Published in final edited form as:

Breast Cancer Res Treat. 2007 July ; 104(1): 57–66. doi:10.1007/s10549-006-9388-4.

A strong association between body fat mass and protein profiles in nipple aspirate fluid of healthy premenopausal non-lactating women

Yafei Huang¹, Manubai Nagamani², Karl E. Anderson¹, Alexander Kurosky³, Anthony M. Haag³, James J. Grady¹, and Lee-Jane W. Lu¹

¹Department of Preventive Medicine and Community Health, The University of Texas Medical Branch, Galveston, TX 77555-1109

²Department of Obstetrics and Gynecology, The University of Texas Medical Branch, Galveston, TX 77555-1109

³Department of Biochemistry & Molecular Biology, The University of Texas Medical Branch, Galveston, TX 77555-1109

Abstract

Fluid can be aspirated from the nipples of most non-lactating women. This nipple aspirate fluid (NAF) is a potential source for the discovery of new breast cancer biomarkers. NAF has two distinct protein profiles. Type I NAF is similar to the fluid associated with cystic disease of the breast, whereas type II NAF is enriched in milk-associated proteins. The prevalence of these two profiles differs in healthy women and in breast cancer patients. This study investigated the relationship of these two NAF profiles to reproductive history, body composition, diet, and levels of lipids, steroids and thyroid hormones in healthy premenopausal women (age 30–40 years) who had regular menstrual cycles and normal mammograms and were not taking contraceptive medications. On average, women with the type I NAF profile were older, had more years since last childbirth, were less likely to have breastfed their babies and had higher dietary saturated fat intake, body mass index, body fat mass, and levels of plasma low density lipoproteins than women with the type II profile ($P < 0.05$). Using multiple logistic regression, type I NAF was predicted independently ($P < 0.05$) by higher body fat mass [Odds Ratio=3.0; 95% Confidence Interval (CI): 1.5–6.1], more years since last childbirth (OR=2.6; 95% CI: 1.3–5.2) and a higher percentage of calories from saturated fat (OR=4.1; 95% CI: 1.1–14.6). These results suggest that protein profiles of NAF might be influenced by amounts or types of dietary and body fat, but further study of the relationship of the two profiles to breast cancer risk is needed.

Keywords

breast cancer biomarker; proteome; obesity; milk proteins; gross cystic disease fluid proteins

Introduction

Nipple aspirate fluid (NAF) represents the secretion from the breast ducts and lobules, from which breast cancer arises. NAF can be collected non-invasively and repeatedly from women

Address correspondence to Lee-Jane W. Lu, Ph.D. at the Department of Preventive Medicine and Community Health, The University of Texas Medical Branch, 700 Harborside Drive, Galveston, TX 77555-1109; Tel: (409) 772 1730; Fax: (409) 772 6287; Email: llj@utmb.edu.

who are not lactating, and may reflect physiological and pathological changes in the breast. Therefore, NAF is considered an ideal source for discovering new biomarkers of breast cancer.

NAF is known to contain proteins, lipids, steroid hormones, growth factors, and carbohydrates [1], at concentrations much higher than in plasma. For example, the average level of proteins in NAF is 71–170 mg/ml, which is much higher than the 6–8 mg/ml reported for plasma [2, 3]. Several NAF proteins, including hormone-regulated proteins and cancer markers such as prostate-specific antigen and carcinoembryonic antigen, have been identified using antibodies [4,5]. Sanchez *et al.* [6] described two distinct types of NAF proteomes based on the profiles of highly abundant proteins. Type I NAF is enriched in the same proteins found in gross cystic disease fluid, including zinc α 2-glycoprotein (ZAG), apolipoprotein-D (apoD), and gross cystic disease fluid protein-15 (GCDFP-15), which is also known as prolactin-inducible protein. Type II NAF is characterized by the great abundance of milk-associated proteins, including lactoferrin, α -lactalbumin, and lysozyme C.

Type I NAF proteins have also been identified in other human body fluids, such as saliva, sweat, and nasal lavage fluids [7–9]. Higher levels of ZAG, apoD, and GCDFP-15 in serum and other body fluids have been associated with apocrine activity [10]. The expression of type I NAF proteins has been studied in cultures of malignant mammary epithelial cells. In these cultures, androgens (dihydrotestosterone or fluoxymesterone) stimulated the secretion of ZAG, apoD and GCDFP-15 into the culture media [11–14], progesterone and prolactin up-regulated GCDFP-15 expression [15,16], and estradiol decreased the synthesis of GCDFP-15 and apoD [12–14]. In contrast, Zhou *et al.* showed that testosterone, progesterone, and tamoxifen decreased apoD mRNA expression in the mammary glands of ovariectomized rhesus monkeys [17].

Regulation of the synthesis of milk proteins by hormones during pregnancy and lactation is complex. Prolactin is believed to play a primary role in stimulating the expression of milk proteins [18]. Several type II NAF proteins, such as lactoferrin, lysozyme and lactalbumin, are inducible by prolactin, and also by insulin and corticosterone, in cultures of breast cells and in the mammary glands of laboratory animals [19]. In mice, estrogen up-regulated the expression of lactoferrin in the uterus, but not in the mammary glands [20].

Limited studies showed that both types of NAF are found in healthy women and in breast cancer patients, but with different prevalence. Sanchez *et al.* found type II NAF in only 10% of women without breast pathology or with benign breast disease, but in slightly more than half of women with breast cancer [6]. More recent studies from the same group showed that cancer patients with type II NAF generally had a better prognosis compared with those with type I [21]. Therefore, these protein profiles might be predictors for both breast cancer risk and prognosis. Little is known, however, about the mechanisms and factors affecting the secretion of these proteins in women. This study investigated the demographic, reproductive, anthropometric, hormonal, and nutritional factors that are associated with the secretion of these two types of NAF.

Experimental Subjects

This was a cross-sectional study of healthy, premenopausal, non-pregnant, non-lactating, 30–40 year old women who had volunteered for a dietary intervention study at the University of Texas Medical Branch (UTMB). Subjects had been recruited from communities within a 50 mile radius of Galveston, Texas, by posted advertisements and postal mailings. The baseline data obtained prior to the beginning of the dietary intervention were used for this study. The subjects had never used contraceptive medications (pills, injections, or patches), or had stopped taking them at least 6 months prior to enrollment. Women who were vegetarians, had irregular

menstrual cycles or had first degree relatives with breast cancer were excluded from the study. All women had normal mammograms at entry into the study. Written informed consent was obtained from all subjects. A total of six study visits were scheduled during the luteal phases of two separate menstrual cycles, usually on the days between cycle day 20 and 24, with three visits during each menstrual cycle. The study protocol was approved by the Institutional Review Board of UTMB and the Human Subject Research Review Board of the US Army Medical Research and Materiel Command.

Materials and Methods

Body weight and height were measured at each study visit. At one study visit, total body mass, body lean mass, body fat mass, and %-body fat were measured in duplicate, with the subject in a supine position, by dual energy X-ray absorptiometry (DEXA) (Model QDR4500A, Hologic, Waltham, MA). Average values for the two measurements were used for statistical analyses. Reproductive history was obtained using a self-administered standard gynecologic clinic questionnaire. Nutrient intakes, based on three 24-hour food records, were analyzed using the Nutrition Data System for Research software (the Nutrition Coordinating Center, University of Minnesota, MN), and averaged for statistical analyses. NAF was obtained by gentle suction using a modified breast pump made in our laboratory. Two or three aspirations from each breast were attempted during each of three separate study visits. Fasting blood samples were obtained at every study visit, and the first three samples, all from the same luteal phase, were used for measuring steroid hormones. All samples collected were stored immediately at -80°C until analysis.

Plasma samples were analyzed for levels of progesterone, testosterone, and 17β -estradiol using commercial immunoassay kits (Diagnostics Labs, Webster, TX), according to the manufacturers' instructions. A radioactive immunoassay kit was used to examine progesterone levels (sensitivity 0.1 ng/ml), and enzymatic immunoassay kits were used for testosterone levels (0.04 ng/ml) and 17β -estradiol levels (7 pg/ml). For all the assays, the average intra- and inter-assay coefficients of variation (CV) were less than 15%. Each sample was assayed at least twice (and repeated if the intra-assay CV was $>15\%$), and the results were averaged. Hormone levels from three different cycle days of the same luteal phase were averaged for each subject for statistical analyses. Fasting blood samples from two different study visits were also measured for triiodothyronine (T3), tetraiodothyronine (T4), thyroid stimulating hormone (TSH), high density lipoprotein (HDL), low density lipoprotein (LDL), very low density lipoprotein (VLDL), total cholesterol, and triglycerides by the certified UTMB hospital clinical laboratory, and the average of the two measurements were used for statistical analyses.

Protein profiling of NAF

One dimensional (1D) SDS-PAGE (12.5 %, 1.0 mm, 200 Volts, 1 hr) was used to separate abundant NAF proteins (30 μg) for the classification of NAF types. After staining with Coomassie blue (Biosafe Coomassie, Biorad, Hercules, CA), visible bands of the gels were excised and trypsin-digested (15 $\mu\text{g}/\text{ml}$ of trypsin) at 37°C overnight. Mass spectra of peptide digests were obtained using a matrix-assisted laser desorption/ionization-time of flight mass spectrometer (MALDI-TOF MS) (Applied Biosystems, Foster City, CA). Proteins were identified by peptide mass fingerprinting with a search against the National Center for Biology Information (NCBI) protein database using the Profound algorithm. Positive protein identifications were accepted for those with scores above the 95% confidence level.

Four out of sixteen breast milk samples and four out of forty-two type II NAF samples from age-matched women were randomly chosen for comparison of protein compositions. Human milk samples (N=16) were obtained and donated by the Mother's Milk Bank at Austin (Austin, TX) and separated on 1D gels. After staining with Coomassie blue, the gel (randomly chosen

from breast milk and type II NAF) was scanned using the AlphaImager™ 2200 (Alpha Innotech Co., San Leandro, CA), and the density and percentage of the total density for each individual band were measured using the Alpha Ease FC software. The major bands were excised, digested, and subjected to protein identification by MALDI-TOF MS. The mean percentage staining density for the major proteins, including lactoferrin, albumin, β -casein, and lactalbumin, were compared between type II NAF and human breast milk samples.

Statistical analysis

Means and standard deviations were computed for continuous data, and percentages for categorical data. For univariate analyses, the χ^2 test was performed for categorical variables and the Student's *t*-test for continuous variables for comparisons between groups with type I and type II NAF. Variables associated with NAF protein profiles by univariate analysis (using $P < 0.10$ as a cut-off) were further entered into age-adjusted multi-level logistic regression models, allowing the identification of independent predictors. For the multivariate logistic regression models, a new ordinal variable was created to code "years since last childbirth". Specifically, the first quartile (0 to 3.3 yrs) was coded as "0", second quartile (3.3 to 7.2 yrs) as "1", third quartile (7.2 to 10.5 yrs) as "2", the 4th quartile (over 10.5 yrs) as "3", and nulliparous women as "4". The linearity of the effect of the new variable was assessed and confirmed. Stepwise selection was used to determine predictors associated with secretion types of NAF. Variables with a two-tailed *P* value of < 0.05 were kept in the final models. Correlations among parameters of anthropometrics and among demographic and reproductive variables were tested using Pearson correlation analysis. All statistical analyses were performed using SAS® (Version 9.1, SAS Institute Inc., Cary, NC).

Results

Identification of NAF proteins and classification of NAF types

The abundant proteins in NAF that were visible in a 1D gel after Commassie blue staining (Figure 1) were subjected to identification by peptide mass fingerprinting [22]. Two protein bands, albumin (abbreviated as Alb in Figure 1, 67 kDa) and Ig heavy chain (Ig-H, 59 kDa), were found in all NAF samples. ZAG (41 kDa), apoD (24 kDa), and GCDFP-15 (17 kDa) were typically found and were more abundant in type I NAF, while lactoferrin (Lf, 78 kDa), lysozyme C (Lyz, 15 kDa), and α -lactalbumin (Lalb, 14 kDa) were typically found and were more enriched in type II NAF, as shown in Figure 1. β -Casein (25 kDa, multiple bands in Figure 1) was also present but with varying abundance in type II NAFs. Based on the gel mobility, the molecular weights (shown in parentheses) of these NAF proteins were consistent with published data. Type I NAFs were found in 58%, and type II in 33% of the women in our study (see Table 1 below). Some women (9%) had a protein profile that was a mixture of types I and II, showing the presence of lactoferrin, ZAG, apoD, GCDFP-15 and lactalbumin (mixed type in Figure 1).

Figure 1 also shows the 1D-gel profile of breast milk samples for comparison with type II NAF. Breast milk samples contained lactoferrin, albumin, casein, lysozyme and lactalbumin. Although type II NAF contains mostly milk-associated proteins, their relative proportions in breast milk and type II NAF differed. As shown in Figure 1, breast milk generally had a higher content of casein proteins ($29\% \pm 7.0\%$ of the total protein staining-density for milk vs. $2.3\% \pm 4.6\%$ for type II NAF, $P < 0.05$, $N = 4$ per group), whereas type II NAF had more albumin ($9.3\% \pm 1.6\%$ of the total density for milk vs. $24.7\% \pm 10.0\%$ for type II NAF, $P < 0.05$, $N = 4$ per group). There were no differences in the mean percentage densities of lactoferrin and lactalbumin.

Prevalence of the protein profiles in the study population

Nipple aspiration was attempted on three separate study visits from a total of 219 subjects. Women were classified as secretors if fluid was obtained at least once after three attempts. Of the 219 subjects, 148 (68%) were secretors, and 128 of these samples were large enough for protein profiling. The success rate for obtaining NAF in our study is consistent with previous reports [2,23].

Table 1 shows the distribution of NAF protein profiles in our subjects (N=128). Type I NAF was more common than type II (ratio 1.8 to 1). Mixed type NAF (containing both type I and II signature proteins) was found in less than 10% of the subjects, and these were not included in *t*-test and multivariate analyses due to the small sample size. Among women who had given birth within the last four years, 63% showed type II NAF (Table 1).

Characteristics of the study population by NAF secretion types

Table 2 describes demographic, reproductive, nutritional, and hormonal characteristics of our study subjects, related to secretion types. Ethnicity did not differ significantly between the type I and type II groups. There were no differences in the distribution of secretion types between past users and those who have never used contraceptive medications. Women with type II NAF had a significantly higher rate of breastfeeding compared to those with type I NAF (by 29.8%, $P=0.002$).

Women with type I NAF were older in age compared to women with type II NAF (by 1.2 years, $P=0.01$), slightly younger at first childbirth (by 1.8 years, $P=0.08$), and had a lower number of complete pregnancies (2.1 compared to 2.6, $P=0.04$), and a longer time since last childbirth (by 3.9 years, $P<0.0001$). Women with type I NAF were also heavier (by 5.8 kg, $P=0.03$), had higher BMIs (by 2.1 kg/m², $P=0.02$), more %-body fat (by 3.1%, $P=0.02$), more total body fat mass (by 4.9 kg, $P=0.008$), more total body lean mass (by 2.1 kg, $P=0.09$), higher blood LDL levels (by 11.4 mg/dl, $P=0.05$), and higher plasma 17 β -estradiol levels (by 10.3 pg/ml, $P=0.08$), and they consumed more calories from saturated fat (by 0.9%, $P=0.09$). Age and years since last childbirth for women with a mixed NAF type were values that fell between those of women with type I and type II NAF (Table 2).

There were no significant differences between the groups in age of menarche, height, plasma levels of testosterone and progesterone, serum levels of T3, T4, and TSH, total caloric intake, %-caloric intake from total fat, monounsaturated fat and polyunsaturated fat, and blood levels of lipids other than LDL. In addition, there were 20 secretors who did not have samples large enough for protein profiling. The characteristics shown in Table 2 for subjects with adequate samples were not significantly different from those with inadequate samples ($P>0.05$, *t*-test, results not shown in Table 2).

Logistic regression model to predict secretion types

Multi-level age-adjusted logistic regression models were constructed to predict NAF secretion types (I or II) as shown in Table 3. Model 1 included reproductive variables (i.e. age, age at first childbirth, history of breastfeeding, parity and years since last childbirth). Years since last childbirth was the only significant predictor for secretion types ($P=0.0008$). Age at first childbirth, history of breastfeeding, and parity were not predictors. Model 2 included variables in model 1 plus anthropometric parameters (i.e. weight, BMI, %-body fat, total body fat mass, and total body lean mass). Both years since last childbirth ($P=0.003$) and total body fat mass ($P=0.0007$) were found to be significant predictors for secretion types. BMI, body lean mass, %-body fat, and weight were not predictors. Model 3 included variables in model 2 plus dietary fat intake. Percentage caloric intake from saturated fat ($P=0.03$) was another significant predictor for secretion types, in addition to years since last childbirth ($P=0.008$) and total body

fat mass ($P=0.002$). Model 4 included the variables in model 3 plus estradiol and blood LDL levels, neither of which was associated with secretion types. Estradiol level as a continuous variable was not a significant predictor of secretion types in model 4 ($P=0.11$). However, when the estradiol level was divided into quartiles and coded as a categorical variable, a significantly higher percentage of type I NAF was found among women in the highest quartile (estradiol level >97.5 pg/ml) compared to those in the lowest quartile (<55.7 pg/ml), after adjustment for body fat mass, years since last complete pregnancy, and %-caloric intake from saturated fat ($OR=5.9$, 95% CI from 1.3 to 27.7, $P<0.01$). In summary, results shown in Table 3 indicate that incremental changes of about 3 years since last childbirth ($OR=2.6$), every 10 kg in total body fat mass ($OR=3.0$), and every 5% in caloric intake from saturated fat ($OR=4.1$) increased the odds of being type I secretors.

It is noteworthy that there were strong correlations among some variables entered into the logistic regression models. Anthropometric variables were positively correlated with one another, with r values ranging from 0.22 to 0.93 (Pearson correlation, $P<0.05$). There were also strong correlations between parity and age at first childbirth ($r=-0.47$, $P<0.001$), and between age at first childbirth and years since last childbirth ($r=-0.52$, $P<0.001$). Age was also positively correlated with the variable “years since last childbirth” in the age range of our study subjects ($r=0.47$, $P<0.0001$), which was significant in each of the regression models. This may be the reason why age was not an independent predictor in the multiple logistic regression models. In addition, history of breastfeeding was significantly associated with secretion type in univariate analysis, but was not a predictor that was independent of years since last childbirth in model 1. However, if the variable “years since last childbirth” was not included in the models, breastfeeding ($OR=16.0$; CI: 3.0–84.1) was an independent predictor, while the influence of total body fat mass ($OR=2.2$; CI: 1.1–4.4) and %-calorie from saturated fat ($OR=5.6$; 95% CI: 1.7–18.5) on secretion types remained the same as when years since last childbirth was included. This may be because breastfeeding had a strong association with the ordinal variable “years since last childbirth” in our study. Women with fewer years since last childbirth were more likely to have a past history of breastfeeding (P trend <0.0001).

Discussion

Protein profiles of NAF are a potential source for the discovery of new breast cancer biomarkers. There is some limited evidence from the study of Vizoso *et al.* [6,21] that both the prevalence and the prognosis of breast cancer may be related to a woman’s NAF protein profile. Vizoso *et al.* found that recent childbirth was a strong predictor of NAF protein composition [24]. Other possible determinants for protein patterns in NAF have not, to our knowledge, been examined. We studied a variety of potential influences, such as reproductive history, hormones, nutrition, and body composition. We found that body fat mass (not body lean mass) and %-caloric intake from saturated fat are strong predictors of type I NAF and that their effects are independent of a recent history of childbirth.

Interestingly, other potential factors such as BMI and weight, that included body lean mass and differed in women with type I and II NAF in univariate analysis, were not independent predictors in multivariate analyses when body fat mass was a variable in the model. Of the various types of fat, only saturated fat had an influence on NAF protein profiles. Because body fat mass and fat intake had strong influences on NAF proteomes, the influence of circulating lipids, which can partly reflect fat intake, were also investigated. LDL was a weak predictor for secretion types in univariate analysis, but was not an independent predictor in multivariate analyses. Indices of thyroid function, e.g. T3, T4 and TSH, were not associated with NAF types, consistent with results previously reported by Vizoso *et al.* [25]. It was the total body fat mass, not body fat as a percentage of body weight, that independently predicted secretion types in multivariate analyses. Total body lean mass is also not a predictor of secretion type.

These results suggest that dietary fat or fat stores have an influence on NAF protein profiles. Adipose tissue has recently been viewed as an endocrine organ, capable of synthesizing and secreting hormones (such as androgens, estrogens, and leptin), and cytokines (such as interleukin-6 and tumor necrosis factor- α) [26–29]. These substances may act in a paracrine/autocrine manner to regulate breast fluid protein.

Because testosterone and progesterone were shown to up-regulate and estradiol to down-regulate ZAG, apoD and GCDFP-15 in breast cancer cell cultures [11–16], we hypothesized that levels of these hormones might be significant predictors for NAF protein profiles. However, no differences in levels of testosterone and progesterone were detected, and only estradiol was marginally higher in type I compared to type II secretors ($P=0.08$) in univariate analyses. Nevertheless, in an age-adjusted logistic regression model, a significantly higher prevalence of type I NAF was found in women within the highest quartile of estradiol levels compared to those within the lowest quartile ($P<0.01$). These data suggest a threshold effect of circulating estradiol levels on NAF types. A more detailed dose-response study is needed to establish the relationship between estradiol and type I proteins. The weak stimulatory effect of estradiol for type I proteins observed in our human study, while not consistent with *in vitro* results from breast cancer cell culture studies, is in line with *in vivo* effects in rhesus monkeys where tamoxifen, an anti-estrogen, decreased the mRNA expression of apoD (a type I NAF protein) [17]. Considering the potential autocrine/paracrine effects, it is also possible that steroid levels in NAF, which were not measured in this study due to insufficient sample quantity, may be a better predictor of NAF profiles than hormone levels in serum.

ApoD and ZAG are involved in the metabolism and disposition of fat. ApoD is a component of the HDL complex and serves as a carrier protein for transporting cholesterol and other lipids from tissues via the blood stream to the liver for further metabolism and disposition. Liu *et al.* found an interaction of apoD with leptin receptor B in hypothalamic neurons, and an association of hypothalamic apoD expression with body fat and circulating levels of leptin [30]. ZAG is also a carrier protein for various hydrophobic molecules, including lipids, and is well recognized as a lipid mobilizing factor [31]. ApoD and ZAG may also regulate body fat accrual [30], and their secretion into the circulation may be part of an effort to remove and metabolize excess fat [32]. These data are consistent with our finding that higher body fat mass and dietary saturated fat intake were associated with elevated levels of the two type I proteins, apoD and ZAG, in NAF.

The finding that recent childbirth was strongly associated with type II NAF ($P<0.0001$) is not surprising, since type II proteins are mostly milk-associated proteins. Our results are consistent with those of Vizoso *et al.*, who reported that type II NAF was found mostly in women who had given birth within the last four years [24]. The prevalence of type II secretors, however, was higher in our study population (33%, Table 1), even after excluding women who had given birth within the last four years (20%, Table 1), when compared to the 9% reported by Vizoso *et al.*... This might be due to differences in the selection of study subjects. Our subjects were healthy women between 30 and 40 years of age, while Vizoso and coworkers selected subjects, from a hospital-based population, that were between 20 and 50 years of age. Current usage of oral contraceptives was also a predictor for type II secretion in the study of Vizoso *et al.*, whereas we excluded current users of all contraceptive medications. In addition, our study showed that past usage of contraceptive medications did not affect the NAF protein profile.

The proteomes of type II NAF and breast milk are qualitatively similar, though there are quantitative differences (Figure 1 and Results). We found that caseins are more abundant than albumin in breast milk, whereas albumin was more abundant than caseins in type II NAF. Moreover, caseins were not always visible in type II NAF, suggesting a gradual shift from milk fluid to type II NAF after childbirth. This, coupled with the presence of a mixed type NAF,

leads us to speculate that type I and type II NAF may not be a constant physiological state in women, and that longitudinal studies are needed to confirm this possibility. Our observation that breastfeeding *per se* is associated with differences in NAF protein profiles, but is not independent from years since last child birth is consistent with this potential time-dependent change from type II to type I NAF.

The few available studies examining proteins in NAF as potential breast cancer markers have found mixed and sometimes contradictory results. For example, GCDFP-15 has been proposed as a plasma marker of proliferative breast disease [33], and was also found to be elevated in the serum of metastatic breast cancer patients [34], but, in another study, it was reported to be decreased in the NAF of breast cancer patients [35]. This inconsistency may be simply attributable to a difference in the prevalence of type I NAF (which is positive for GCFDP-15) and type II NAF in the two study populations. The prevalence of the two NAF types is modifiable by pregnancy history, body fat mass, and dietary fat intake, as shown in our study, and by current usage of birth control medications and age, as was shown in other studies [24, 36]. This underscores the importance of a mechanistic understanding of the presence and relevance of type I and II NAF in pathological and physiological states of the breast before breast cancer biomarkers can be effectively identified.

The major strength of our study is a well-defined study population with a variety of variables available for analysis. However, because of the narrow age range of the study subjects who were all premenopausal, our findings may have limited applicability to postmenopausal women.

In conclusion, the secretion of NAF proteins was influenced strongly by body fat mass, reproductive history, dietary fat intake, and possibly by estradiol levels. Obesity is an important public health issue, and also plays a critical role in breast cancer risk. Studies with larger samples and longitudinal observations may help determine the relevance of body fat mass and NAF proteomes in breast cancer prevention and prognosis.

Acknowledgements

We thank the staff of the GCRC at UTMB for nursing and dietary research assistance. Special thanks to study volunteers, Dr. Astrid Inness for the analysis of food records, Dr. Marinel Ammenheuser for critical review of the manuscript, and the Mother's Milk Bank at Austin for providing human breast milk samples.

Grant support: U.S. Army MRMCM under DAMD17-01-1-0417, NIH NCRR GCRC M01 RR00073, NIH R01 CA95545, U.S. Army MRMCM under W81XWH-04-1-0345, NIH 2 P30 ES06676, 1 R24 CA88317, AICR grant 01B110, and USPHS CA65628.

Abbreviations used

NAF, nipple aspirate fluid
 ZAG, zinc α 2-glycoprotein
 apoD, apolipoprotein D
 GCDFP-15, gross cystic disease fluid protein 15 (prolactin-inducible protein)
 BMI, body mass index
 DEXA, dual energy X-ray absorptiometry
 T3, triiodothyronine
 T4, tetraiodothyronine
 TSH, thyroid stimulating hormone
 HDL, high density lipoprotein
 LDL, low density lipoprotein
 VLDL, very low density lipoprotein
 MALDI-TOF MS, matrix-assisted laser desorption/ionization-time of flight mass spectrometry

Reference List

1. Petrakis NL. Physiologic, biochemical, and cytologic aspects of nipple aspirate fluid. *Breast Cancer Res Treat* 1986;1:7–19. [PubMed: 3539239]
2. Klein P, Glaser E, Grogan L, Keane M, Lipkowitz S, Soballe P, Brooks L, Jenkins J, Steinberg SM, DeMarini DM, Kirsch I. Biomarker assays in nipple aspirate fluid. *Breast J* 2001;6:378–387. [PubMed: 11843848]
3. Djuric Z, Visscher DW, Heilbrun LK, Chen G, Atkins M, Covington CY. Influence of lactation history on breast nipple aspirate fluid yields and fluid composition. *Breast J* 2005;2:92–99. [PubMed: 15730453]
4. Lee MM, Petrakis NL, Wensch MR, King EB, Miike R, Sickles E. Association of abnormal nipple aspirate cytology and mammographic pattern and density. *Cancer Epidemiol Biomarkers Prev* 1994;1:33–36. [PubMed: 8118383]
5. Sauter ER, Chervoneva I, Diamandis A, Khosravi JM, Litwin S, Diamandis EP. Prostate-specific antigen and insulin-like growth factor binding protein-3 in nipple aspirate fluid are associated with breast cancer. *Cancer Detect Prev* 2002;2:149–157. [PubMed: 12102150]
6. Sanchez LM, Vizoso F, Diez-Itza I, Lopez-Otin C. Identification of the major protein components in breast secretions from women with benign and malignant breast diseases. *Cancer Res* 1992;1:95–100. [PubMed: 1727390]
7. Hu S, Xie Y, Ramachandran P, Ogorzalek Loo RR, Li Y, Loo JA, Wong DT. Large-scale identification of proteins in human salivary proteome by liquid chromatography/mass spectrometry and two-dimensional gel electrophoresis-mass spectrometry. *Proteomics* 2005;6:1714–1728. [PubMed: 15800970]
8. Tada T, Ohkubo I, Niwa M, Sasaki M, Tateyama H, Eimoto T. Immunohistochemical localization of Zn-alpha 2-glycoprotein in normal human tissues. *J Histochem Cytochem* 1991;9:1221–1226. [PubMed: 1918940]
9. Ghafouri B, Kihlstrom E, Stahlbom B, Tagesson C, Lindahl M. PLUNC (palate, lung and nasal epithelial clone) proteins in human nasal lavage fluid. *Biochem Soc Trans Pt* 2003;4:810–814.
10. Bundred NJ, Scott WN, Davies SJ, Miller WR, Mansel RE. Zinc alpha-2 glycoprotein levels in serum and breast fluids: a potential marker of apocrine activity. *Eur J Cancer* 1991;5:549–552. [PubMed: 1828959]
11. Haagensen DE, Stewart P, Dilley WG, Wells SA. Secretion of breast gross cystic disease fluid proteins by T47D breast cancer cells in culture--modulation by steroid hormones. *Breast Cancer Res Treat* 1992;1–2:77–86.
12. Simard J, de Launoit Y, Haagensen DE, Labrie F. Additive stimulatory action of glucocorticoids and androgens on basal and estrogen-repressed apolipoprotein-D messenger ribonucleic acid levels and secretion in human breast cancer cells. *Endocrinology* 1992;3:1115–1121. [PubMed: 1537279]
13. Boutteville C, Revillion F, Vandewalle B, Lefebvre J. [Hormonal regulation of GCDFP-15 secretion in breast tumors. Study of cultured cells]. *Bull Cancer* 1989;8:805–811. [PubMed: 2620105]
14. Simard J, Hatton AC, Labrie C, Dauvois S, Zhao HF, Haagensen DE, Labrie F. Inhibitory effect of estrogens on GCDFP-15 mRNA levels and secretion in ZR-75-1 human breast cancer cells. *Mol Endocrinol* 1989;4:694–702. [PubMed: 2725533]
15. Simard J, Dauvois S, Haagensen DE, Levesque C, Merand Y, Labrie F. Regulation of progesterone-binding breast cyst protein GCDFP-24 secretion by estrogens and androgens in human breast cancer cells: a new marker of steroid action in breast cancer. *Endocrinology* 1990;6:3223–3231. [PubMed: 2351114]
16. Murphy LC, Tsuyuki D, Myal Y, Shiu RP. Isolation and sequencing of a cDNA clone for a prolactin-inducible protein (PIP). Regulation of PIP gene expression in the human breast cancer cell line, T-47D. *J Biol Chem* 1987;31:15236–15241. [PubMed: 3667631]
17. Zhou J, Ng S, Adesanya-Famuiya O, Anderson K, Bondy CA. Testosterone inhibits estrogen-induced mammary epithelial proliferation and suppresses estrogen receptor expression. *FASEB J* 2000;12:1725–1730. [PubMed: 10973921]
18. Topper YJ, Freeman CS. Multiple hormone interactions in the developmental biology of the mammary gland. *Physiol Rev* 1980;4:1049–1106. [PubMed: 7001510]

19. Ray DB, Horst IA, Jansen RW, Kowal J. Normal mammary cells in long term culture. I. development of hormone-dependent functional monolayer cultures and assay of alpha-lactalbumin production. *Endocrinology* 1981;2:573–583. [PubMed: 7192624]
20. Teng CT, Pentecost BT, Chen YH, Newbold RR, Eddy EM, McLachlan JA. Lactotransferrin gene expression in the mouse uterus and mammary gland. *Endocrinology* 1989;2:992–999. [PubMed: 2463910]
21. Vizoso F, Plaza E, Vazquez J, Serra C, Lamelas ML, Gonzalez LO, Merino AM, Mendez J. Lysozyme expression by breast carcinomas, correlation with clinicopathologic parameters, and prognostic significance. *Ann Surg Oncol* 2001;8:667–674. [PubMed: 11569783]
22. Webster J, Oxley D. Peptide mass fingerprinting: protein identification using MALDI-TOF mass spectrometry. *Methods Mol Biol* 2005;227–240. [PubMed: 16350956]
23. Rose DP. Hormones and growth factors in nipple aspirates from normal women and benign breast disease patients. *Cancer Detect Prev* 1992;1:43–51. [PubMed: 1551137]
24. Vizoso F, Sanchez LM, Diez-Itza I, Luz LM, Lopez-Otin C. Factors affecting protein composition of breast secretions from nonlactating women. *Breast Cancer Res Treat* 1992;3:251–258. [PubMed: 1463865]
25. Vizoso F, Diez-Itza I, Sanchez LM, Tuya AF, Ruibal A, Lopez-Otin C. Relationship between serum prolactin levels and protein composition of breast secretions in nonlactating women. *J Clin Endocrinol Metab* 1994;2:525–529. [PubMed: 8045972]
26. Bray GA. The underlying basis for obesity: relationship to cancer. *J Nutr* 2002;11(Suppl):3451S–3455S. [PubMed: 12421869]
27. Bray GA. Medical consequences of obesity. *J Clin Endocrinol Metab* 2004;6:2583–2589. [PubMed: 15181027]
28. Irwin ML, McTiernan A, Bernstein L, Gilliland FD, Baumgartner R, Baumgartner K, Ballard-Barbash R. Relationship of obesity and physical activity with C-peptide, leptin, and insulin-like growth factors in breast cancer survivors. *Cancer Epidemiol Biomarkers Prev* 2005;12:2881–2888. [PubMed: 16365005]
29. Quinkler M, Sinha B, Tomlinson JW, Bujalska IJ, Stewart PM, Arlt W. Androgen generation in adipose tissue in women with simple obesity--a site-specific role for 17beta-hydroxysteroid dehydrogenase type 5. *J Endocrinol* 2004;2:331–342. [PubMed: 15531721]
30. Liu Z, Chang GQ, Leibowitz SF. Apolipoprotein D interacts with the long-form leptin receptor: a hypothalamic function in the control of energy homeostasis. *FASEB J* 2001;7:1329–1331. [PubMed: 11344130]
31. Bing C, Bao Y, Jenkins J, Sanders P, Manieri M, Cinti S, Tisdale MJ, Trayhurn P. Zinc-alpha2-glycoprotein, a lipid mobilizing factor, is expressed in adipocytes and is up-regulated in mice with cancer cachexia. *Proc Natl Acad Sci U S A* 2004;8:2500–2505. [PubMed: 14983038]
32. Gohda T, Makita Y, Shike T, Tanimoto M, Funabiki K, Horikoshi S, Tomino Y. Identification of epistatic interaction involved in obesity using the KK/Ta mouse as a Type 2 diabetes model: is Zn-alpha2 glycoprotein-1 a candidate gene for obesity? *Diabetes* 2003;8:2175–2181. [PubMed: 12882938]
33. Manni A, Santen RJ, Boucher AE, Lipton A, Harvey H, Drago J, Rohner T, Haagensen D, Glode M, Santner SJ. Evaluation of CEA and GCDFP-15 plasma level during hormonally induced cancer stimulation. *Anticancer Res* 1984;3:141–144. [PubMed: 6380398]
34. Haagensen DE Jr, Dilley WG, Mazoujian G, Wells SA Jr. Review of GCDFP-15. An apocrine marker protein. *Ann N Y Acad Sci* 1990;161–173. [PubMed: 2192632]
35. Alexander H, Stegner AL, Wagner-Mann C, Du Bois GC, Alexander S, Sauter ER. Proteomic analysis to identify breast cancer biomarkers in nipple aspirate fluid. *Clin Cancer Res* 2004;22:7500–7510. [PubMed: 15569980]
36. Wrensch MR, Petrakis NL, Gruenke LD, Ernster VL, Miike R, King EB, Hauck WW. Factors associated with obtaining nipple aspirate fluid: analysis of 1428 women and literature review. *Breast Cancer Res Treat* 1990;1:39–51. [PubMed: 2183892]

Figure 1.

1D gel (12.5% polyacrylamide) image of different types of NAF proteomes and breast milk proteomes. Lf, lactoferrin; Alb, albumin; Ig-H, immunoglobulin heavy chain; Ig-L, immunoglobulin light chain; Lyz, lysozyme C; Lalb, α -lactalbumin.

Table 1

The frequency and %-distribution of different protein profiles in our study population by status of recent childbirth

Secretion type	All subjects	Childbirth within the last 4 years	
		Yes	No
Type I	74 (58%)	9 (24%)	65 (72%)
Type II	42 (33%)	24 (63%)	18 (20%)
Mixed type	12 (9%)	5 (13%)	7 (8%)
Total	128 (100%)	38 (100%)	90 (100%)

Table 2
Characteristics (Mean $\pm$ SD, unless otherwise specified) of our study population by secretion types

Variables	All subjects N=128	Type I N=74	Type II N=42	Mixed type N=12	P ^a
Demographics					
Age	36.5 $\pm$ 2.5	36.9 $\pm$ 2.4	35.7 $\pm$ 2.3	36.0 $\pm$ 3.2	0.01
Race/Ethnicity, column percentage					
Caucasian	58.3	57.5	57.1	66.7	0.21
Hispanic	28.4	27.4	31	25	
African-American	10.2	13.7	4.8	8.3	
Asian	1.6	0	4.8	0	
Other	1.6	1.4	2.4	0	
Reproductive history					
Contraceptive usage, column percentage					
Never	27	28.4	24.4	27.3	0.64
Past	73	71.6	75.6	72.7	
Breastfeeding, column percentage					
No (including nulliparous subjects)	28.7	42.3	12.5	0	0.002
Yes	71.3	57.7	87.5	100	
Age of menarche	12.5 $\pm$ 1.6	12.5 $\pm$ 1.6	12.5 $\pm$ 1.5	11.8 $\pm$ 1.3	0.88
Parity	2.3 $\pm$ 1.3	2.1 $\pm$ 1.3	2.6 $\pm$ 1.4	2.2 $\pm$ 1.1	0.04
Age at first childbirth (parous subjects only)	24.2 $\pm$ 5.2	23.6 $\pm$ 4.8	25.4 $\pm$ 5.5	23.9 $\pm$ 6.0	0.08
Years since last childbirth (parous subjects only)	7.2 $\pm$ 4.8	8.8 $\pm$ 4.7	4.9 $\pm$ 4.1	6.4 $\pm$ 4.1	<0.0001
Anthropometrics					
Weight (kg)	72.5 $\pm$ 14.1	75.6 $\pm$ 13.4	69.8 $\pm$ 15.0	62.0 $\pm$ 7.8	0.03
Height (cm)	161.5 $\pm$ 7.2	161.8 $\pm$ 7.6	161.6 $\pm$ 7.3	159.4 $\pm$ 4.7	0.91
BMI (kg/m ²)	27.8 $\pm$ 5.3	28.9 $\pm$ 5.3	26.8 $\pm$ 5.2	24.5 $\pm$ 3.4	0.02
%-Body fat	35.5 $\pm$ 6.7	37.1 $\pm$ 6.2	34.0 $\pm$ 6.8	31.1 $\pm$ 6.6	0.02
Total fat mass (kg)	26.3 $\pm$ 9.4	28.7 $\pm$ 9.0	23.8 $\pm$ 9.5	19.6 $\pm$ 5.8	0.008
Total lean mass (kg)	45.8 $\pm$ 6.3	47.0 $\pm$ 6.2	44.9 $\pm$ 6.8	42.4 $\pm$ 3.5	0.09

Variables	All subjects N=128	Type I N=74	Type II N=42	Mixed type N=12	p ^a
Nutrient intake					
Total calorie (kcal)	1747 ± 457	1754 ± 477	1694 ± 403	1879 ± 511	0.5
Total fat (g)	74.9 ± 24.9	76.7 ± 24.8	71.6 ± 26.7	75.3 ± 20.0	0.3
%-Calorie from fat	38.0 ± 6.6	38.8 ± 6.5	37.2 ± 7.0	35.8 ± 5.9	0.24
Saturated fat (g)	24.9 ± 8.3	25.9 ± 8.9	23.4 ± 7.1	24.0 ± 8.0	0.13
%-Calorie from saturated fat	12.7 ± 2.8	13.1 ± 2.9	12.2 ± 2.3	11.3 ± 2.6	0.09
Monounsaturated fat (g)	29.2 ± 10.1	29.8 ± 9.4	28.2 ± 11.7	29.1 ± 7.3	0.41
%-Calorie from monounsaturated fat	14.8 ± 2.9	15.1 ± 2.7	14.5 ± 3.4	14.2 ± 2.1	0.31
Polyunsaturated fat (g)	14.8 ± 6.9	15.2 ± 7.1	14.0 ± 7.3	15.0 ± 4.5	0.36
%-Calorie from polyunsaturated fat	7.4 ± 2.2	7.5 ± 2.3	7.1 ± 2.3	7.3 ± 1.8	0.38
Hormones					
Testosterone (ng/ml)	0.71 ± 0.38	0.73 ± 0.42	0.59 ± 0.35	0.65 ± 0.24	0.41
17β-Estradiol (pg/ml)	75.7 ± 31.8	77.9 ± 32.7	67.6 ± 24.0	92.3 ± 23.3	0.08
Progesterone (ng/ml)	11.0 ± 5.6	10.8 ± 5.5	11.4 ± 4.9	10.6 ± 7.8	0.61
T3 (ng/dl)	121.2 ± 36.1	119.7 ± 32.7	123.7 ± 41.6	101.5 ± 22.2	0.59
T4 (mcg/dl)	8.3 ± 1.5	8.2 ± 1.3	8.5 ± 1.9	7.3 ± 0.6	0.57
TSH (μIU/ml)	1.9 ± 1.1	1.9 ± 1.1	2.0 ± 1.3	1.8 ± 1.3	0.15
Lipids					
VLDL (mg/dl)	20.6 ± 12.6	21.2 ± 13.7	20.1 ± 11.6	18.9 ± 7.5	0.67
LDL (mg/dl)	107.1 ± 29.6	111.0 ± 29.2	99.6 ± 32.0	107.3 ± 18.5	0.05
HDL (mg/dl)	53.5 ± 10.9	52.2 ± 9.7	55 ± 12.9	57.4 ± 9.8	0.18
TRIG (mg/dl)	103.1 ± 62.7	105.9 ± 68.3	100.5 ± 58.2	94.6 ± 37.5	0.67
CHOL (mg/dl)	181.1 ± 33.0	184.2 ± 33.7	174.7 ± 34.5	183.6 ± 19.7	0.15

^aall comparisons are between women with type I and type II NAF profiles

Table 3

Multi-level logistic regression models for predicting type I secretors of NAF

Models ^a	Variables ^b			
	Age OR (95% CI)	Years since last childbirth OR (95% CI)	Total fat mass OR (95% CI)	%-Calorie from saturated fat OR (95% CI)
Model 1 (N=104) Reproductive	1.1 (0.9, 1.4)	2.8 (1.5, 5.1)		
Model 2 (N=99) Anthropometrics	1.0 (0.8, 1.3)	2.9 (1.4, 5.8)	2.9 (1.6, 5.5)	
Model 3 (N=98) Nutrients	1.1 (0.9, 1.4)	2.6 (1.3, 5.1)	3.2 (1.6, 6.3)	4.2 (1.2, 15.1)
Model 4 (N=96, final) Hormones and Lipids	1.1 (0.9, 1.4)	2.6 (1.3, 5.2)	3.0 (1.5, 6.1)	4.1 (1.1, 14.6)

^a Model 1: reproductive variables entered in the model included parity, breastfeeding, age at first childbirth, and years since last childbirth

Model 2: model 1 plus anthropometric parameters including weight, BMI, %-body fat, total fat mass and total lean mass

Model 3: model 2 plus %-calorie from saturated fat

Model 4: model 3 plus estradiol and LDL.

^b Age: included in all the models; Years since last childbirth: ordinal data

Total fat mass: increment of 10 kg; %-Calorie from saturated fat: increment of 5%