1

Transcription in the *Xenopus* Oocyte Nucleus

Joseph G. Gall
Department of Embryology, Carnegie Institution for Science, Baltimore, MD

**Abstract:** The mature oocyte of *Xenopus* is a gigantic cell with a diameter of 0.8 mm in *Xenopus tropicalis* and 1.2 mm in *Xenopus laevis*. It stores a large number of stable mRNAs for use during early development, all of which are transcribed by the giant lampbrush chromosomes inside the equally giant oocyte nucleus or germinal vesicle. The lampbrush chromosomes are specialized for an unusually high rate of transcription, but even so they require months to produce the enormous number of stable transcripts needed for early embryogenesis. Deep sequencing of oocyte mRNA reveals a wide variety of transcripts made by the lampbrush chromosomes during oogenesis.

**Introduction**

Oocytes of animals vary greatly in size, rate of growth, presence or absence of a quiescent stage, and association with supporting or nurse cells of various types (Davidson 1986; Voronina and Wessell 2003). These factors influence the nature of the transcription that takes place in the oocyte nucleus or germinal vesicle (GV). The *Xenopus* oocyte represents one extreme. Its oocyte grows to an enormous size, up to 1.2 mm in *Xenopus laevis* and 0.8 mm in *Xenopus tropicalis*, and there are no nurse cells (Figure 1.1). At their maximal size, the oocytes of *X. laevis* and *X. tropicalis* have volumes some $10^5$–$10^6$ times that of a typical somatic cell. All of the transcripts for this enormous cell must be synthesized by the single GV. The strategy used by the oocyte to accomplish this prodigious task involves three major components. First, the chromosomes in the GV transcribe at what is probably close to the theoretical maximum, giving rise to the remarkable lampbrush chromosomes (LBCs) (http://projects.exeter.ac.uk/lampbrush/), which will be a major focus of this chapter. Second, and equally importantly, transcription continues for several months during the long period of oocyte development. Finally, the transcripts produced by the GV and stored in the cytoplasm are unusually stable. Only by a combination of these three features is the *Xenopus* oocyte able to make and store...
the transcripts needed for oogenesis and early embryogenesis.

LBCs similar to those of *Xenopus* are found in a wide range of organisms, both vertebrate and invertebrate (Callan 1986), and have even been described from a plant, the single-celled alga *Acetabularia* (Spring et al. 1975; Berger et al. 1994). It is worth emphasizing, however, that LBCs have been described only from large meiotic nuclei that provide transcripts to a large oocyte without help from nutritive cells. The situation can be very different in other organisms. For instance, the *Drosophila* oocyte is large but the GV is small and transcriptionally silent, or nearly so. In this case, there are no LBCs and transcripts are supplied to the growing oocyte by polyploid nurse cells (Spradling 1993). The example of *Drosophila* and other organisms with transcriptionally inactive GVs emphasizes the fact that LBCs are not required for meiosis or more generally for oogenesis (Gall 2012).

**LBC structure: The standard model**

Extensive studies on the LBCs of many organisms over the past 50–60 years have established what can be called the “standard model” of their physical structure. LBCs consist of four chromatids in the diplotene stage (G2) of the first meiotic division. Each chromatid is fundamentally a single, very long DNA double helix, which, if fully extended, would be centimeters in length (Callan and Maegregor 1958; Callan 1963; Gall 1963). The two homologues of each bivalent are independent of each other, except at the chiasmata, whose physical structure is almost completely obscure. It is the unique and variable association of sister chromatids that gives rise to the classic “lampbrush” condition. Specifically, there are condensed, transcriptionally inactive regions (chromomeres) along the major axis of each homologue, where sister chromatids are associated with each other. And there are transcriptionally active regions (loops) where sisters extend laterally from the axis independently of each other (Figure 1.2A and B). Each loop consists of one or more transcription units (TUs) that are visible at the light optical level as “thin-to-thick” regions, the thin end being where transcription initiates and the thick end where it terminates. The entire structure is visible primarily because the nascent RNA transcripts are associated with massive amounts of protein. These relationships are shown diagrammatically in Figure 1.3, variations of which have been published many times before (Gall 1956; Callan and Lloyd 1960; Hess 1971; Morgan 2002; Austin et al. 2009; Gaginskaya et al. 2009).

**Chromomeres and loops**

Beginning with the transcriptionally inactive axis of each homologue, we immediately run into unanswered structural issues. The more or less accepted view is that the axis consists of a series of DNA-rich chromomeres within which the sisters are tightly wound up in some fashion. They can be stained by various DNA-specific dyes, such as Feulgen or DAPI (Figure 1.2B and D). The chromomeres are separated by exceedingly delicate interchromomeric regions that are either invisible or barely
Figure 1.2  LBCs of the newt Notophthalmus viridescens (A and B) and X. tropicalis (C and D). (A) A short segment of an LBC stained with antibodies against pol II (green) and the RNA-binding protein CELF1 (red) (Morgan 2007). The axes of all loops appear as diffraction-limited green lines, because they are covered with closely spaced pol II molecules. One pair of sister chromatids is preferentially stained with CELF1, revealing the prominent thin-to-thick orientation of the associated loop matrix (RNP transcripts). (B) The same segment of LBC stained with the DNA-specific dye DAPI reveals the axis of transcriptionally inactive chromomeres. (C) Bivalent No. 2 of X. tropicalis stained with antibodies against pol II (green) and pol III (red). The vast majority of loops are transcribed by pol II. The loops of X. tropicalis are much shorter than those of the newt, and only a few are recognizable as loops in this image (arrow). (D) The same bivalent showing strong staining of the chromomere axes with DAPI. DAPI also reveals two amplified rDNA cores (arrowheads) in each of two extrachromosomal nucleoli. Regions of high protein concentration in the nucleoli also bind DAPI to a lesser extent. The same is true of two moderately stained structures near the middle of this bivalent (arrows), which represent loop pairs whose matrix has fused into a single large mass (lumpy loops). To see a color version of this figure; see Plate 1.
visible at the light optical level. By electron microscopy, these regions usually appear as a single fiber about 10 nm in thickness (Tomlin and Callan 1951; Mott and Callan 1975). Although an analogy of the chromomeres and interchromomeric regions to the bands and interbands of polytene chromosomes is often made, this analogy breaks down when examined closely. Specifically, the number of chromomeres varies greatly during development of the oocyte, there being dozens of chromomeres in an amphibian or avian LBC at maximal extension, but a decreasing number as the chromosomes contract in length for the first meiotic division. It is possible to construct maps of individual chromosomes based on the chromomere pattern at maximal extension, as has been done for avian LBCs (Rodionov, Galkina, and Lukina in Schmid et al. 2005), but it is often difficult to recognize a reproducible chromomere pattern in amphibian LBCs, even between the homologues of a given bivalent (Callan and Lloyd 1960). Macgregor (2012) discusses the “chromomere problem” in a recent essay.

To say that we are woefully ignorant about the internal structure of chromomeres is an understatement. The first question we might ask is whether sister chromatids are intimately paired inside the chromomere, as they are in the interchromomeric regions. Although we do not have an answer to that question, we can say definitively that a single chromatid can form either an entire LBC or part of one. The most direct evidence comes from LBCs that form when sperm heads are injected into a GV (Gall and Murphy 1998; Liu and Gall 2012). In such experiments, the single chromatids inside the sperm head are released within minutes and develop gradually into morphologically recognizable LBCs with transcriptionally active loops. Except that their loops are not paired, these LBCs are similar in overall organization to the normal LBCs in the same nucleus (Figure 1.4). A similar argument comes from the existence of “double-axis” regions of normal LBCs. Double-axis regions are segments of an LBC in which sisters are completely unpaired. Although rare, they are a regular feature of specific regions of certain chromosomes: one end of the shortest chromosome of Triturus cristatus (Callan and Lloyd 1960), near the middle of chromosome Nos. 8 and 9 of X. laevis (Figure A1.1), and roughly half of chromosome No. 10 of X. tropicalis (Figure A1.2). Although LBCs that consist of single chromatids, as well as the double-axis regions of otherwise typical LBCs, demonstrate that chromatids need not be paired to form typical “lampbrushes”, they do not directly address the organization of sister chromatids within the chromomeres of typical LBCs.

One structural issue on which there is no question is that sister chromatids form independent transcription loops. There is both observational and experimental evidence for this model, going back to Callan’s original stretching experiment (Callan 1957). Basically, Callan showed that an LBC chromosome “breaks” in a stereotypical and counterintuitive fashion when stretched between microneedles. Instead of breaking in the thinnest regions between the chromomeres, the chromosome doesn’t really break.
Transcription in the Xenopus Oocyte Nucleus

at all. Instead, something happens at the bases of the loops such that a pair of loops, which originally extended laterally, comes to lie along the main axis of the chromosome. Such “double-loop bridges (dlb)” also occur when chromosomes are accidentally stretched during preparation for microscopical examination (Figure 1.5). Moreover, certain pairs of identifiable loops exist normally in the dlb configuration (Callan 1954; Callan and Lloyd 1960). An interesting example is found on chromosome No. 3 of *X. laevis* (Figure A1.1). Here, a prominent dlb near the centromere contains an unusually high concentration of the RNA-editing enzyme ADAR1 (Eckmann and Jantsch 1999).

Callan’s experiment provided what is arguably the single most important insight into the LBC structure: that each lateral loop is part of an extraordinarily long and continuous chromatid. Coupled with the demonstration that a loop contains one DNA double helix, whereas the main axis contains two helices, LBCs provided critical evidence that the largest known chromosomes are not multistranded, but instead conform to the unineme hypothesis of chromosome structure (Gall 1963, 1981).

**Transcription on LBC loops**

The lateral loops are the most distinctive feature of LBCs and gave rise to the name “lampbrush”, which was coined by Rückert (1892) by analogy to the then familiar brushes used to clean soot from kerosene lamp chimneys. There is no question that the loops represent transcriptionally active regions of the chromosome, as opposed to the transcriptionally inactive chromomeres. The first hint came from the demonstration of RNase-sensitive staining in these regions (Gall 1954), followed by autoradiographic experiments showing that the loops incorporate RNA precursors such as adenine and uridine (Gall 1958; Gall and Callan 1962).

Well before there was detailed molecular evidence for transcription on the loops, the beautiful electron micrographs of Oscar Miller and his colleagues provided stunning images of TUs in amphibian oocytes at unprecedented resolution. Because “Miller spreads” involve disruption of the GV in distilled water, the overall organization of the chromosomes is lost. Nevertheless, it was abundantly evident that the (nonribosomal) “Christmas trees” were derived from the loops of LBCs (Miller and Hamkalo 1972; Hamkalo and Miller 1973; Scheer et al. 1976).
Immunofluorescent staining, especially when coupled with confocal or superresolution microscopy, now provides textbook images of active transcription on intact LBCs (Figures 1.2A, C, and 1.6). RNA polymerase II molecules form a diffraction-limited line along the axis of each loop, whereas ribonucleoprotein (RNP) transcripts appear as a massive coating around this axis. The thin-to-thick organization of loops early suggested the direction of transcription, and in the case of the histone loops of the newt Notophthalmus, it was even possible to correlate the direction of transcription with the strand of DNA being transcribed (Stephenson et al. 1981). Multiple thin-to-thick regions within a single loop demonstrated that a one-to-one correlation between the loops and TUs is not possible.
Instead, a loop consists of one or more TUs, not necessarily oriented in the same direction (Scheer et al. 1976; Gall et al. 1983).

Interestingly, pol III transcription also occurs on loops. Because pol III transcripts are usually short, they do not form a matrix detectable by phase contrast or differential interference contrast microscopy. Nevertheless, pol III loops can be seen when they are immunostained with antibodies against pol III (Figure 1.2C). If the loops are extended, they appear as diffraction-limited lines; otherwise, they are seen as irregular masses of stain close to the chromosome axis (Figures A1.1 and A1.2) (Murphy et al. 2002). What are possibly pol III loops can also be recognized in electron micrographs by their very short transcripts (Scheer 1981).

It is not known what holds sister chromatids together at the bases of the loops. One would imagine this to be a protein or more likely a complex of proteins, but no one has been lucky enough to find an antibody that stains just the bases of the loops. Perhaps this hypothetical glue at the bases of the loops corresponds to the insulators that separate the functional units of the chromosome (Giles et al. 2010).

As just noted, a loop is not the same as a TU, since many loops contain multiple TUs. Moreover, a repeated gene locus can be represented by multiple loops, as is true for the histone gene loci of Notophthalmus (Diaz et al. 1981). There are other cases where loops of similar morphology occur not in pairs but in clusters, again suggesting a complex and variable relationship among TUs, loops, and the underlying genes or gene clusters.

**Transcripts produced during oogenesis**

Transcripts stored in the cytoplasm

Ribosomal RNA is the most abundant RNA present in the cytoplasm of the oocyte, and it occurs at about the same concentration as in cells of normal size (Brown and Littna 1964). In X. laevis, there are about 500–800 copies of the rDNA genes at a single nucleolus organizer (Wallace and Birnstiel 1966), a number that is physically incapable of transcribing the total amount of rRNA produced during oogenesis. As shown a number of years ago, the genes coding for rRNA are amplified during the early stages of meiosis, giving rise to hundreds of transcriptionally active nucleoli (Figure 1.2D), which are physically separate from the LBCs (Peacock 1965; Miller 1966; Brown and Dawid 1968; Gall 1968; Perkowska et al. 1968). The 5S rRNA, which must be produced during oogenesis in equimolar amounts to the 18S and 28S rRNAs, is not generated from extrachromosomal copies. Instead, the X. laevis genome carries about 24,000 copies of a special oocyte-type 5S gene, which are transcribed specifically during oogenesis (Brown et al. 1971).

For protein-coding genes, the corresponding mRNAs are presumably all transcribed on the loops of the LBCs. It is beyond the scope of this chapter to consider the complexity of the mRNA stored in the cytoplasm, much of it for use during early embryogenesis, when transcription is shut down. The nature of this stored RNA has been the subject of investigation for many years; earlier studies are ably summarized in Davidson’s text *Gene Activity in Early Development* (Davidson 1986). With the advent of deep sequencing, it is now possible to examine the totality of stored transcripts in great detail. A recently published study from John Gurdon’s group detected cytoplasmic transcripts from over 11,000 genes of X. tropicalis (Simeoni et al. 2012), more than half of the 20,000 annotated genes in the genome (Hellsten et al. 2010). As shown by RT-PCR analysis for a selected subset, these transcripts range in abundance from more than 10⁷ copies per oocyte to less than a few hundred. We have also examined transcripts from mature X. tropicalis oocytes and found a similar wide range of abundance (Gardner et al. 2012). These data revive – or rather continue – an old debate about LBC transcription: do LBCs simply transcribe a set of oocyte-specific genes at an unusually high rate, or do they transcribe most or all genes as part of specific germline reprogramming of the genome?

We have recently addressed a more limited question about oocyte transcription. Are there major changes in the relative abundance of transcripts stored in the oocyte during the course of oogenesis? To answer this question,
we sequenced total oocyte RNA from X. tropicalis oocytes of different sizes, from less than 100 µm diameter to full-grown oocytes of about 800 µm (Figure 1.7). These data demonstrate three essential facts. First, from the beginning of oocyte development, the oocyte produces and stores transcripts from a wide variety of genes. Figure 1.7 shows data for approximately 9000 transcribed genes (specifically all genes with log_2 FPKM (fragments per kilobase per million reads) scores above 0). Second, these transcripts vary greatly in relative abundance, from transcripts that are just detectable at the read depth of our samples to some that are extremely abundant. Finally, the relative abundance of most transcripts changes very little during development of the oocyte, from well before the onset of yolk formation (oocytes about 100 µm diameter) all the way through until the mature oocyte.

Nascent transcripts on the LBCs

As just discussed, quantitative data are now available on the population of cytoplasmic transcripts stored during oocyte development. These transcripts are produced by the LBCs and in this respect they give insight into the nature of LBC transcription. However, fundamental questions will remain until there is detailed information about the nascent transcripts themselves and the nature of their processing. In an attempt to gain such data, we carried out a deep sequence analysis of GV RNA from X. tropicalis oocytes (Gardner et al. 2012). To our surprise, we found that the bulk of GV RNA consists of stable intronic sequences (sisRNA) derived from the same set of genes whose transcripts are found in the cytoplasm. There is a rough correlation between the abundance of a given mRNA and the abundance of sisRNA from the same gene, although the absolute amount of mRNA is much greater (molar ratio roughly 100 : 1). For technical reasons, it was not possible to analyze sisRNA after GV breakdown by deep sequencing, but RT-PCR analysis of specific sequences demonstrated that sisRNA persists in the embryo until at least the blastula stage, at which time transcription resumes. At present, the functional significance of sisRNA is completely unknown.

We should not have been surprised that nascent transcripts were missing from our deep sequence data. Despite its enormous size, the GV of X. tropicalis contains only four sets of chromosomes with a total of 6–8 pg of genomic DNA (Gregory 2006). On the basis of incorporation data, Davidson earlier estimated that a X. laevis GV (with about twice the amount of genomic DNA as X. tropicalis) transcribes roughly 1.4 ng of chromosomal RNA per day. The total amount of RNA in nascent transcripts must be still smaller. Thus, even in a sample of RNA derived from several hundred GVs, the total amount of nascent transcripts will be no more than a few picograms, below the detection level in our experiments.

**In situ hybridization of nascent transcripts on individual LBC loops**

Although global information about nascent transcripts must await the results of deep sequencing, specific transcripts have been investigated by in situ hybridization. The most complete analysis, carried out some...
years ago, involved the histone gene clusters in the newt *Notophthalmus* (Diaz et al. 1981; Stephenson et al. 1981; Gall et al. 1983; Diaz and Gall 1985). The basic finding was that individual LBC loops contain one or more clusters of the five histone genes, the clusters being separated by extremely long tracts of a 221-bp repeated “satellite” DNA. *In situ* hybridization with probes specific for the histone genes and for the satellite DNA showed that most of the RNA on the loops is derived from the satellite DNA, presumably by read-through transcription from promoters in the histone gene clusters. Unfortunately, we do not have comparable data on other specific genes, although there is considerable evidence for transcription of repeated sequences on LBCs of other amphibians (Macgregor and Andrews 1977; Varley et al. 1980a, 1980b) and birds (Solovei et al. 1996; Deryusheva et al. 2007; Gaginskaya et al. 2009).

On the basis of this admittedly incomplete evidence, it is reasonable to suppose that the long length of LBC loops relative to the lengths of “ordinary” genes results at least in part from read-through transcription into downstream noncoding regions. The disparity between loop size and the length of genes, already an issue for the relatively modest-sized LBC loops of *Xenopus*, becomes even more problematic for the gigantic loops of salamander LBCs (Figures 1.2 and 1.6). Many loops in these organisms are 25–50 µm in length and some reach the almost unbelievable length of 1 mm. Because 1 µm of B-form DNA corresponds to about 3 kb, many loops (and hence TUs) of salamander LBCs must be hundreds of kb long. There is already convincing evidence for very long introns in some salamander genes (Casimir et al. 1988; Smith et al. 2009). Detailed analysis of a few highly transcribed genes in salamander (and *Xenopus*) LBCs by *in situ* hybridization would add greatly to our understanding of LBC structure and function during oogenesis. It may well turn out that the majority of RNA transcribed on LBCs consists of either intronic or downstream noncoding regions.

**Appendix**

The majority of LBC loops are similar in general morphology within a given organism, as exemplified by the relatively short loops of anurans like *X. tropicalis* and the enormously longer loops of salamanders (Figure 1.2). As first shown in detail by Callan and Lloyd (1960) for the LBCs of the newt *Triturus*, it is possible to identify specific loops on the basis of their size and unique morphology of the RNP matrix. At present, we have almost no clue as to the functional significance of such differences among loops. It is possible to identify the transcripts being made on specific loops by correlating genetic maps and RNAseq data with fluorescent *in situ* hybridization analysis. To make such correlations easier, it is useful to have physical maps of the *X. laevis* LBCs, concentrating primarily on the distribution of the 5S and ribosomal RNA genes (Callan et al. 1988). In the interim, a good deal of additional mapping has been done, and updated maps are presented in Figure A1.1. More recently, *X. tropicalis* has become the favorite organism for sequence analysis, its major advantage being that it is a diploid species (*n* = 10), whereas *X. laevis* is an allotetraploid (*n* = 18). For that reason, it is useful to have LBC maps for this species as well. In Figure A1.2, we present our most current maps for *X. tropicalis*. Similar maps were recently published by Penrad-Mobayed et al. (2009). There are slight discrepancies in numbering between our maps and those of Penrad-Mobayed, resulting from the difficulty in determining relative lengths of the similarly sized chromosome. There are also discrepancies in numbering between both the LBC maps and the mitotic maps published earlier (Wells et al. 2011). These discrepancies will need to be resolved by *in situ* hybridization of specific sequences on the LBCs.

**Acknowledgments**

I thank the members of my laboratory for data and helpful discussions, especially Svetlana Deryusheva, Zehra Nizami, Jun Wei Pek, and Zheng'an Wu. Many of the ideas presented here were developed during conversations with Herbert Macgregor, Gary Morgan, and Michel Bellini. Special thanks go to Sidney Shaw and James Powers for help with super-resolution microscopy. The research reported in this publication was supported by the
Figure A1.1 Cytological maps of the 18 LBCs of *X. laevis*, based on the analysis of 41 complete or nearly complete spread preparations. Lengths are given as fraction of the total length of all chromosomes. The numbering system is the same as that given in Murphy et al. (2002), differing slightly from the original maps in Callan et al. (1987). Centromere positions (large solid circles) were determined from a subset of 15 preparations in which the oocytes had been injected with a myc-tagged transcript of the centromere-specific protein CENP-C, and centromeres detected with an antibody against the tag. Pol III sites are shown as elongated ovals at positions described earlier in Murphy et al. (2002). Three chromosomes (Nos. 8, 11, and 16) bear histone locus bodies (HLB) at the histone gene loci (Callan et al. 1991). The nucleolus organizer is located near the centromere of chromosome No. 3 (Callan et al. 1988), although a nucleolus is only rarely seen at this locus. Oocyte-specific 5S genes are located at or near the end of the long arm of all chromosomes except Nos. 10, 17, and 18 (Callan et al. 1988). These regions are recognizable by the presence of a small terminal granule (solid circle) and pol III-labeled loops. Bodies identical in morphology and immunostaining properties to extrachromosomal speckles (B-snurposomes) are regularly seen at specific chromosome termini and at a few interstitial sites (small open circles). A dlb near the nucleolus organizer of chromosome No. 3 is associated with the RNA-editing enzyme ADAR1 (Eckmann and Jantsch 1999). Double-axis regions of unknown significance occur near the centromeres of chromosome Nos. 8 and 9.
Figure A1.2  Cytological maps of the 10 LBCs of X. tropicalis, based on the analysis of 29 complete or nearly complete spread preparations. Lengths are given as fraction of the total length of all chromosomes. Centromere positions (large solid circles) were determined from a subset of 10 preparations in which the oocytes had been injected with a myc-tagged transcript of the centromere-specific protein CENP-C, and centromeres detected with an antibody against the tag. Terminal spheres of unknown nature are present on 15 of the 20 telomeres. These stain with an antibody against pol III, as do multiple internal sites (small solid circles). Four pol III sites on chromosome Nos 3, No. 4, and No. 10 frequently have pearls (P) associated with them (Nizami and Gall 2012). Three chromosomes (Nos. 3, 7, and 9) bear HLBs, presumably at the histone loci (not independently verified). The single nucleolus is located near the middle of chromosome No. 8, and the position of the nucleolus organizer (NO) has been verified by in situ hybridization. The large gray masses on several chromosomes are presumed to be “lumpy loops” as described originally by Callan in the newt Triturus (Callan and Lloyd 1960).

References


Macgregor H, Andrews C (1977) The arrangement and transcription of ‘middle repetitive’ DNA