Seminal Plasma ExLncRNA Pairs: Updating Perspectives in the Search for Testicular Spermatozoa Retrieval Biomarkers in Nonobstructive Azoospermia Patients with mTESE by WGCNA

Haiming Cao[1]#, Chuntao Wang[1]#, Ruilin Cai[2], Zi Wan[3]*, Lin Ma[4]**

Purpose: This study aims to find candidates for testicular spermatozoa retrieval biomarkers among the seminal plasma exLncRNA pairs.

Materials and Methods: A set of exLncRNA pairs with the best potential biomarkers was selected and validated in 96 NOA samples. Weighted correlation network analysis (WGCNA) and Least Absolute Shrinkage and Selection Operator were used to identify possible biomarkers for these pairs (LASSO). These pairs' potential biomarkers were identified using receiver operating curves. Confusion matrices and sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), FP, false-negative rates (FNR), and F1 scores are calculated. Through F1 scores, we selected the best threshold value.

Results: The relative differential expression of each pair in testicular spermatozoa retrieval (+) and testicular spermatozoa retrieval (-) men were validated. The six pairs displayed the best biomarker potential. Among them, CCDC37.DT-LOCI00505685 pair and LOC440934- LOCI01929088 (XR_001745218.1) pair showed the most significant potential and stability for detecting testicular spermatozoa retrieval in the selected and validated cohort.

Conclusion: CCDC37.DT-LOCI00505685 pair and LOC440934- LOCI01929088 (XR_001745218.1) pair have the potential to become new molecular biomarkers that could help to select clinical strategies for microdissection testicular sperm extraction.

Keywords: nonobstructive azoospermia, seminal plasma, LncRNA, prediction

INTRODUCTION

In the case of azoospermia, there is no sperm in the ejaculate. The total number of males with Azoospermia is approximately 1%, with the condition impacting approximately 15% of all infertile males[(1)]. It is broadly classified into two types (nonobstructive azoospermia (NOA) and obstructive azoospermia (OA)). The incidence of NOA is higher than that of OA, at almost 60% [(2)]. Testicular causes of NOA include congenital causes such as Klinefelter syndrome, Y chromosome microdeletions, cryptorchidism, postnatal factors such as exposure to radiotherapy and chemotherapy, genital trauma, and infectious diseases such as mumps orchitis and SARS-CoV-2[(3)]. Furthermore, more than 15% of cases of NOA are idiopathic[(4)].

Most cases of NOA are related to spermatogenic failure, which is not curable.

The NOA patients who achieved fertility need to seek the sole technique which is intracytoplasmic sperm injection (ICSI) combined with surgical sperm retrieval (SR) techniques. For surgical sperm retrieval (SR), various techniques contain conventional testicular sperm extraction (cTESE), testicular sperm aspiration (TESA), and microsurgical testicular sperm extraction (micro-TESE). Among them, the SR rate with micro-TESE in NOA improved to 77%[(5-8)].

Various factors influenced the success rate of SR techniques in NOA patients in couples pursuing fertility treatment. To some extent, the experience of surgeons and laboratory expertise influence SR success. Among them, SR predictors of pre-operation include clinical parameters, laboratory parameters, and the utility of adjuvant therapy before SR. By using ejaculated sperm, studies have shown that younger men can have a better chance of conceiving than older men[(9)]. Seminiferous tubules account for approximately 75–85 % of the mass of the testes[(10)]. In general, testicular volume indicates spermatogenetic activity, with lower testicular sizes corresponding to lower SR rates. Historical studies have regarded the FSH level as a testicular function marker [(10-13)]. Inhibin B levels have shown a certain degree of sensitivity and specificity in predicting SR suc-

[1]The Andrology department, Reproductive Medicine Center, the Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen 518000, PR China.
[2]The Clinical medicine department, the Medical School of Sun Yat-sen University, Shenzhen 518000, PR China.
[3]The Andrology department, the First Affiliated Hospital of Sun Yat-sen University, Guangzhou 510080,PR China.
[4]Reproductive Medicine Center, the Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen 518000, PR China.
# The authors contributed equally to this work.
* Correspondence: 58 Zhongshan Second Road Guangzhou, The Andrology Department, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou 510080, People's Republic of China. zi_wan2012@163.com.
**NO.628 of Zhenyuan Road,518000.The reproductive medicine center, the Seventh Affiliated Hospital of Sun Yat-sen University, Shenzhen 518000, PR China.E mail: malin8@mail.sysu.edu.cn.
Received June 2022 & Accepted February 2023

Table 1. Characteristics of NOA

characteristics	sperm(-)	sperm(+)	*P* value
n	32	64	
Age (mean (SD))	28.75 (3.86)	33.22 (7.85) 0.003	
FSH (IU/I) (median [IQR])	20.13 [14.20, 31.57]	12.86 [5.94, 23.80]	0.001
LH (IU/I) (median [IQR])	10.04 [7.01, 13.72]	7.90 [5.07, 11.62]	0.016
Total testosterone (ng/ml) (median [IQR])	3.00 [2.20, 4.12]	3.91 [2.24, 5.78]	0.047
Inhibin B (pg/ml) (median [IQR])	7.98 [5.28, 15.36]	31.44 [16.20, 60.16]	0.014
Risk disease of NOA (%)			< 0.001
AZFc microdeletions	0 (0.0)	4 (6.2)	
Cryptorchidism-associated NOA	0 (0.0)	13 (20.3)	
Idiopathic NOA	8 (25.0)	5 (7.8)	
Idiopathic testicular atrophy	9 (28.1)	14 (21.9)	
Klinefelter's syndrome	5 (15.6)	2 (3.1)	
Mumps orchitis	0 (0.0)	13 (20.3)	
Partial AZFb + c microdeletions	1 (3.1)	0 (0.0)	
Partial AZFb microdeletions	0 (0.0)	1 (1.6)	
Varicocele-associated NOA	9 (28.1)	12 (18.8)	

Abbreviations: SD, standard deviation; IQR, interquartile range; NOA, non-obstructive azoospermia.

cess [14,15]. Moreover, many models based on FSH, LH, testosterone, and Inhibin B, may be applied to predict the success SR rates among NOA patients[8,16-18].

Currently, the search for new effective biomarkers of the success rates of SR is driven by the need for alternative diagnostic approaches. The expression level of ZMYND15 and SPEM1 may have the potential for the prediction of successful SR[19,20]. The expression level of these genes (ZMYND15, TNP1, and PRM1) may have the potential for the prediction of successful sperm retrieval[21]. But the analysis may be invasive; we need many non-invasive methods to predict successful SR. The seminal plasma is rich in exosomes. A population of small non-coding RNAs is found in exosomes in semen. The semen exosomes serve as rich noninvasive biomarkers, such as testis-specific RNA and other molecules, which indicate a better pathological process [22]. Long noncoding RNA (LncRNA) is a class of non-coding RNA molecules with a length > 200 nt, which is closely related to the occurrence and development of many human diseases[23]. In recent years, a large number of studies have shown that lncRNA is particular in testicular tissue and can be used as an important regulatory molecule to participate in spermatogenesis[24-26]. Many of them display restricted expression in the testis, suggesting that LncRNAs may be ideal biomarkers for male reproductive system disorders[27,28].

Yet, there exists no data with regards to the expression ratios of sperm LncRNAs about successful SR concerning pairwise comparison. Considering the prevalence of LncRNA expression which is strongly correlated with stability, LncRNA expression profiles might also reflect the stability of these correlations. Infertility biomarkers might be based on 16 LncRNA pairs detected in highly fertile patients with highly stable relative expression[29]. The goal of this research was to find biomarkers that are most suited to assessing successful SR among stable sperm LncRNA pairs. Each group chose the best LncRNA pairings to use in the correct diagnosis. The pairs chosen were validated using an independent cohort.

MATERIALS AND METHODS

Extracellular vesicle LncRNA profiles in seminal plasma

Extracellular vesicle LncRNA profiles from 96 NOA samples were included. The data was downloaded from the previous study[29]. Totally, ninety-six NOA patient data sets were randomly divided into screen (30 patients) and validation (66 patients) set by Approx. 3:7. All NOA patients experienced related medical therapy (such as varicocele repair or drug therapy). According to the committee opinions of the American Society for Reproductive Medicine, non-obstructive azoospermia

Table 2. Characteristics of screen and validation set

character	screen set	validation set	*P* value
n	30	66	
Age (mean (SD))	30.63 (4.90)	32.23 (7.86)	0.309
FSH (IU/I) (median [IQR])	16.74 [12.54, 25.28]	14.41 [6.16, 28.27]	0.402
LH (IU/I) (median [IQR])	8.39 [6.48, 10.96]	8.76 [5.52, 13.18]	0.797
Total testosterone (ng/ml) (median [IQR])	3.75 [2.29, 4.26]	3.58 [2.21, 5.14]	0.915
Inhibin B (pg/ml) (median [IQR])	17.43 [4.58, 32.47]	31.44 [16.86, 73.60]	0.04
Risk disease of NOA (%)			0.133
AZFc microdeletions	3 (10.0)	1 (1.5)	
Cryptorchidism-associated NOA	4 (13.3)	9 (13.6)	
Idiopathic NOA	4 (13.3)	9 (13.6)	
Idiopathic testicular atrophy	8 (26.7)	15 (22.7)	
Klinefelter's syndrome	1 (3.3)	6 (9.1)	
Mumps orchitis	1 (3.3)	12 (18.2)	
Partial AZFb + c microdeletions	1 (3.3)	0 (0.0)	
Partial AZFb microdeletions	1 (3.3)	0 (0.0)	
Varicocele-associated NOA	7 (23.3)	14 (21.2)	

Abbreviations: SD, standard deviation; IQR, interquartile range; NOA, non-obstructive azoospermia.

Table 3. The predictive value of exlncRNA pairs and clinical parameters

characteristics	screen set		validation set		screen set		validation set	
	AUC	95%CI	AUC	95%CI	AUC-adjusted	95%CI	AUC-adjusted	95%CI
Age (Year)	0.739	0.613-0.865	0.574	0.365-0.783	NA	NA	NA	NA
FSH(IU/I)	0.74	0.618-0.863	0.634	0.427-0.841	NA	NA	NA	NA
Total testosterone (ng/ml)	0.617	0.476-0.758	0.639	0.434-0.844	NA	NA	NA	NA
SPATA42-LOCI00505685	0.636	0.491-0.782	0.806	0.646-0.965	0.825	0.724-0.926	0.861	0.723-0.999
SPATA42-LOCI01929088(XR_927561.2)	0.686	0.551-0.82	0.843	0.628-1	0.852	0.762-0.943	0.866	0.737-0.995
CCDC37.DT-LOCI00505685	0.86	0.764-0.957	0.782	0.608-0.957	0.929	0.871-0.988	0.824	0.674-0.974
GABRG3.ASI-LOCI00505685	0.599	0.437-0.762	0.734	0.547-0.921	0.815	0.712-0.918	0.801	0.643-0.959
LOC440934-LOCI00505685	0.634	0.483-0.786	0.75	0.573-0.927	0.842	0.743-0.942	0.875	0.748-1
LOC440934-LOCI01929088(XR_001745218.1)	0.81	0.705-0.914	0.81	0.654-0.967	0.879	0.79-0.968	0.856	0.724-0.989
LOCI01929088(XR_927561.2)-LINC00343	0.778	0.664-0.892	0.736	0.551-0.921	0.915	0.85-0.91	0.819	0.655-0.984

was diagnosed. The age and Sperm concentration (106/mL) of all patients were collected. For patients diagnosed with NOA, experienced andrologists performed a first mTESE. Through patients' clinical history and pathological examination, the risk disease of NOA patients was divided into idiopathic NOA, Klinefelter's syndrome, partial AZFb microdeletions, AZFc microdeletions, partial AZFb b+c microdeletions, cryptorchidism-associated NOA, varicocele-associated NOA, mumps orchitis, idiopathic testicular atrophy (testicular volume < 6 mL). FSH (IU/I), LH (IU/I), total testosterone (ng/mL), inhibin B (pg/mL) were tested for NOA patients. In the long arm of the human Y chromosome (Yq11) are genomic deletions (AZF deletions, azoospermia factor) associated with NOA. AZF mirco-deletions (AZFa, AZFb and AZFc) are caused by intrachromosomal (Yq11) recombinations (deletions, duplications, inversions).[30]

The procedure of mTESE

The mTESE Inclusion criteria: patients diagnosed as NOA wish to assistant reproductive technology. Exclusion criteria: patients diagnosed as NOA refuse mTESE. To gain easy access to both testes, an incision is made in the median raphe. By incising through the dartos and tunica vaginalis, the larger of two testicles are delivered first. Cut the skin, flesh membrane, sheath successively, squeeze out unilateral testis, epididymis, observe the size of the testis and epididymis development is normal. Under the operating microscope with magnification x20, extract some relatively good seminiferous tubules (thick, plump and opaque seminiferous tubules). In the sperm transport buffer, IVF lab doctors ripped the seminiferous tubules under high magnification to observe the presence of sperm. If no sperm was found, the testicle was separated into six regions, and the testicle tissue in the testicle lobule was completely exposed until the tunica albuginea. Each location produced relatively good seminiferous tubules. Testicular tissue was sent to the pathological department for examination. But if sperm is not identified, then dissection of the contralateral side proceeds.

Construct seminal plasma exLncRNA pairs

Based on a previous study (29), 16 exlncRNAs in seminal plasma were determined as testis-specific after evaluating the tissue expression of lncRNAs (supplemental Table 1.1 and 1.2, the primers used for LncRNA in supplemental table 4). The statistical description of exLncRNA pairs with a constant expression was assessed over normalized FPKM values of every possible exLncRNA- exLncRNA combination (R language base function: combn). The difference by subtracting values (normalized exLncRNA- exLncRNA) of the obtained pairs was calculated.

Weighted correlation network analysis (WGCNA)

With the assistance of the R environment, the WGCNA was performed using the WGCNA package.(31) It was determined that the exLncRNA pairings should minimize noise and computational cost while guaranteeing that no significant information is lost. A weighted correlation network analysis was carried out on each exLncRNA pair (screen set) and Pearson correlation coefficients were calculated between how the module correlated with the sperm retrieval (+/-) to identify the module most closely related to the sperm retrieval for further analysis.

Least Absolute Shrinkage and Selection Operator regression (LASSO regression)

Based on the glmnet package (version 4.1-2), the logistic LASSO model is an active selection method in which the variables can be selected from a large and possibly multicollinear set to derive an interpretably relevant set of predictors[32]. The function of LASSO

Table 4. Multiple variables logistics regression analysis

Variables	Multiple varibles OR[95%CI]	*P* value	Inclusion variabls OR[95%CI]	*P* value
Age	1.20 [1.07, 1.37]	0.003	1.16 [1.06, 1.31]	0.004
FSH (IU/I)	0.97 [0.92, 1.01]	0.091	NA	NA
Total testosterone (ng/ml)	1.09 [0.86, 1.52]	0.568	NA	NA
CCDC37.DT-LOCI00505685	0.73 [0.59, 0.88]	0.002	0.70 [0.56, 0.83]	< 0.001
Age	1.22 [1.08, 1.41]	0.003	1.18 [1.06, 1.34]	0.004
FSH (IU/I)	0.97 [0.92, 1.01]	0.121	NA	NA
Total testosterone (ng/ml)	1.16 [0.87, 1.61]	0.356	NA	NA
LOC440934- LOCI01929088 (XR_001745218.1)	0.60 [0.46, 0.77]	< 0.001	0.58 [0.44, 0.73]	< 0.001

p = exp(-3.25+0.152*Age-0.258*CCDC37.DT-LOCI00505685)/(1+exp(-3.25+0.152*Age-0.258*CCDC37.DT-LOCI00505685))
p = exp(-5.33+0.167*Age-0.54*LOC440934- LOCI01929088 (XR_001745218.1))/(1+exp(-5.33+0.167*Age-0.54*LOC440934- LOCI01929088 (XR_001745218.1)))
OR:odds ratio; CI:confident index

Table 5. Thresholds, sensitivities and specificities analysis

Variables	Sensitivities[95%CI]	specificities[95%CI]	thresholds	F1
CCDC37.DT-LOCI00505685	0.703[0.703 ,0.875]	0.813 [0.813 ,0.938]	0.686	0.754
	0.719 [0.703 ,0.875]	0.813 [0.813 ,0.938]	0.684	0.763
	0.734 [0.703 ,0.875]	0.813 [0.781 ,0.938]	0.682	0.771
	0.750 [0.703 ,0.875]	0.813 [0.781 ,0.938]	0.677	0.78
	0.750 [0.750 ,0.906]	0.781 [0.781 ,0.938]	0.671	0.765
	0.750 [0.766 ,0.938]	0.750 [0.781 ,0.938]	0.662	0.75
	0.750 [0.797 ,0.938]	0.719 [0.781 ,0.938]	0.654	0.734
	0.766 [0.797 ,0.938]	0.719 [0.750 ,0.906]	0.653	0.741
LOC440934- LOCI01929088 (XR_001745218.1)	0.609 [0.234 ,0.828]	0.938 [0.844 ,1.000]	0.848	0.739
	0.625 [0.234 ,0.828]	0.938 [0.813 ,1.000]	0.842	0.75
	0.641 [0.234 ,0.828]	0.938 [0.813 ,1.000]	0.825	0.761
	0.656 [0.234 ,0.828]	0.938 [0.781 ,1.000]	0.802	0.772
	0.672 [0.234 ,0.828]	0.938 [0.750 ,1.000]	0.791	0.783
	0.688 [0.234 ,0.828]	0.938 [0.719 ,1.000]	0.788	0.793
	0.703 [0.234 ,0.828]	0.938 [0.688 ,1.000]	0.783	0.804
	0.703 [0.297 ,0.859]	0.906 [0.688 ,1.000]	0.773	0.792
	0.719 [0.297 ,0.859]	0.906 [0.688 ,1.000]	0.761	0.802
	0.719 [0.359 ,0.875]	0.875 [0.688 ,1.000]	0.729	0.789
	0.734 [0.359 ,0.875]	0.875 [0.656 ,1.000]	0.7	0.799
	0.734 [0.625 ,0.891]	0.844 [0.656 ,1.000]	0.696	0.785
	0.734 [0.641 ,0.906]	0.813 [0.656 ,1.000]	0.694	0.771
	0.750 [0.641 ,0.906]	0.813 [0.625 ,1.000]	0.689	0.78
	0.766 [0.641 ,0.906]	0.813 [0.625 ,0.969]	0.677	0.788
	0.781 [0.641 ,0.906]	0.813 [0.594 ,0.969]	0.662	0.797
	0.781 [0.656 ,0.922]	0.781 [0.594 ,0.969]	0.65	0.781
	0.797 [0.656 ,0.922]	0.781 [0.563 ,0.969]	0.638	0.789
	0.797 [0.672 ,0.938]	0.750 [0.563 ,0.969]	0.627	0.773
	0.813 [0.672 ,0.938]	0.750 [0.500 ,0.969]	0.616	0.78

F1=2*(sensitivities*specificities)/(sensitivities+specificities).
CCDC37.DT-LOCI00505685 with age's cutoff value:0.677; without age's cutoff value:1.29;
LOC440934- LOCI01929088 (XR_001745218.1)with age's cutoff value:0.783;without age's cutoff value:-3.95.

is to minimize regression coefficients by continuously shrinking them. The sum of regression coefficients is reduced by constant shrinking until the coefficients obtained are precisely zero, allowing non-zero variables to remain in the model.

To assess the accuracy of these selected pairs when discerning sperm retrieval, the receiver operating characteristic (ROC) curve analyses (packages: 1.18.0) were performed with the use of the R (version 4.1.1). In the screen set, the area under the ROC curve analysis (AUC value) of each pair of exlncRNAs was calculated.
Validation of exLncRNA pairs (validation set)
In these analyses, the difference by subtracting values (normalized exLncRNA- exLncRNA) was compared with values of sperm (+/-). Based on the AUC of each pair of exlncRNAs, indications were drawn of the discriminatory capacity of the pairs. In this study, they were classified into excellent (AUC more than 0.9), good (AUC more than 0.8), fair (AUC more than 0.7), poor (AUC more than 0.6), and failed (AUC less than 0.60).

The recommended cutoff and sensitivity analysis

We further analyzed the LOC440934- LOCI01929088 (XR_001745218.1) pair and the CCDC37.DT-LOCI00505685 pair by the multiple variables logistic regression analysis. Confusion matrices comprise true positive (TP), true negative (TN), false positive (FP), and false-negative (FN) results. Results are collated to output sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), FP, and false-negative rates (FNR). For performance grading, F1 scores are calculated using F1=2*(sensitivities*specificities)/ (sensitivities+specificities). When the F1 score is equal to the maximum, the cutoff value is most optimal. *P* values of <.05 were considered to be significant.

RESULTS

Clinical and demographic characteristics

Totally 96 NOA patients (n [AZFc microdeletions] = 4; n [Cryptorchidism-associated NOA] = 13; n [Idiopathic NOA] = 13; n [Idiopathic testicular atrophy] = 23; n [Klinefelter's syndrome] = 7; n [Mumps orchitis] = 13; n [Varicocele-associated NOA] = 21; n [Partial AZFb + c microdeletions and Partial AZFb microdeletions] = 2). The flowchart of this study was observed in Figure 1. The clinical and demographic characteristics are summarized and compared between sperm retrieval (+/-) in Table 1. The patient's age and FSH were the most highly related to the discrimination between sperm retrieval (+/-) ($P = 0.003$ and $P = 0.001$, respectively) in **Table 1**. 96 NOA patients' data sets were randomly divided into screen (30 patients) and validation (66 patients) set. But from statistical theory, patient size is insufficient for subgroup-based a on the risk disease of NOA analysis. What is more, there were no significant differences between the two cohorts in **Table 2**. *P* values of <.05 were considered to be significant.

Establishment of pairs based on seminal plasma exlncRNAs

Based on a previous study, the FPKM values by high-throughput sequencing of long RNAs in EVs were transformed by the Z-score. We construct seminal plasma exlncRNA pair's matrix (in supplemental table 1.1 and 1.2).
These findings confirmed that lncRNAs were prevalent in seminal plasmaEVs and that the 120 testis-specific exlncRNA pairs-candidate may be used as biomarkers.
Selection of Potential exlncRNA pairs (screen set)
A co-expression network of 120 exlncRNA pairs was constructed based on the WGCNA algorithm. The grey

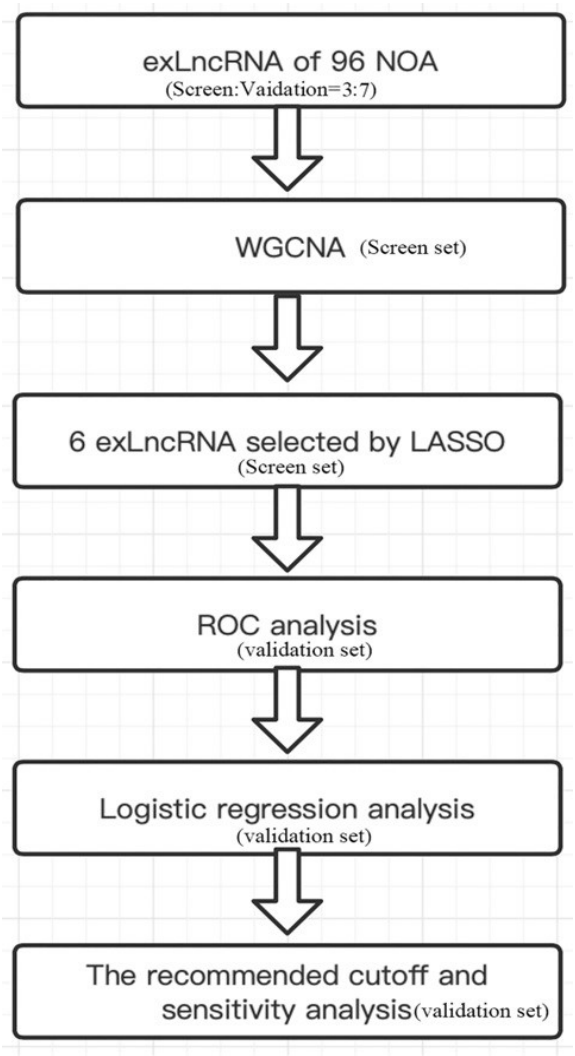

Figure 1. Flowchart of this study.

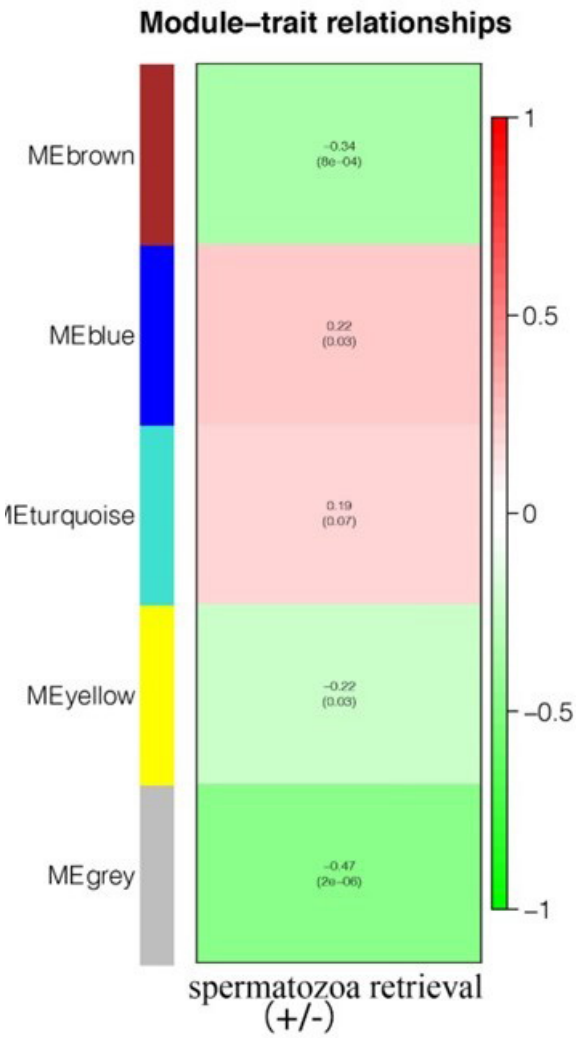

Figure 2. Weighted gene co-expression network analysis (WGCNA) of the exLncRNA. Correlation analysis between modules, and the grey and brown module are most related to sperm retrieval (+/-).

module (10 exlncRNA pairs) is highlighted because it has the most significant correlation (r = -0.47, P < 0.001) with sperm retrieval (+/-). Additionally, sperm retrieval (+/-) was negatively correlated with the brown module (r = -0.34, P < 0.001, 6 exLncRNA pairs). Therefore, the 16 pairs of the two modules were identified for further analysis in **Figure 2**.

Then we applied logistic LASSO regression in Figures 3 and 4, which minimizes multi-collinearity between 16 pairs, to assess the relationship between sperm retrieval (+/-). Using LASSO, we also accounted for well-established sperm retrieval (+/-) predictive factors. We found 6 pairs may be potential for sperm retrieval (+/-).

The six selected exLncRNA Pairs (SPATA42-LOCI00505685, SPATA42-LOCI01929088 (XR_927561.2), CCDC37.DT-LOCI00505685, LOC440934-LOCI00505685, LOC440934-LOCI01929088 (XR_001745218.1), LOCI01929088 (XR_927561.2)-LINC00343) that were obtained from analysis above. In the screen set, the AUC value and AUC-adjusted (age, FSH, total testosterone)) value of each pair of exlncRNAs was calculated and shown in **Table 3**.

Validation of the Selected seminal plasma exlncRNA Pairs (validation set)

Ultimately, six pairs were considered for further analysis based on the pair with the high and stable AUC. The LOC440934- LOCI01929088 (XR_001745218.1) pair AUC and AUC-adjusted (age, FSH, total testosterone) both more than 80%. The CCDC37.DT-LOCI00505685 pair possessed a slight fluctuation and a high AUC value. We found CCDC37.DT-LOCI00505685 pair and LOC440934- LOCI01929088 (XR_001745218.1) pair could become new molecular biomarkers because both were higher and stable AUC values in **Table 3**.

The recommended cutoff and sensitivity analysis

According to the multiple variables logistic regression analysis, we found (age+ CCDC37.DT-LOCI00505685) and (age+ LOC440934- LOCI01929088 (XR_001745218.1)) were independent potential predictors in **Table 4**.

Utilizing cutoff of age and CCDC37.DT-LOCI00505685 > 0.677 (F1=0.78 is maximum) as a predictor of SSR, the cutoff point had a sensitivity of 75%

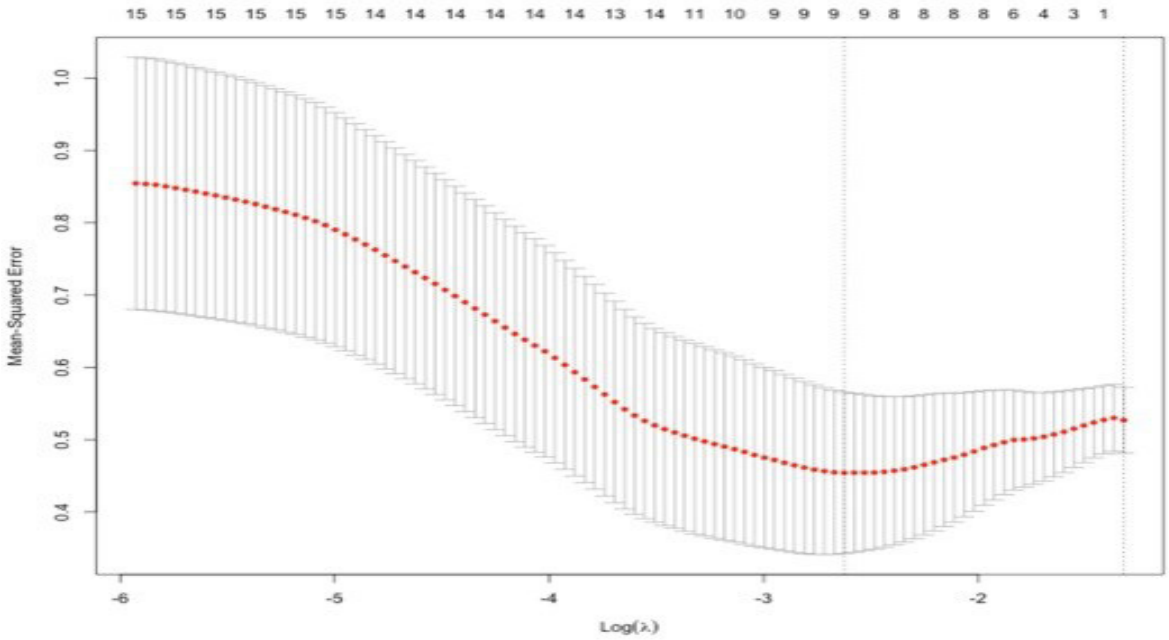

Figure 3. Partial likelihood deviation of LASSO coefficient distribution. The two vertical dashed lines represent lambda. min and lambda.1se, respectively.

and a specificity of 81.3%. Utilizing cutoff of age and LOC440934- LOCI01929088 (XR_001745218.1) > 0.783 (F1 = 0.804 is maximum) as a predictor of SSR, the cutoff point had a sensitivity of 70.3% and a specificity of 93.8% in Table 5 and supplemental **table 2,3**.

DISCUSSION

NOA refers to impaired spermatogenesis in the testes, excluding obstructive factors that lead to azoospermia. Partial NOA patients through hormone therapy may be significantly improved or even successful in infertility. Other patients may need to try surgical sperm retrieval in combination with intracytoplasmic sperm injection (ICSI) to conceive the next generation. One of the most common sperm retrieval procedures for NOA patients is microdissection testicular sperm extraction (MD-TESE). Its advantage is that surgeons can selectively identify the seminiferous tubules that may contain sper-

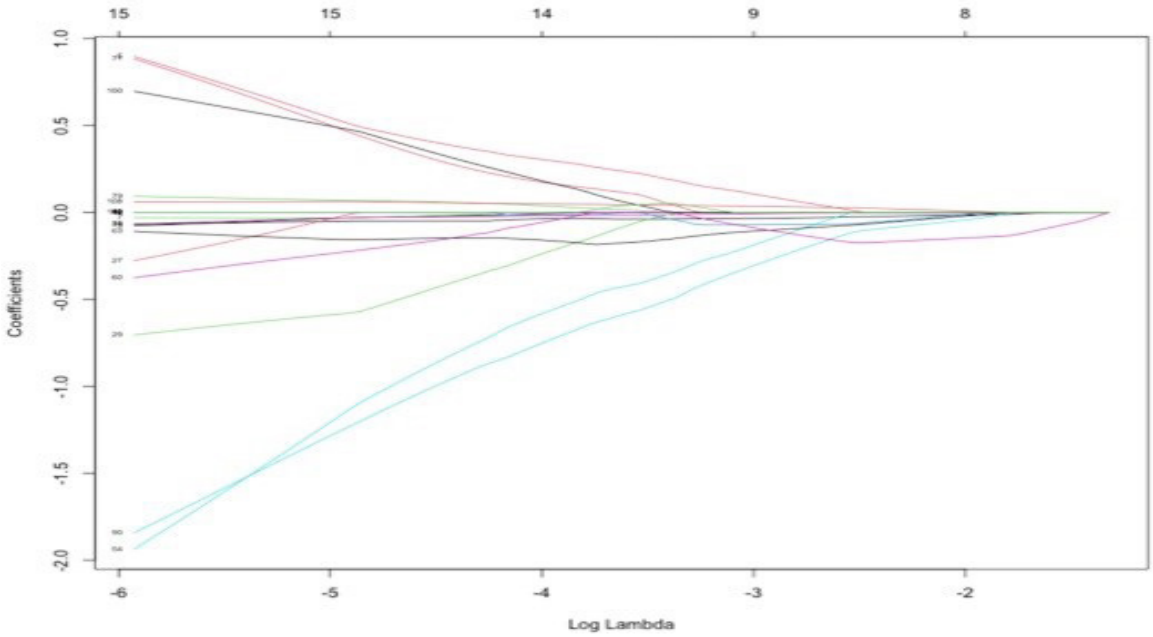

Figure 4. Plots for LASSO regression coefficients over different values of the penalty parameter.

matozoa. The MD-TESE is widely accepted compared to other sperm acquisition techniques. At present, there are many types of research about predicting SRR before the operation, but there is still a lack of consensus predictors.

This study demonstrates that seminal plasma lncRNA pairs may become precise and noninvasive predictors of sperm retrieval in NOA patients. The predictors assist clinicians in choosing a better treatment to accomplish precision medical treatment, reduce harm to patients, and save money for patients. The approach of considering the differential expression of two exlncRNAs as a biomarker value instead of the FPKM level of a single exlncRNA is mainly based on the biological implications of these pairwise fluctuations in NOA patients. The analysis of the association between two exlncRNA (difference of FPKM values) is more robust than the analysis of a single lncRNA because the pair provides more information about spermatogenesis. So, however, the expression level of exLncRNA in pairs, the global difference will reflect a variance that will be found by this pairwise analysis.

AUC values were used to estimate the predictive power of the selected pairings to evaluate predictors. Furthermore, only the couples with high and steady AUC values were considered valid potential biomarkers. The CCDC37.DT-LOCI00505685 pair and LOC440934-LOCI01929088 (XR_001745218.1) pair achieved this goal. Exosomes, multivesicular nanovesicles (50–500 nm) released by the fusion of the plasma membrane with multivesicular bodies, are released into the extracellular space and body fluids[33,34]. Many kinds of cells produce exosomes, which are nanosized membrane vesicles found in abundance in the fluids of living organisms and semen. They are packed with lipids, proteins, microRNAs and mRNAs, and are known to be essential for intracellular communication. Their content is derived from the target cells' endosomal compartments. A wide variety of cells secrete exosomes containing epithelial cells and testis tissue. The exosomes are found in biological fluids such as plasma, saliva, and semen under normal and abnormal conditions[35,36]. Exosomal proteins may indicate a biomarker of male infertility. They may be essential in sperm motility, capacitation, acrosome reaction, and fertilization[37]. With the advance of the global molecules tools and purification of exosomes, it is possible to find many exosomal molecules. Yang et al. showed 1474 exosomal proteins in normozoospermic men. The exosomal proteins primarily relate to protein metabolism, cell growth, and maintenance[38].

In the male reproductive system, altered expression of seminal plasma exosomes containing sncRNA has been reported to participate in molecular events. The miRNAs, such as miR-539-5p, have the potential to predict the presence of spermatozoa in the testes of NOA patients[39]. The miR-539-5p cannot be used to determine the existence of spermatozoa in the testes of NOA patients, normozoospermic, and oligozoospermia. Many long noncoding RNAs play an essential role in spermatogenesis, which gives them an advantage compared with miRNAs for estimating the likelihood of sperm retrieval when an individual has an abnormal sperm cycle[27].

The term LncRNA refers to non-protein-coding RNAs with a length of at least 200 nucleotides[40]. According to Zhao et al., 157 have different expression levels among diverse populations of testicular cells[41]. In a similar study, Jan et al. found a total of 137 lncRNAs were differentially expressed between different testicular cells[42]. Rolland et al. identified 113 known lncRNAs and 20 novel unannotated transcripts important for spermatogenesis[43]. New research avenues can be explored to understand the regulation of spermatogenesis through the signaling involved. Human sperm contains a small amount of lncRNA. Sendler et al. mapped the transcriptome of fertile men's sperm and identified 155 lncRNAs in human sperm[44]. Zhang et al. identified the testis-specific Lnc32058, lnc09522, and Lnc98497 showed high expression between motile and immotile human sperm[45]. Based on a previous study, only 16 testis-specific lncRNA was identified by sequence[29]. A decision-making process for NOA patients has been established based on the scores of the 9-exlncRNA prediction model. Compared to the model, we only need one pair and age to predict, and both pairs were an effective and money-saved method for assessing sperm retrieval in patients with NOA.

Unfortunately, although the tremendous and stable biomarker potential of the CCDC37.DT-LOCI00505685 pair and LOC440934- LOCI01929088 (XR_001745218.1) pair were found in our study, we needed the comprehensive analysis of pairs combined with clinical practice to make clinical decisions for NOA. The NOA sample size and preparation process must be further expanded, simplified and standardized. For example, the seminal plasma exosomal RNA isolation kit is studied to simplify and standardize. We can cooperate with many medical centers to further expand the sample size. This is the study's key limitation since this exLncRNA pair's biomarker suitability must be confirmed before it can be used for clinical diagnosis.

CONCLUSIONS

This study suggested that the LOC440934-LOCI01929088 (XR_001745218.1) pair AUC and AUC-adjusted were more than 80%. The CCDC37.DT-LOCI00505685 pair possessed a slight fluctuation and a high AUC value. Both pairs were an effective and money-saved method for assessing sperm retrieval in patients with NOA. But these findings must be applied to clinical practice by performing validation in more extensive and more diverse NOA patients.

ACKNOWLEDGEMENT

We thank for Prof. Deng et al. group’s contribution.

CONFLICT OF INTEREST

The authors report no conflict of interest.

REFERENCES

1. Cocuzza M, Alvarenga C, Pagani R. The epidemiology and etiology of azoospermia. Clinics (Sao Paulo). 2013;68 Suppl 1:15-26.
2. Kolettis PN. The evaluation and management of the azoospermic patient. J Androl. 2002;23:293-305.
3. Liu X, Chen Y, Tang W, et al. Single-cell transcriptome analysis of the novel coronavirus (SARS-CoV-2) associated gene ACE2 expression in normal and non-

obstructive azoospermia (NOA) human male testes. Sci China Life Sci. 2020;63:1006-15.
4. Esteves SC. Clinical management of infertile men with nonobstructive azoospermia. Asian J Androl. 2015;17:459-70.
5. Ishikawa T, Nose R, Yamaguchi K, Chiba K, Fujisawa M. Learning curves of microdissection testicular sperm extraction for nonobstructive azoospermia. Fertil Steril. 2010;94:1008-11.
6. Okada H, Dobashi M, Yamazaki T, et al. Conventional versus microdissection testicular sperm extraction for nonobstructive azoospermia. J Urol. 2002;168:1063-7.
7. Ramasamy R, Yagan N, Schlegel PN. Structural and functional changes to the testis after conventional versus microdissection testicular sperm extraction. Urology. 2005;65:1190-4.
8. Tsujimura A, Matsumiya K, Miyagawa Y, et al. Prediction of successful outcome of microdissection testicular sperm extraction in men with idiopathic nonobstructive azoospermia. J Urol. 2004;172:1944-7.
9. Klonoff-Cohen HS, Natarajan L. The effect of advancing paternal age on pregnancy and live birth rates in couples undergoing in vitro fertilization or gamete intrafallopian transfer. Am J Obstet Gynecol. 2004;191:507-14.
10. Jarvi K, Lo K, Fischer A, et al. CUA Guideline: The workup of azoospermic males. Can Urol Assoc J. 2010;4:163-7.
11. Aydos K, Unlu C, Demirel LC, Evirgen O, Tolunay O. The effect of pure FSH administration in non-obstructive azoospermic men on testicular sperm retrieval. Eur J Obstet Gynecol Reprod Biol. 2003;108:54-8.
12. Ezeh UI, Moore HD, Cooke ID. A prospective study of multiple needle biopsies versus a single open biopsy for testicular sperm extraction in men with non-obstructive azoospermia. Hum Reprod. 1998;13:3075-80.
13. Gnessi L, Scarselli F, Minasi MG, et al. Testicular histopathology, semen analysis and FSH, predictive value of sperm retrieval: supportive counseling in case of reoperation after testicular sperm extraction (TESE). BMC Urol. 2018;18:63.
14. Bohring C, Schroeder-Printzen I, Weidner W, Krause W. Serum levels of inhibin B and follicle-stimulating hormone may predict successful sperm retrieval in men with azoospermia who are undergoing testicular sperm extraction. Fertil Steril. 2002;78:1195-8.
15. Ballesca JL, Balasch J, Calafell JM, et al. Serum inhibin B determination is predictive of successful testicular sperm extraction in men with non-obstructive azoospermia. Hum Reprod. 2000;15:1734-8.
16. Cissen M, Meijerink AM, D'Hauwers KW, et al. Prediction model for obtaining spermatozoa with testicular sperm extraction in men with non-obstructive azoospermia. Hum Reprod. 2016;31:1934-41.
17. Yang Q, Huang YP, Wang HX, et al. Follicle-stimulating hormone as a predictor for sperm retrieval rate in patients with nonobstructive azoospermia: a systematic review and meta-analysis. Asian J Androl. 2015;17:281-4.
18. Li H, Chen LP, Yang J, et al. Predictive value of FSH, testicular volume, and histopathological findings for the sperm retrieval rate of microdissection TESE in nonobstructive azoospermia: a meta-analysis. Asian J Androl. 2018;20:30-6.
19. Hashemi MS, Mozdarani H, Ghaedi K, Nasr-Esfahani MH. Expression of ZMYND15 in Testes of Azoospermic Men and Association With Sperm Retrieval. Urology. 2018;114:99-104.
20. Hashemi MS, Mozdarani H, Ghaedi K, Nasr-Esfahani MH. Among seven testis-specific molecular markers, SPEM1 appears to have a significant clinical value for prediction of sperm retrieval in azoospermic men. Andrology. 2018;6:890-5.
21. Hashemi MS, Mozdarani H, Ghaedi K, Nasr-Esfahani MH. Could analysis of testis-specific genes, as biomarkers in seminal plasma, predict presence of focal spermatogenesis in non-obstructive azoospermia? Andrologia. 2020;52:e13483.
22. Vojtech L, Woo S, Hughes S, et al. Exosomes in human semen carry a distinctive repertoire of small non-coding RNAs with potential regulatory functions. Nucleic Acids Res. 2014;42:7290-304.
23. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet. 2009;10:155-9.
24. Zhang C, Gao L, Xu EY. LncRNA, a new component of expanding RNA-protein regulatory network important for animal sperm development. Semin Cell Dev Biol. 2016;59:110-7.
25. Wen K, Yang L, Xiong T, et al. Critical roles of long noncoding RNAs in Drosophila spermatogenesis. Genome Res. 2016;26:1233-44.
26. Li L, Wang M, Wang M, et al. A long non-coding RNA interacts with Gfra1 and maintains survival of mouse spermatogonial stem cells. Cell Death Dis. 2016;7:e2140.
27. Hong SH, Kwon JT, Kim J, et al. Profiling of testis-specific long noncoding RNAs in mice. BMC Genomics. 2018;19:539.
28. Satoh Y, Takei N, Kawamura S, Takahashi N, Kotani T, Kimura AP. A novel testis-specific long noncoding RNA, Tesra, activates the Prss42/Tessp-2 gene during mouse spermatogenesisdagger. Biol Reprod. 2019;100:833-48.
29. Xie Y, Yao J, Zhang X, et al. A panel of extracellular vesicle long noncoding RNAs in seminal plasma for predicting testicular spermatozoa in nonobstructive azoospermia patients. Hum Reprod. 2020;35:2413-27.
30. Vogt PH. AZF deletions and Y chromosomal haplogroups: history and update based on sequence. Hum Reprod Update. 2005;11:319-36.

31. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.
32. Liu R, Yuan M, Xu H, Chen P, Xu XS, Yang Y. Adaptive weighted sum tests via LASSO method in multi-locus family-based association analysis. Comput Biol Chem. 2020;88:107320.
33. Valadi H, Ekstrom K, Bossios A, Sjostrand M, Lee JJ, Lotvall JO. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nat Cell Biol. 2007;9:654-9.
34. Zhang Y, Liu Y, Liu H, Tang WH. Exosomes: biogenesis, biologic function and clinical potential. Cell Biosci. 2019;9:19.
35. Poliakov A, Spilman M, Dokland T, Amling CL, Mobley JA. Structural heterogeneity and protein composition of exosome-like vesicles (prostasomes) in human semen. Prostate. 2009;69:159-67.
36. Nilsson J, Skog J, Nordstrand A, et al. Prostate cancer-derived urine exosomes: a novel approach to biomarkers for prostate cancer. Br J Cancer. 2009;100:1603-7.
37. Burden HP, Holmes CH, Persad R, Whittington K. Prostasomes--their effects on human male reproduction and fertility. Hum Reprod Update. 2006;12:283-92.
38. Yang C, Guo WB, Zhang WS, et al. Comprehensive proteomics analysis of exosomes derived from human seminal plasma. Andrology. 2017;5:1007-15.
39. Barcelo M, Mata A, Bassas L, Larriba S. Exosomal microRNAs in seminal plasma are markers of the origin of azoospermia and can predict the presence of sperm in testicular tissue. Hum Reprod. 2018;33:1087-98.
40. Batista PJ, Chang HY. Long noncoding RNAs: cellular address codes in development and disease. Cell. 2013;152:1298-307.
41. Zhu Z, Li C, Yang S, et al. Dynamics of the Transcriptome during Human Spermatogenesis: Predicting the Potential Key Genes Regulating Male Gametes Generation. Sci Rep. 2016;6:19069.
42. Jan SZ, Vormer TL, Jongejan A, et al. Unraveling transcriptome dynamics in human spermatogenesis. Development. 2017;144:3659-73.
43. Rolland AD, Evrard B, Darde TA, et al. RNA profiling of human testicular cells identifies syntenic lncRNAs associated with spermatogenesis. Hum Reprod. 2019;34:1278-90.
44. Sendler E, Johnson GD, Mao S, et al. Stability, delivery and functions of human sperm RNAs at fertilization. Nucleic Acids Res. 2013;41:4104-17.
45. Zhang X, Zhang P, Song D, et al. Expression profiles and characteristics of human lncRNA in normal and asthenozoospermia spermdagger. Biol Reprod. 2019;100:982-93.