mechanical liquefaction with either Pasteur or plastic wide orifice transfer pipettes on capacitation ability.

Whereas liquefaction timing and mechanical liquefaction had little impact on Cap-Score, enzymatic liquefaction with chymotrypsin or bromelain reduced the Cap-Score in samples stimulated for capacitation. If the impairment resulted from changes in membrane curvature or surface lipid presentation that altered G<sub>M1</sub> labeling, a similar reduction in Cap-Score should have been observed in the Non-Cap treatment group; however, this was not the case. Thus, proteolytic enzymes might inhibit capacitation or mechanisms associated with capacitation, and the mechanism(s) warrant additional investigation. Potential targets for the enzymes include any of the numerous plasma membrane proteins, particularly ion channels that are needed for volume regulation or capacitation. Chymotrypsin and bromelain might also damage proteins found in the seminal plasma that positively regulate capacitation in a manner similar to the heparin-binding proteins found in cattle (Miller, Winer, & Ax, 1990).

Although not a common or recommended practice, incubation for 3 hr with chymotrypsin provided further evidence that these enzymes negatively affected proteins within the sperm membrane. Sperm incubated under both Non-Cap and Cap conditions exhibited significant damage, which was consistent with inability to regulate volume and prevented CTB labeling for determination of G<sub>M1</sub> localization patterns. This extended exposure time experiment was intended not only to simulate samples containing residual chymotrypsin following washing (or exposure to very high concentrations for shorter time periods), but also to exaggerate the presence and to suggest the potential nature of damage done to the sperm during a typical 0.25 hr liquefaction with chymotrypsin. Volume is regulated in sperm cells by potassium and chloride ion channels involved in cyclic adenosine monophosphate (cAMP)-mediated protein tyrosine phosphorylation pathways (Petrunkina, Harrison, Tsovala, Jebe, & Topfer-Petersen, 2007; Petrunkina, Jebe, & Topfer-Petersen, 2005; Yeung, Barfield, Anapolski, & Cooper, 2004). Water channels have also been detected on the cytoplasmic droplet in mouse (Yeung, Callies, Rojek, Nielsen, & Cooper, 2009) and rat sperm (Yeung & Cooper, 2010), and in the midpiece membranes in human sperm (Yeung, Callies, Tuttelmann, Kliesch, & Cooper, 2010). The swollen appearance of the cells suggests that exposing sperm to chymotrypsin or bromelain can damage these membrane channels, leading to the dysregulation of volume and capacitation in samples exposed to capacitating stimuli.

Despite the effects on cell volume and membrane architecture, the percentage of motile sperm increased in both Non-Cap and Cap samples incubated with chymotrypsin, when compared to control populations without the protease. The increase in motility might have resulted from altered intracellular concentration of ions, such as calcium or bicarbonate, that regulate sperm motility (Fakih, MacLusky, DeCherney, Wallimann, & Huszar, 1986; Turner, 2006). Alternatively, treatment with chymotrypsin might have damaged motility-inhibitory substances, such as protein kinase A (PKA)-anchoring inhibitor peptides (Vijayaraghavan, Goueli, Davey, & Carr, 1997), or might have digested semenogelin I (Lilja et al., 1989; Robert & Gagnon, 1999) or its binding partner EPPIN (Silva et al., 2013). Targeted research beyond the scope of this report would be needed to identify chymotrypsin's specific effects on membrane channels, cAMP activity, and/or tyrosine phosphorylation.

Assessment of male fertility is plagued by the inability to assess sperm function—namely, the ability of a individual's sperm to fertilize an egg (Oehninger et al., 2014; Wang & Swerdloff, 2014). A simple diagnostic assay would provide a needed functional complement to the descriptive assessments of traditional semen evaluations (World Health Organization, 2010). Identifying sperm with deficiencies in fertilizing ability will allow for a more specific understanding of what is now categorized as “idiopathic infertility.” Of much greater practical importance, such a physiological assessment would enable a clinician to effectively counsel a couple toward the most appropriate form of assisted reproduction to achieve pregnancy. To meet this pressing clinical need, many assays of sperm function have been suggested—e.g., hamster zona pellucida-penetration assays (Barros, Gonzalez, Herrera, & Bustos-Obregon, 1979; Rogers et al., 1979), sperm-zona pellucida-binding tests (Liu, Garrett, & Baker, 2004), and cervical mucus penetration assays (Alexander, 1981; Eggert-Kruse, Leinhos, Gerhard,

Tilgen, & Runnebaum, 1989; Menge & Beitner, 1989)—but their use in the clinic is limited by the difficulty in obtaining needed materials in a logistically practical fashion. Filling the current void, data presented here validate the Cap-Score as an assay that can determine the ability of sperm to undergo the physiological changes required to fertilize an egg. Complementing these findings, the clinical utility of the Cap-Score assay was also tested in an independent, post-hoc clinical trial, in which it was found that capacitation status strongly tracked with a man's history of fertility (Cardona et al., 2017).

Even when standardized according to the recommendations of national or international organizations, the methods traditionally used for semen analysis remain subjective and variable (Auger et al., 2000; Jørgensen et al., 2001). Therefore, it has been recommended that internal and external quality controls be developed and the variations observed within and between persons performing semen analysis be evaluated to reduce confounding the assessment of semen quality (Cooper, Neuwinger, Bahrs, & Nieschlag, 1992; Matson, 1995; Mortimer, Shu, & Tan, 1986; Neuwinger, Behre, & Nieschlag, 1990). In contrast, the data presented in the current study demonstrate that the Cap-Score is highly reproducible and reliable within and between readers. The data and image files acquired could serve as a foundation for the continued quality control and quality assurance in the evaluation of Cap-Score. For example, 2 of the 20 image files, one each from the "lower Cap-Scores" and "presumed fertile" groups could be selected at random and scored each day to demonstrate a reader's daily ability to read Cap-Scores. If values are obtained that are outside of acceptable ranges from the established mean (Westgard, Barry, Hunt, & Groth, 1981), the laboratory director can be consulted for remediation. These data can also be used to track individual readers over time and to identify potential changes in Cap-Score determination. Similarly, as new personnel are trained and incorporated into the reading rotation, their reading ability can be evaluated by scoring multiple image files and comparing their Cap-Scores to established values. Only through continued internal and external quality assurance and quality control will high standards of evaluating sperm function be maintained.

## 4 | MATERIALS AND METHODS

### 4.1 | Specimen collection

All procedures were approved by the Western Institutional Review Board (<https://www.wirb.com/Pages/Default.aspx>; Protocol #20152233). Semen samples were collected by manual masturbation from consenting men with known fertility (evidence of fathering a child within the past 3 years or current pregnancy), after a minimum of 2 and a maximum of 5 days of sexual abstinence. Any samples having fewer than  $10 \times 10^6$  motile sperm were discarded from this study.

### 4.2 | Standard sample processing

Ejaculates within a sealed specimen container (Fisher Scientific, 14-375-462) were liquefied at 37°C in an air incubator, and then

removed from the seminal plasma by centrifugation through Enhance S-Plus Cell Isolation Media (Vitrolife, reference: 15232 ESP-100-90%) at 300g for 10 min. The cells were collected, resuspended in ~4 ml of Modified Human Tubal Fluid medium (mHTF) (Irvine Scientific; reference 90126 [97.8 mM NaCl; 4.69 mM KCl; 0.20 mM MgSO<sub>4</sub>; 0.37 mM KH<sub>2</sub>PO<sub>4</sub>; 2.04 mM CaCl<sub>2</sub>; 4 mM NaHCO<sub>3</sub>; 21 mM HEPES; 2.78 mM C<sub>6</sub>H<sub>12</sub>O<sub>6</sub>; 0.33 mM sodium pyruvate; 21.4 mM sodium lactate; 10 µg/ml gentamicin; 5 mg/L phenol red]), and pelleted at 600g for 10 min. This washing protocol was optimized through extensive preliminary studies (data not shown) to remove seminal plasma, which can interfere with assessment of G<sub>M1</sub> localization. The sperm were resuspended in mHTF with (Cap) and without (Non-Cap) a stimulus for capacitation, and incubated for 3 hr at 37°C. The capacitation stimulus consisted of 2-hydroxypropyl-β-cyclodextrin (Sigma; reference C0926) (Osheroff et al., 1999; Parinaud, Vieitez, Vieu, Collet, & Perret, 2000). Pilot studies showed that this stimulus was as effective in promoting capacitation in human sperm, as measured by G<sub>M1</sub> localization patterns, at 3 hr of incubation as albumin was at 6 hr (Vairo et al., 2013). Following incubation, all samples were fixed with paraformaldehyde (Electron Microscopy Sciences; Hatfield, PA), as previously described (Selvaraj et al., 2006). The fixed samples were maintained at room temperature overnight prior to labeling.

### 4.3 | Cap-Score and acrosome exocytosis

Two treatments were prepared for each of 10 semen samples. Cap and Non-Cap treatments were incubated, respectively, with and without 2-hydroxypropyl-β-cyclodextrin for 3 hr. For seven of these samples, a third treatment, Cap + ionophore, was prepared in which sperm were incubated with 2-hydroxypropyl-β-cyclodextrin for 2.75 hr, then the calcium ionophore A23187 (Sigma-Aldrich, Allentown, PA; reference C7522) was added to a final concentration of 20 µM and the cells were incubated for another 0.25 hr.

In a second set of experiments, Cap ( $n = 4$ ) and Non-Cap ( $n = 5$ ) treatments were incubated, respectively, with and without 2-hydroxypropyl-β-cyclodextrin for 3 hr, plus a third treatment of Non-Cap + ionophore was prepared. For this treatment, sperm were incubated in basal non-capacitating media for 2.75 hr, then the calcium ionophore A23187 (Sigma-Aldrich) was added to a final concentration of 20 µM and the cells were incubated for another 0.25 hr.

Following incubation, the sperm were attached to slides for 0.25 hr, labeled for 10 min with 10 µg/ml of Alexa Fluor® 647-conjugated PNA from *Arachis hypogaea* (Thermo Fisher, Allentown, PA; reference L32460), washed 1× with mHTF, fixed for 0.5 hr, and then labeled with 2 µg/ml of Alexa Fluor 488-conjugated CTB (Thermo Fisher, reference C34775). All labeling and slide work was done in a humidified chamber maintained at 37°C.

### 4.4 | Calculating sperm motility

The percentage of motile sperm was evaluated by counting non-motile and motile sperm in 10 squares on a Makler chamber (Sefi-Medical Instruments; Haifa, Israel), and then the number of

motile sperm was divided by the total number of sperm counted. Motility was determined four separate times for each treatment. This was achieved by having two unique samples of the treatment assessed by two different readers. The average of the four motility readings was recorded after liquefaction (initial), after standard processing (post-wash), and after the 3-hr incubation (post-incubation).

#### 4.5 | Varying time of liquefaction

Ejaculates of 10 fertile individuals were split into two portions. The first portion was liquefied for 0.25 hr while the second portion was liquefied for 1.25 hr in a sealed 15-ml polypropylene conical tube (FALCON; reference 352096) placed in a 37°C water bath. In later experiments, nine ejaculates were liquefied for 0.25 and 2 hr. All tests of liquefaction duration were performed at 37°C.

#### 4.6 | Mechanical liquefaction

Five samples from four fertile men were split into three portions. The first portion was liquefied for 0.25 hr (standard processing control). The second portion was liquefied for 0.25 hr then diluted with pre-warmed mHTF (1:1 dilution). After the addition of mHTF, the sample was gently pulled into and expelled 10 times from a non-sterile Pasteur pipette (VWR; reference 14672-380). The third portion was liquefied for 0.25 hr, diluted with an equal volume of mHTF, and then gently pulled into and expelled 10 times from a non-sterile wide orifice transfer pipette (VWR; reference 14670-147).

#### 4.7 | Liquefaction with chymotrypsin

Samples from five fertile men were split into three portions. The first portion was liquefied following standard processing methods (control); the second portion was liquefied with chymotrypsin (VitroLife; reference 15524) according to the manufacturer's direction, using a fixed 5 mg/ejaculate for 0.25 hr; and the third portion was liquefied following standard processing methods, but was incubated for 3 hr in the presence of 3 mg/ml chymotrypsin.

#### 4.8 | Liquefaction with bromelain

Samples from five fertile men were split into two portions. The first portion was liquefied for 0.25 hr at 37°C; and the second portion was liquefied with bromelain (Sigma; reference B5144-100UN), as recommended by the World Health Organization (2010), for 0.25 hr at 37°C (10 units/ml). Both portions were then processed following standard operating procedures.

#### 4.9 | Sample labeling

Following incubation and fixation, liquefied samples were labeled with 2 µg/ml of Alexa Fluor 488-conjugated CTB. After 10 min, 5 µl of the labeled sperm was placed on a microscope slide, overlaid with a cover slip, and moved to an imaging station.

#### 4.10 | Image acquisition

Imaging was performed on Nikon Eclipse NI-E microscopes equipped with CFI60 Plan Apochromat Lambda 40× Objectives; C-FL AT GFP/FITC Long-Pass Filter Sets; Hamamatsu ORCA-Flash 4.0 cameras; H101F—ProScan III Open Frame Upright Motorized H101F Flat Top Microscope Stages; and 64-bit imaging workstations running NIS Elements software (Nikon; Melville, NY). For the reliability studies, these systems were programmed to automatically capture sets of 15 × 15 stitched images containing up to 5,000 sperm. For the acrosome-exocytosis studies, the C-FL AT GFP/FITC Long-Pass Filter Set was modified to have a band-pass emission filter (ET535/50M). A C-FL At Cy5/Fluor 647/Draq5 Filter set in combination with a ND16 filter were employed to visualize the Alexa Fluor 647.

#### 4.11 | Cap-Score determination

Two independent readers were trained to identify  $G_{M1}$  localization patterns associated with capacitation of human sperm. The proportion of sperm within a sample having undergone capacitation was determined and reported as the Cap-Score (# of sperm with capacitation patterns/(# of sperm with capacitation patterns + number of sperm with other patterns)). For additional tests, readers were trained to identify  $G_{M1}$  localization patterns and were required to determine the Cap-Score for 10 of the stitched images. New readers had to be within 2 standard deviations of the mean Cap-Score established by the original two readers to qualify for scoring.

#### 4.12 | Statistical analyses

Power Analysis was done using  $G^+$ power (Faul, Erdfelder, Lang, & Buchner, 2007). The following four quantities have an intimate relationship: (i) sample size; (ii) effect size; (iii) significance level = probability (Type I error) = incorrect rejection of a true null hypothesis (a "false positive"); and (iv) power = 1 - probability (Type II error) = incorrectly retaining a false null hypothesis ("false negative"). Given any three, the fourth can be determined. For the purposes of this study, power analysis was used to estimate the number of images to sample for evaluating Cap-Score accuracy.

Student's *t*-test was performed using Microsoft Excel (2013), and was used to compare Cap-Score means between two readers. In this example 20 different means were compared, increasing the chance of incorrectly rejecting the null hypothesis that there is no difference between the means (i.e., making a Type I error). A Bonferroni correction was applied to counteract this problem. Similarly, Student's *t*-test was used to compare the Cap-Scores for liquefaction timing, mechanical liquefaction, and enzymatic liquefaction using the online statistical calculator in silico (*t*-test; available online at: <http://in-silico.net/tools/statistics/ttest>).

Fisher's LSD was used for comparing means in the acrosome-exocytosis data, and was carried out in XLSTAT (Version 2015.5.01.22912). Linear Regression and Bartlett's test of homoscedasticity were carried out in XLSTAT (Version 2015.5.01.22912). The non-parametric Mann-Whitney test and Kruskal-Wallis H test were

calculated for enzymatic liquefaction motilities by using XLSTAT (Version 2015.6.01.25740).

## ACKNOWLEDGMENT

This study was funded by Androvia LifeSciences.

## DISCLOSURES

Melissa A. Moody, Alana J. Simpson, Cristina Cardona, and G. Charles Ostermeier are employees of Androvia LifeSciences, LLC. T. Timothy Smith is a consultant working under contract for Androvia regarding validation of the laboratory developed test and laboratory certification. Alexander J. Travis and his laboratory at Cornell University developed the technology which Androvia licensed. He serves as a consultant to Androvia LifeSciences with duties of Chief Scientific Officer.

## AUTHORS' CONTRIBUTION

MM and GCO analyzed the data. MM, AT, and GCO wrote the paper. AS, CC, MM, and GCO performed the research. MM, CC, AS, TTS, AT, and GCO designed the research study.

## REFERENCES

- Aboulghar, M., Mansour, R., Serour, G., Abdrazek, A., Amin, Y., & Rhodes, C. (2001). Controlled ovarian hyperstimulation and intrauterine insemination for treatment of unexplained infertility should be limited to a maximum of three trials. *Fertility and Sterility*, 75(1), 88–91.
- Alexander, N. J. (1981). Evaluation of male infertility with an in vitro cervical mucus penetration test. *Fertility and Sterility*, 36(2), 201–208.
- Alvarez, J. G., Lasso, J. L., Blasco, L., Nuñez, R. C., Heyner, S., Caballero, P. P., & Storey, B. T. (1993). Centrifugation of human spermatozoa induces sublethal damage; separation of human spermatozoa from seminal plasma by a dextran swim-up procedure without centrifugation extends their motile lifetime. *Human Reproduction*, 8(7), 1087–1092.
- Asano, A., Nelson-Harrington, J. L., & Travis, A. J. (2013). Phospholipase B is activated in response to sterol removal and stimulates acrosome exocytosis in murine sperm. *Journal of Biological Chemistry*, 288(39), 28104–28115.
- Auger, J., Eustache, F., Ducot, B., Blandin, T., Daudin, M., Diaz, I., ... Kolbezen, M. (2000). Intra- and inter-individual variability in human sperm concentration, motility and vitality assessment during a workshop involving ten laboratories. *Human Reproduction*, 15(11), 2360–2368.
- Austin, C. R. (1952). The capacitation of the mammalian sperm. *Nature*, 170(4321), 326.
- Balasuriya, A., Serhal, P., Doshi, A., & Harper, C. (2012). Processes involved in assisted reproduction technologies significantly increase sperm DNA fragmentation and phosphatidylserine translocation. *Andrologia*, 46(2), 86–97.
- Baldi, E., Casano, R., Falsetti, C., Krausz, C., Maggi, M., & Forti, G. (1991). Intracellular calcium accumulation and responsiveness to progesterone in capacitating human spermatozoa. *Journal of Andrology*, 12(5), 323–330.
- Barros, C., Gonzalez, J., Herrera, E., & Bustos-Obregon, E. (1979). Human sperm penetration into zona-free hamster oocytes as a test to evaluate the sperm fertilizing ability. *Andrologia*, 11(3), 197–210.
- Bedu-Addo, K., Lefièvre, L., Moseley, F., Barratt, C., & Publicover, S. (2005). Bicarbonate and bovine serum albumin reversibly 'switch' capacitation-induced events in human spermatozoa. *Molecular Human Reproduction*, 11(9), 683–691.
- Burd, E. M. (2010). Validation of laboratory-developed molecular assays for infectious diseases. *Clinical Microbiology Reviews*, 23(3), 550–576.
- Cardona, C., Neri, Q. V., Simpson, A. J., Moody, M. A., Ostermeier, G. C., Seaman, E. K., ... Travis, A. J. (2017). Localization patterns of the ganglioside GM1 in human sperm are indicative of male fertility and independent of traditional semen measures. *Molecular Reproduction and Development*, 9999, 1–13. <https://doi.org/10.1002/mrd.22803>
- Chang, M. C. (1951). Fertilizing capacity of spermatozoa deposited into the fallopian tubes. *Nature*, 168(4277), 697–698.
- Chen, F., Lu, J., Xu, H., Huang, Y., & Lu, N. (2006). Chymotrypsin effects on the determination of sperm parameters and seminal biochemistry markers. *Clinical Chemistry and Laboratory Medicine*, 44(11), 1335–1339.
- Chobotova, K., Vernallis, A. B., & Majid, F. A. A. (2010). Bromelain's activity and potential as an anti-cancer agent: Current evidence and perspectives. *Cancer Letters*, 290(2), 148–156.
- Cleophas, T. J., & Zwinderman, A. H. (2012). Validating qualitative diagnostic tests. *Statistics applied to clinical studies* (pp. 509–517). Dordrecht, Heidelberg, London, New York: Springer Science & Business Media.
- Cohen, R., Buttke, D. E., Asano, A., Mukai, C., Nelson, J. L., Ren, D., ... Travis, A. J. (2014). Lipid modulation of calcium flux through CaV2.3 regulates acrosome exocytosis and fertilization. *Developmental Cell*, 28(3), 310–321.
- Cohen, R., Mukai, C., & Travis, A. (2016). Lipid regulation of acrosome exocytosis. In M. G. Buffone (Ed.), *Sperm acrosome biogenesis and function during fertilization* (pp. 107–127). Springer.
- Cooper, T., Neuwinger, J., Bahrs, S., & Nieschlag, E. (1992). Internal quality control of semen analysis. *Fertility and Sterility*, 58(1), 172–178.
- Cross, N. L. (1998). Role of cholesterol in sperm capacitation. *Biology of Reproduction*, 59(1), 7–11.
- Eggert-Kruse, W., Leinhos, G., Gerhard, I., Tilgen, W., & Runnebaum, B. (1989). Prognostic value of in vitro sperm penetration into hormonally standardized human cervical mucus. *Fertility and Sterility*, 51(2), 317–323.
- Fakih, H., MacLusky, N., DeCherney, A., Wallimann, T., & Huszar, G. (1986). Enhancement of human sperm motility and velocity in vitro: Effects of calcium and creatine phosphate. *Fertility and Sterility*, 46(5), 938–944.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.
- Gadella, B., & Van Gestel, R. (2004). Bicarbonate and its role in mammalian sperm function. *Animal Reproduction Science*, 82, 307–319.
- JCGM/WG2 (Joint Committee for Guides on Metrology; Working Group 2). (2008). ISO/IEC Guide 200:2008, The international vocabulary of metrology (VIM) – basic and general concepts and associated terms. p. 90.
- Jin, M., Fujiwara, E., Kakiuchi, Y., Okabe, M., Satouh, Y., Baba, S. A., ... Hirohashi, N. (2011). Most fertilizing mouse spermatozoa begin their acrosome reaction before contact with the zona pellucida during in vitro fertilization. *Proceedings of the National Academy of Sciences*, 108(12), 4892–4896.
- Jørgensen, N., Andersen, A.-G., Eustache, F., Irvine, D. S., Suominen, J., Petersen, J. H., ... Horte, A. (2001). Regional differences in semen quality in Europe. *Human Reproduction*, 16(5), 1012–1019.
- Katkov, I. I., & Mazur, P. (1998). Influence of centrifugation regimes on motility, yield, and cell associations of mouse spermatozoa. *Journal of Andrology*, 19(2), 232–241.
- Kim, K.-S., & Gerton, G. L. (2003). Differential release of soluble and matrix components: Evidence for intermediate states of secretion during spontaneous acrosomal exocytosis in mouse sperm. *Developmental Biology*, 264(1), 141–152.

- Kim, K. S., Foster, J. A., Kvasnicka, K. W., & Gerton, G. L. (2011). Transitional states of acrosomal exocytosis and proteolytic processing of the acrosomal matrix in guinea pig sperm. *Molecular Reproduction and Development*, 78(12), 930–941.
- Knuth, U., Neuwinger, J., & Nieschlag, E. (1989). Bias to routine semen analysis by uncontrolled changes in laboratory environment—detection by long-term sampling of monthly means for quality control. *International Journal of Andrology*, 12(5), 375–383.
- Kumar, N., & Singh, A. K. (2015). Trends of male factor infertility, an important cause of infertility: A review of literature. *Journal of Human Reproductive Sciences India*, 8(4), 191–196.
- Kussler, A. P. S., Pimentel, A. M., Alcoba, D. D., Liu, I. P., Brum, I. S., Capp, E., & Corleta, V. H. (2014). Mechanical processing of hyperviscous semen specimens can negatively affect sperm DNA fragmentation. *International Urology and Nephrology*, 46(4), 737–742.
- Lamb, D. J. (2010). Semen analysis in 21st century medicine: The need for sperm function testing. *Asian Journal of Andrology*, 12(1), 64–70.
- Lilja, H., Abrahamsson, P. A., & Lundwall, A. (1989). Semenogelin, the predominant protein in human semen. Primary structure and identification of closely related proteins in the male accessory sex glands and on the spermatozoa. *The Journal of Biological Chemistry*, 264(3), 1894–1900.
- Liu, D. Y., Garrett, C., & Baker, H. G. (2004). Clinical application of sperm-oocyte interaction tests in in vitro fertilization-embryo transfer and intracytoplasmic sperm injection programs. *Fertility and Sterility*, 82(5), 1251–1263.
- Matson, P. (1995). Andrology: External quality assessment for semen analysis and sperm antibody detection: Results of a pilot scheme. *Human Reproduction*, 10(3), 620–625.
- Maurer, H. (2001). Bromelain: Biochemistry, pharmacology and medical use. *Cellular and Molecular Life Sciences CMLS*, 58(9), 1234–1245.
- Menge, A. C., & Beitner, O. (1989). Interrelationships among semen characteristics, antisperm antibodies, and cervical mucus penetration assays in infertile human couples. *Fertility and Sterility*, 51(3), 486–492.
- Metzig, C., Grabowska, E., Eckert, K., Rehse, K., & Maurer, H. R. (1999). Bromelain proteases reduce human platelet aggregation in vitro, adhesion to bovine endothelial cells and thrombus formation in rat vessels in vivo. *In Vivo*, 13(1), 7–12.
- Miller, D. J., Winer, M. A., & Ax, R. L. (1990). Heparin-binding proteins from seminal plasma bind to bovine spermatozoa and modulate capacitation by heparin. *Biology of Reproduction*, 42(5–6), 899–915.
- Monaco, D., Fatnassi, M., Padalino, B., Hammadi, M., Khorchani, T., & Lacalandra, G. M. (2016). Effect of alpha-amylase, papain, and spermfluid(R) treatments on viscosity and semen parameters of dromedary camel ejaculates. *Research in Veterinary Science*, 105, 5–9.
- Mortimer, D. (1994). Sperm recovery techniques to maximize fertilizing capacity. *Reproduction, Fertility and Development*, 6(1), 25–31.
- Mortimer, D., Curtis, E., & Miller, R. (1987). Specific labelling by peanut agglutinin of the outer acrosomal membrane of the human spermatozoon. *Journal of Reproduction and Fertility*, 81(1), 127–135.
- Mortimer, D., Shu, M. A., & Tan, R. (1986). Standardization and quality control of sperm concentration and sperm motility counts in semen analysis. *Human Reproduction*, 1(5), 299–303.
- Mortimer, S. T., Swan, M. A., & Mortimer, D. (1998). Effect of seminal plasma on capacitation and hyperactivation in human spermatozoa. *Human Reproduction*, 13(8), 2139–2146.
- Neuwinger, J., Behre, H., & Nieschlag, E. (1990). External quality control in the andrology laboratory: An experimental multicenter trial. *Fertility and Sterility*, 54(2), 308–314.
- Oehninger, S., Franken, D. R., & Ombelet, W. (2014). Sperm functional tests. *Fertility and Sterility*, 102(6), 1528–1533.
- Osheroff, J. E., Visconti, P. E., Valenzuela, J. P., Travis, A. J., Alvarez, J., & Kopf, G. S. (1999). Regulation of human sperm capacitation by a cholesterol efflux-stimulated signal transduction pathway leading to protein kinase A-mediated up-regulation of protein tyrosine phosphorylation. *Molecular Human Reproduction*, 5(11), 1017–1026.
- Parinaud, J., Vieitez, G., Vieu, C., Collet, X., & Perret, B. (2000). Enhancement of zona binding using 2-hydroxypropyl- $\beta$ -cyclodextrin. *Human Reproduction*, 15(5), 1117–1120.
- Petrunkina, A. M., Harrison, R. A., Tsoлова, M., Jebe, E., & Topfer-Petersen, E. (2007). Signalling pathways involved in the control of sperm cell volume. *Reproduction*, 133(1), 61–73.
- Petrunkina, A. M., Jebe, E., & Topfer-Petersen, E. (2005). Regulatory and necrotic volume increase in boar spermatozoa. *Journal of Cellular Physiology*, 204(2), 508–521.
- Practice Committee of the American Society for Reproductive Medicine. (2006). Effectiveness and treatment for unexplained infertility. *Fertility and Sterility*, 86(5 Suppl 1), S111–S114.
- Robert, M., & Gagnon, C. (1999). Semenogelin I: A coagulum forming, multifunctional seminal vesicle protein. *Cellular and Molecular Life Sciences*, 55(6–7), 944–960.
- Rogers, B., Van Campen, H., Ueno, M., Lambert, H., Bronson, R., & Hale, R. (1979). Analysis of human spermatozoal fertilizing ability using zona-free ova. *Fertility and Sterility*, 32(6), 664–670.
- Said, L., Galeraud-Denis, I., Carreau, S., & Saad, A. (2009). Relationship between semen quality and seminal plasma components: Alpha-glucosidase, fructose and citrate in infertile men compared with a normospermic population of Tunisian men. *Andrologia*, 41(3), 150–156.
- Selvaraj, V., Asano, A., Buttke, D. E., McElwee, J. L., Nelson, J. L., Wolff, C. A., . . . Lisanti, M. P. (2006). Segregation of micron-scale membrane sub-domains in live murine sperm. *Journal of Cellular Physiology*, 206(3), 636–646.
- Selvaraj, V., Asano, A., Buttke, D. E., Sengupta, P., Weiss, R. S., & Travis, A. J. (2009). Mechanisms underlying the micron-scale segregation of sterols and GM1 in live mammalian sperm. *Journal of Cellular Physiology*, 218(3), 522–536.
- Selvaraj, V., Buttke, D. E., Asano, A., McElwee, J. L., Wolff, C. A., Nelson, J. L., . . . Travis, A. J. (2007). GM1 dynamics as a marker for membrane changes associated with the process of capacitation in murine and bovine spermatozoa. *Journal of Andrology*, 28(4), 588–599.
- Shao, Y., Ge, Y., Zhao, X., Yao, B., Lu, J., Cui, Y., . . . Huang, Y. (2010). Effects of semen analysis on human sperm movement parameters at different times after semen collection. *National Journal of Andrology*, 16(7), 631–634.
- Silva, E. J., Hamil, K. G., & O’Rand, M. G. (2013). Interacting proteins on human spermatozoa: Adaptive evolution of the binding of semenogelin I to EPPIN. *PLoS ONE*, 8(12), e82014.
- Taussig, S. J., & Batkin, S. (1988). Bromelain, the enzyme complex of pineapple (*Ananas comosus*) and its clinical application. An update. *Journal of Ethnopharmacology*, 22(2), 191–203.
- Taylor, J. R., & Cohen, E. (1998). An introduction to error analysis: The study of uncertainties in physical measurements. *Measurement Science and Technology*, 9(6), 1015.
- Tournaye, H. (2012). Male factor infertility and ART. *Asian Journal of Andrology*, 14(1), 103–108.
- Turner, R. M. (2006). Moving to the beat: A review of mammalian sperm motility regulation. *Reproduction, Fertility, and Development*, 18(1–2), 25–38.
- Vairo, L., Neri, Q., Rosenwaks, Z., Schlegel, P., Travis, A., & Palermo, G. (2013). A novel, biomarker-based assay to screen for dysfunctional spermatozoa. *Fertility and Sterility*, 100(3), S224–S225.

- Varisli, O., Uguz, C., Agca, C., & Agca, Y. (2009). Various physical stress factors on rat sperm motility, integrity of acrosome, and plasma membrane. *Journal of Andrology*, 30(1), 75–86.
- Vázquez, J. M., Martínez, E., Pastor, L. M., Roca, J., Matas, C., & Calvo, A. (1996). Lectin histochemistry during in vitro capacitation and acrosome reaction in boar spermatozoa: New lectins for evaluating acrosomal status of boar spermatozoa. *Acta Histochemica*, 98(1), 93–100.
- Vijayaraghavan, S., Goueli, S. A., Davey, M. P., & Carr, D. W. (1997). Protein kinase A-anchoring inhibitor peptides arrest mammalian sperm motility. *The Journal of Biological Chemistry*, 272(8), 4747–4752.
- Wang, C., & Swerdloff, R. S. (2014). Limitations of semen analysis as a test of male fertility and anticipated needs from newer tests. *Fertility and Sterility*, 102(6), 1502–1507.
- Westgard, J. O., Barry, P. L., Hunt, M. R., & Groth, T. (1981). A multi-rule Shewhart chart for quality control in clinical chemistry. *Clinical Chemistry*, 27(3), 493–501.
- World Health Organization. (1999). *WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction* (4th ed.). Cambridge, UK: Cambridge University Press.
- World Health Organization. (2010). *WHO laboratory manual for the examination and processing of human semen* (5th ed., pp. 1–271). Geneva 27, Switzerland: World Health Organization.
- Xu, H.-R., Lu, J.-C., Chen, F., Huang, Y.-F., Yao, B., & Lu, N.-Q. (2006). The effect of chymotrypsin on the determination of total alpha-glucosidase activity in seminal plasma and the correlation between alpha-glucosidase level and semen parameters. *Archives of Andrology*, 52(6), 441–446.
- Yeung, C. H., Barfield, J. P., Anapolski, M., & Cooper, T. G. (2004). Volume regulation of mature and immature spermatozoa in a primate model, and possible ion channels involved. *Human Reproduction*, 19(11), 2587–2593.
- Yeung, C. H., Callies, C., Rojek, A., Nielsen, S., & Cooper, T. G. (2009). Aquaporin isoforms involved in physiological volume regulation of murine spermatozoa. *Biology of Reproduction*, 80(2), 350–357.
- Yeung, C. H., Callies, C., Tuttelmann, F., Kliesch, S., & Cooper, T. G. (2010). Aquaporins in the human testis and spermatozoa – identification, involvement in sperm volume regulation and clinical relevance. *International Journal of Andrology*, 33(4), 629–641.
- Yeung, C. H., & Cooper, T. G. (2010). Aquaporin AQP11 in the testis: Molecular identity and association with the processing of residual cytoplasm of elongated spermatids. *Reproduction*, 139(1), 209–216.
- Zöpfgen, A., Priem, F., Sudhoff, F., Jung, K., Lenk, S., Loening, S., & Sinha, P. (2000). Relationship between semen quality and the seminal plasma components carnitine, alpha-glucosidase, fructose, citrate and granulocyte elastase in infertile men compared with a normal population. *Human Reproduction*, 15(4), 840–845.

**How to cite this article:** Moody MA, Cardona C, Simpson AJ, Smith TT, Travis AJ, Ostermeier GC. Validation of a laboratory-developed test of human sperm capacitation. *Mol Reprod Dev*. 2017;84:408–422.  
<https://doi.org/10.1002/mrd.22801>