Insight in development and functional differentiation of the mammary gland; a chromatin perspective. - *Monique Rijnkels*[*], Courtneay Freeman-Zadrowski[*], Joseph Hernandez[*], Vani Potluri[*], Austin Lin[*], Gonzalo Rincon[#], Alma Islas-Trejo[#], Liguo Wang[‡], Wei Li[‡], Juan F. Medrano[#]. [*]USDA/ARS Children's Nutrition Research Center, department of Pediatrics, Baylor College of Medicine, Houston TX 77030, USA.

Introduction

The mammary gland and lactation are unique to mammals; mother's milk provides all the nutrients needed for normal development and growth of the neonate. Comparative and evolutionary analysis have indicated the importance of lactation for the survival of mammals as well as the adaptation of milk composition to nutritional en environmental requirements (Brawand, Wahli & Kaessmann 2008; Elsik, Tellam, Worley *et al*. 2009; Lefevre, Sharp & Nicholas 2009; Lemay, Lynn, Martin *et al*. 2009). At the DNA level both gene sequence and copy number variation contribute to the diversity of milk composition among mammals (Elsik, Tellam, Worley *et al*. 2009; Lemay, Lynn, Martin *et al*. 2009). However, recent studies suggest that other mechanisms can also contribute to the observed diversity (Elsik, Tellam, Worley *et al*. 2009; Lemay, Lynn, Martin *et al*. 2009). Such mechanisms could include genetic variation in regulatory elements and changes at the chromatin level.

Mammary gland development starts during embryonic development however most of the development and functional differentiation occurs after birth. During the developmental windows of puberty, pregnancy, lactation and involution the gland undergoes profound morphological and functional changes (Watson & Khaled 2008). These changes occur in conjunction with changes in gene expression patterns and are considered to comprise a series of cell-lineage commitments and cell-fate decisions (LaMarca & Rosen 2008; Watson & Khaled 2008; Visvader 2009). A number of the factors involved in these morphological and cell commitment changes have been identified. However the fact that none of these factors are mammary gland specific suggests another layer of regulation contributing to the spatial and temporal expression needed to produce a fully functional mammary gland.

Three dimensions of gene regulation.

The underpinnings of functional development are the changes in gene expression that occur and are maintained with cell differentiation. Fundamental to gene expression are the transcription of DNA sequences into RNA and post transcriptional regulation—RNA processing and turnover. However, regulation of gene transcription takes place in the 3-dimensional space of the nucleus. A large amount of DNA is squeezed into a small space and needs to be organized in such a way that it is available for transcription at the right time in the right place and cell type. To achieve this organization DNA is carefully packaged into the nucleus. The way DNA is packaged determines the accessibility of the DNA code to regulatory and transcriptional machinery. Chromatin provides the framework for DNA packaging; the DNA is looped around histone octamers, these octamers are post-translationally modified which affects the further folding of the chromatin fiber. Chromatin plays an important role in both immediate transcriptional regulation and cell memory driving development and differentiation.

[*]USDA/ARS Children's Nutrition Research Center, department of Pediatrics, Baylor College of Medicine, Houston TX 77030, USA.
[‡]Duncan Cancer Center, Baylor College of Medicine, Houston TX 77030, USA
[#]Department of Animal Science, University of California Davis, Davis CA 95616, USA

Gene transcription is regulated at three hierarchical levels (van Driel, Fransz & Verschure 2003; Misteli 2007; Babu, Janga, de Santiago *et al*. 2008) that comprise: DNA sequence and the factors that bind to DNA, and the way the DNA is packaged and organized in the nucleus as described below, and shown in Figure 1.

Level 1: Linear DNA sequence (regulatory elements) and DNA binding factors (TF).
This level is affected by the activity, binding capacity, and cellular localization of TF and by DNA sequence variation.

Level 2: Chromatin conformation.
Chromatin conformation can be closed (more compact) or open; chromatin status can be identified by DNase1 hyper sensitivity (DHS)—open chromatin—, various post-translational histone modification representing open or closed chromatin depending on the modification (Jenuwein & Allis 2001), and DNA methylation (DNAme)—in general associated with closed chromatin conformations. Specific protein complexes can remodel the positioning and histone modification status of nucleosomes. Furthermore, different chromosomal regions can interaction with other regions or the nuclear matrix affecting the higher order organization of the chromatin.

Level 3: Nuclear organization
This level of regulation is associated with the position of a genomic locus in chromosome territories (CT) and the position within the nuclear volume, interaction of chromatin/genomic regions with specialized nuclear bodies (e.g. Transcription factories) or association with heterochromatin (Misteli 2007).

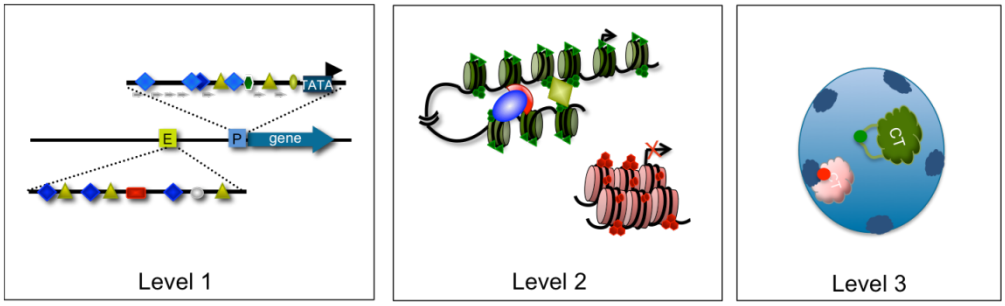

Figure 1: Hierarchical levels of gene regulation.
Level1: DNA sequence and factors that bind to DNA (TF), Enhancer (E) and Promoter (P) diverse shapes and colors indicate TF binding sites.; Level2: chromatin conformation. Green triangles indicate positive histone modifications on active/open chromatin, variously colored shapes depict TF factor binding DNA and their involvement in Chromatin loop formation, red triangles and ovals indicate negative histone modifications and DNA methylation in closed/inactive chromatin; Level 3: nuclear organization, dark blue areas depict heterochromatin red dot a negatively regulated genomic region interacting with heterochromatin, green dot an active genomic region looping out from its chromosome domain (CT)

Insights in development and functional differentiation of the mammary gland from a chromatin perspective.

Milk protein genes are some of the most studied mammary gland expressed genes due to their importance in the dairy industry and the fact that they are markers for functional differentiation of the gland. Current insights on the 3 levels of gene regulation as they apply to milk protein gene regulation are discussed below.

Level 1: linear DNA and transacting-DNA-binding factors (TF)

In the mammary gland many signal transducers and DNA binding factors and the sequences they bind during gene regulation have been identified (Hennighausen & Robinson 2005; Watson & Khaled 2008). For some important roles have been uncovered both in cell lineage/cell fate decisions underlying functional differentiation and in directly regulating expression of milk protein genes during lactation, eg STAT5 (Yamaji, Na, Feuermann *et al*. 2009). Many milk protein genes share similar regulatory motifs (Rosen, Wyszomierski & Hadsell 1999; Martin, Szymanowska, Zwierzchowski *et al*. 2002). Indicating that they are downstream of the same developmental signaling pathways. Recent studies in mammary epithelial cell lines have shown that lactogenic hormone induction leads to the recruitment of the factors that transduce the lactogenic signal (e.g. STAT5, Glucocorticoid receptor (GR), C/EBPbeta etc) to the regulatory elements of milk protein genes, while displacing inhibitory factors (Kabotyanski, Huetter, Xian *et al*. 2006; Xu, Spencer & Bissell 2007; Kabotyanski, Rijnkels, Freeman-Zadrowski *et al*. 2009; Xu, Nelson, Muschler *et al*. 2009).

Spatial organization of DNA in the genome can also contribute to it regulation. 25-30% of genes expressed in the mammary gland have been found to cluster in the bovine genome, similar to findings in other tissues and species (Lemay, Lynn, Martin *et al*. 2009). It is thought that this clustering is favored to facilitate coordinate regulation, which could in part be at the chromatin and nuclear organization level (see below)

Level 2: Chromatin conformation.

Studies in cell lines have shown changes in chromatin modifications and recruitment of chromatin remodeling complexes in the presence of lactogenic hormones and Extra-Cellular-Matrix (ECM) (Plachot & Lelievre 2004; Kabotyanski, Huetter, Xian *et al*. 2006; Le Beyec, Xu, Lee *et al*. 2007; Xu, Spencer & Bissell 2007; Kabotyanski, Rijnkels, Freeman-Zadrowski *et al*. 2009; Lelievre 2009; Xu, Nelson, Muschler *et al*. 2009). Further more it was shown that the beta-casein gene promoter engages in a physical interaction with the distal beta casein enhancer upon lactogenic hormone stimulation (Kabotyanski, Rijnkels, Freeman-Zadrowski *et al*. 2009). Studies on mouse tissue confirm this interaction in lactating mammary gland and its absence in virgin mammary epithelial cells (Kabotyanski, Rijnkels, Freeman-Zadrowski *et al*. 2009). Withdrawal of the lactogenic stimulus results in a decrease in beta-casein expression concomitant with a loss of detectable interaction between the promoter and enhancer (Rijnkels, Kabotyanski, Montazer-Torbati *et al*. 2010). Furthermore several studies using lactating mammary gland tissue of different species have shown that lactation in general corresponds to an open chromatin conformation while earlier developmental stages and non-mammary tissue have a more closed chromatin conformation around milk protein genes, recently reviewed in (Rijnkels, Kabotyanski, Montazer-Torbati *et al*. 2010).

In recent analyses of the casein gene cluster we showed that this genomic region is enriched for histone modifications corresponding to an open chromatin conformation in the lactating mammary gland compared to the non-lactating gland and non-mammary tissue (e.g. liver) ((Rijnkels, Freeman-Zadrowski & Hernandez 2009; Rijnkels, Kabotyanski, Montazer-Torbati *et al*. 2010), manuscript in preparation). The promoters have an open chromatin conformation in the lactating mammary gland—marked by DNAse 1 Hypersensitivity (DHS), certain histone modifications and lower levels of DNA methylation—and a more closed conformation in the non-lactating gland and non-mammary tissue (e.g. liver)—marked by absence of histone modification associated with open chromatin and higher levels of DNAme. Whereas a number of potential distal regulatory elements, previously identified as evolutionary conserved regions (Rijnkels, Elnitski, Miller *et al*. 2003), have or gain a more open chromatin conformation during earlier development, in puberty. This suggests that mammary epithelial cells (MEC) undergo a progressive change in chromatin conformation during functional differentiation of the gland. Singh and colleagues (Singh, Erdman, Swanson *et al*. 2010) recently reported an increase of DNA methylation (marker for closed chromatin) upon cessation of lactation in a regulatory region distal to the bovine alpha s1 casein gene. Vanselow *et al*. (Vanselow, Yang, Herrmann *et al*. 2006) showed an increase of DNA methylation and DNA compaction in this region during mastitis. In both instances the increase of DNAme corresponds to the loss of alpha-s1-casein gene expression.

Similar to the findings for the casein gene cluster the genomic region harboring the mouse Whey Acidic Protein gene (WAP) has open chromatin marks during lactation, which are not present in the non-lactating virgin gland and liver ((Millot, Montoliu, Fontaine *et al*. 2003; Montazer-Torbati, Hue-Beauvais, Droineau *et al*. 2008; Rijnkels, Kabotyanski, Montazer-Torbati *et al*. 2010), Rijnkels unpublished). For the WAP gene it was shown that interactions with the nuclear matrix is hormone dependent as well (Rijnkels, Kabotyanski, Montazer-Torbati *et al*. 2010). These interactions result in chromatin loop formations that are thought to influence access of regulatory complexes to regulatory elements.

Level 3: Nuclear architecture/organization.
It has been observed that the nuclear organization, particularly the organization and distribution of heterochromatic regions (heavily compacted and usually inactive chromatin regions) and certain nuclear proteins, changes with the differentiation status of the MEC (Spencer, Xu & Bissell 2007; Lelievre 2009; Kress, Ballester, Devinoy *et al*. 2010). At the nuclear level it has been shown that the casein gene nuclear position changes upon lactogenic induction. Further more it has been found that the mouse casein gene cluster region is mostly at the periphery or outside of its chromosome 5 CT in a lactogenic hormone stimulated mouse mammary epithelial cell line while residing mostly inside of the CT in non-stimulated or non-mammary cells (Ballester, Kress, Hue-Beauvais *et al*. 2008; Kress, Ballester, Devinoy *et al*. 2010)

Implication for genetics in livestock production
Lactation performance in domestic animals has largely been improved by genetic selection. Presumably selecting for genetic variants that enhance the development and function of the

mammary gland. The availability of the bovine genome sequence and high-density single nucleotide polymorphisms (SNP)-chips has significantly enhanced the identification of SNP associations with production traits and variations in gene expression. However, the actual causal variants have only been identified in rare cases (Grisart, Coppieters, Farnir *et al*. 2002; Olsen, Nilsen, Hayes *et al*. 2007).

Historically the search for causative SNP has been focused on protein encoding sequences. However, genetic variation in (distal) regulatory elements (RE) might be just as important as genetic variation within the protein coding regions of a gene, as is illustrated for cytokine and cytokine receptor polymorphisms in (Smith & Humphries 2009). Genetic variation in RE can affect the binding of TF by directly affecting transcription, or affecting the long-distance interaction of regulatory elements. We recently identified a number of SNP in distal regulatory elements in the casein gene cluster. Several of these SNP are associated with production traits and gene expression (Rincon, Rijnkels, Islas *et al*. 2009); based on the location of the SNP some are predicted to affect binding of TF and the physical interaction between regulatory elements. Therefore these SNP could have a direct effect on the expression of associated genes. This also suggests that SNP could affect regulatory functions of RE (e.g. enhancer; (Smith & Humphries 2009; Wright, Brown & Cole), or proper chromatin conformation (Smith & Humphries 2009) and nuclear positioning.

In the context of the 3-dimensions of gene regulation and the importance of chromatin dynamics for normal gene regulation, one can imagine that genetic changes that affect TF-binding, chromatin or chromatin related genes could underlie the phenotypic variation of production traits. Allele specific TF binding and chromatin signatures have recently been described in humans (Kasowski, Grubert, Heffelfinger *et al*. ; McDaniell, Lee, Song *et al*.).

Furthermore, chromatin status is involved in conveying cellular memory. Interestingly, external signals such as environmental exposures, nutrition and disease can influence chromatin status. Thus past exposures, e.g. during (pre-) puberty or pregnancy, can result in a cellular memory that influence subsequent lactation performance. In genetic evaluation modeling genetic effects can explain only part (21%) of the phenotypic variation in milk production, while permanent environment effects account for ~35% (Bormann, Wiggans, Druet *et al*. 2002)
The interaction of these influences with SNP and chromatin needs to be explored in more detail. Another aspect of this is the contribution of mammary gland Stem- and Progenitor-cells to lactation performance. The number and secretory activity of the MEC underlies milk production. Chromatin plays a role in propagating changes that take place in stem and progenitor cells to mature cells—cellular memory. So early external influences are important in shaping epigenotype of MEC and functioning of the mammary gland.

Non-protein-coding (nc)RNA can play a role in gene regulation at all three levels of gene regulation; Micro (mi)RNA's are thought to regulate RNA post-transcriptionally, but in some cases might regulate it at the transcriptional levels as well, Long intergenic non-coding (Linc)RNAs have been found to interact with chromatin modifying complexes and influence chromatin conformation as well as nuclear organization (Kapranov, Cheng, Dike *et al*. 2007; Amaral & Mattick 2008; Mattick, Amaral, Dinger *et al*. 2009).

Studies on ncRNA in the mammary gland have found differential expression of miRNAs and LincRNAs in the developing mammary gland (Ginger, Shore, Contreras *et al*. 2006; Gu, Eleswarapu & Jiang 2007; Avril-Sassen, Goldstein, Stingl *et al*. 2009; Sdassi, Silveri, Laubier *et al*. 2009). Thus, genetic variation in the genes encoding ncRNA or elements regulating them might also have a profound effect on mammary function. Especially, because gene regulation through miRNAs and Linc RNA can have many targets.

Prospects:

Integration of chromatin status (epigenetics), structural variation (SNP, CNV), gene expression and signaling pathways will be needed to develop a better understanding of the functioning of the mammary gland so it can be applied to Livestock production. Specifically, it is important to examine the effects of farm management practices that can influence the epigenome and their interaction with genetic factors and their effects on lactation performance. Together, this should lead to the development of new strategies to enhance lactation performance in livestock.

References

Amaral P.P. & Mattick J.S. *Mamm Genome* **19**, 454-92 (2008).
Avril-Sassen S., Goldstein L.D., Stingl J., et al. *BMC Genomics* **10**, 548 (2009).
Babu M.M., Janga S.C., de Santiago I., et al. *Curr Opin Genet Dev* **18**, 571-82 (2008).
Ballester M., Kress C., Hue-Beauvais C., et al. *J Cell Biochem* **105**, 262-70 (2008).
Bormann J., Wiggans G.R., Druet T., et al. *J. Dairy Sci*. **85**, 263- (2002).
Brawand D., Wahli W. & Kaessmann H. *PLoS Biol* **6**, e63 (2008).
Elsik C.G., Tellam R.L., Worley K.C., et al. *Science* **324**, 522-8 (2009).
Ginger M.R., Shore A.N., Contreras A., et al. *Proc Natl Acad Sci U S A* **103**, 5781-6 (2006).
Grisart B., Coppieters W., Farnir F., et al. *Genome Res* **12**, 222-31 (2002).
Gu Z., Eleswarapu S. & Jiang H. *FEBS Lett* **581**, 981-8 (2007).
Hennighausen L. & Robinson G.W. *Nat Rev Mol Cell Biol* **6**, 715-25 (2005).
Jenuwein T. & Allis C.D. *Science* **293**, 1074-80 (2001).
Kabotyanski E.B., Huetter M., Xian W., et al. *Mol Endocrinol* **20**, 2355-68 (2006).
Kabotyanski E.B., Rijnkels M., Freeman-Zadrowski C., et al. *J Biol Chem* **284**, 22815-24 (2009).
Kapranov P., Cheng J., Dike S., et al. *Science* **316**, 1484-8 (2007).
Kasowski M., Grubert F., Heffelfinger C., et al. *Science* **328**, 232-5 (2010).
Kress C., Ballester M., Devinoy E., et al. *J Mammary Gland Biol Neoplasia* **15**, 73-83 (2010).
LaMarca H.L. & Rosen J.M. *Endocrinology* **149**, 4317-21 (2008).
Le Beyec J., Xu R., Lee S.Y., et al. *Exp Cell Res* **313**, 3066-75 (2007).
Lefevre C.M., Sharp J.A. & Nicholas K.R. *Reprod Fertil Dev* **21**, 1015-27 (2009).
Lelievre S.A. *Biochim Biophys Acta* **1790**, 925-35 (2009).
Lemay D.G., Lynn D.J., Martin W.F., et al. *Genome Biol* **10**, R43 (2009).
Martin P., Szymanowska M., Zwierzchowski L., et al. *Reprod Nutr Dev* **42**, 433-59 (2002).

Mattick J.S., Amaral P.P., Dinger M.E., et al. *Bioessays* **31**, 51-9 (2009).
McDaniell R., Lee B.K., Song L., et al. *Science* **328**, 235-9 (2010).
Millot B., Montoliu L., Fontaine M.L., et al. *Biochem J* **372**, 41-52 (2003).
Misteli T. *Cell* **128**, 787-800 (2007).
Montazer-Torbati M.B., Hue-Beauvais C., Droineau S., et al. *Exp Cell Res* **314**, 975-87 (2008).
Olsen H.G., Nilsen H., Hayes B., et al. *BMC Genet* **8**, 32 (2007).
Plachot C. & Lelievre S.A. *Exp Cell Res* **298**, 122-32 (2004).
Rijnkels M., Elnitski L., Miller W., et al. *Genomics* **82**, 417-32 (2003).
Rijnkels M., Freeman-Zadrowski C. & Hernandez J. *J.Dairy Sci*. **92**, 325-6 (2009).
Rijnkels M., Kabotyanski E., Montazer-Torbati M.B., et al. *J Mammary Gland Biol Neoplasia* **15**, 85-100 (2010).
Rincon G., Rijnkels M., Islas A., et al. *J. Dairy Sci* **92**, T34 (2009).
Rosen J.M., Wyszomierski S.L. & Hadsell D. *Annu Rev Nutr* **19**, 407-36 (1999).
Sdassi N., Silveri L., Laubier J., et al. *BMC Genomics* **10**, 149 (2009).
Singh K., Erdman R.A., Swanson K.M., et al. *J Mammary Gland Biol Neoplasia* **15**, 101-12 (2010).
Smith A.J.P. & Humphries S.E. *Cytokine & Growth Factor Reviews* **20**, 43-59 (2009).
Spencer V.A., Xu R. & Bissell M.J. *Adv Cancer Res* **97**, 275-94 (2007).
van Driel R., Fransz P.F. & Verschure P.J. *J Cell Sci* **116**, 4067-75 (2003).
Vanselow J., Yang W., Herrmann J., et al. *J Mol Endocrinol* **37**, 463-77 (2006).
Visvader J.E. *Genes Dev* **23**, 2563-77 (2009).
Watson C.J. & Khaled W.T. *Development* **135**, 995-1003 (2008).
Wright J.B., Brown S.J. & Cole M.D. *Mol. Cell. Biol.* **30**, 1411-20 (2010).
Xu R., Nelson C.M., Muschler J.L., et al. *J Cell Biol* **184**, 57-66 (2009).
Xu R., Spencer V.A. & Bissell M.J. *J Biol Chem* **282**, 14992-9 (2007).
Yamaji D., Na R., Feuermann Y., et al. *Genes Dev* **23**, 2382-7 (2009).