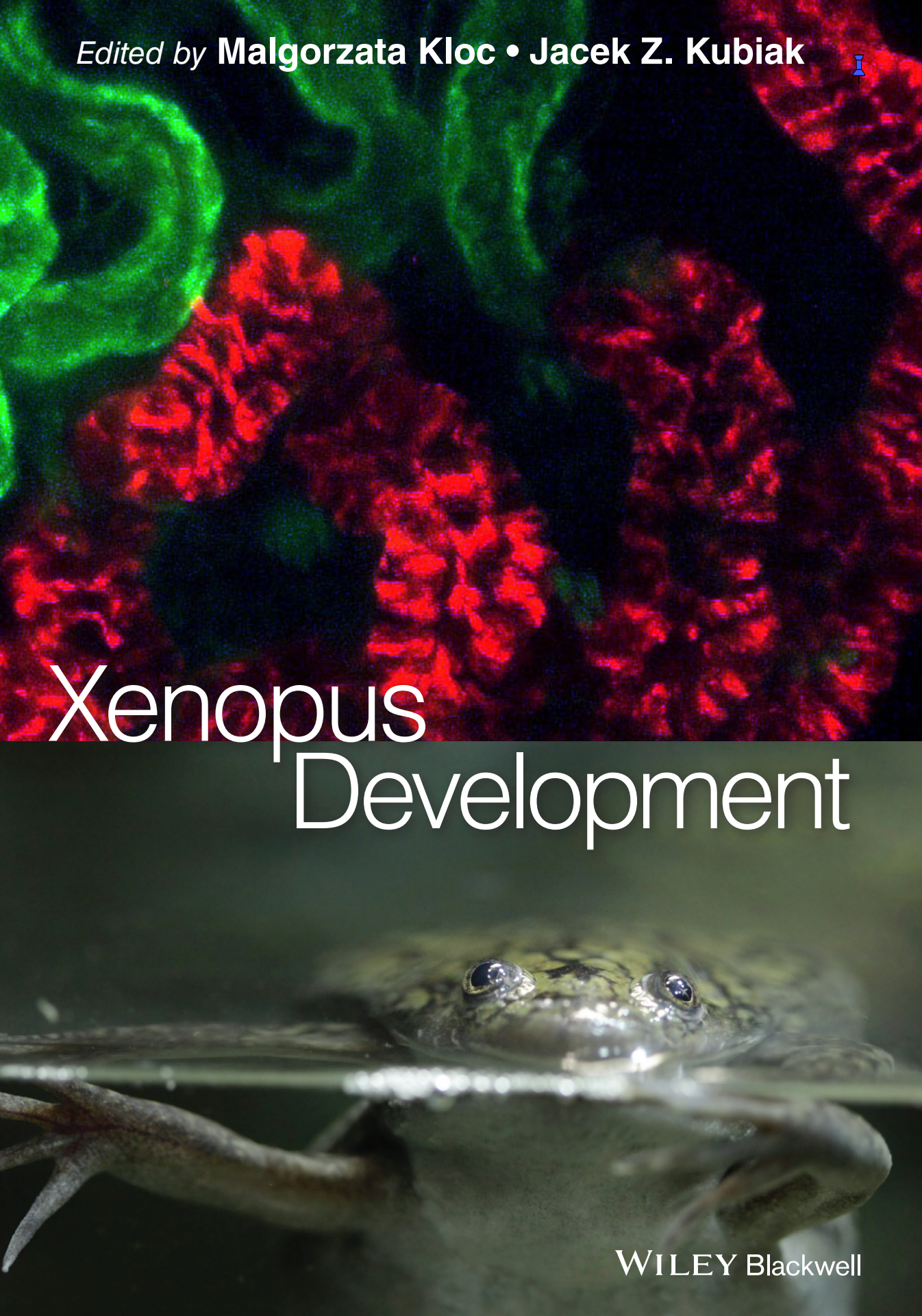

Edited by Malgorzata Kloc • Jacek Z. Kubiak
Xenopus Development
WILEY Blackwell

Xenopus Development

Xenopus Development

Edited by

Malgorzata Kloc

Department of Surgery
Houston Methodist Research Institute
and Houston Methodist Hospital
Houston, TX, USA

Jacek Z. Kubiak

CNRS and University of Rennes 1,
Institute of Genetics & Development of Rennes (IGDR),
UMR 6290, Faculty of Medicine,
Rennes, France

WILEY Blackwell

Copyright © 2014 by John Wiley & Sons, Inc. All rights reserved

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning, or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at http://www.wiley.com/go/permission.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representations or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose.

No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Xenopus development / edited by Malgorzata Kloc, Jacek Z. Kubiak.
pages cm
Includes bibliographical references and index.
ISBN 978-1-118-49281-9 (hardback)
1. Xenopus laevis. 2. Xenopus–Larvae–Microbiology. 3. Microorganisms–Development. 4. Embryology.
I. Kloc, Malgorzata, editor of compilation. II. Kubiak, Jacek Z, editor of compilation.
QL668.E265X46 2014
597.8′654–dc23

2014000055

Printed in Malaysia

10 9 8 7 6 5 4 3 2 1

Contents

Contributors

Masasuke Araki. Department of Biological Sciences, Developmental Neurobiology Laboratory, Nara Women's University, Nara, Japan.

Isabelle Chartrain. CNRS and University of Rennes 1, Institute of Genetics & Development of Rennes (IGDR), UMR 6290, Faculty of Medicine, Rennes, France.

Ying Chen. Stem Cell Institute, University of Minnesota, Minneapolis, MN.

Marie Cibois. Institute of Developmental Biology of Marseille, CNRS, Aix-Marseille Université, Marseille, France.

Chad E. Cragle. Department of Neurobiology and Developmental Sciences, University of Arkansas for Medical Sciences, Little Rock, AR.

James O. Deshler. Directorate for Biological Sciences, National Science Foundation, Arlington, VA.

Aliaksandr Dzementsei. Institute for Developmental Biochemistry, Göttingen, Germany.

Lisa J. Edens. Department of Molecular Biology, University of Wyoming, Laramie, WY.

François Fagotto. Department of Biology, McGill University, Montreal, Québec, Canada.

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

Mary Lou King. Department of Cell Biology, Miller School of Medicine, University of Miami, Miami, FL.

Malgorzata Kloc. Department of Surgery, Houston Methodist Research Institute and Houston Methodist Hospital, Houston, TX.

Laurent Kodjabachian. Institute of Developmental Biology of Marseille, CNRS, Aix-Marseille Université, Marseille, France.

Jacek Z. Kubiak. CNRS and University of Rennes 1, Institute of Genetics & Development of Rennes (IGDR), UMR 6290, Faculty of Medicine, Rennes, France.

Juan Larrain. Department of Cell and Molecular Biology, Centre for Aging and Regeneration, Millennium Nucleus in Regenerative Biology, Faculty of Biological Sciences, Pontificia Universidad Catolica de Chile, Santiago, Chile.

Moonsup Lee. Department of Biochemistry and Molecular Biology, The University of Texas MD Anderson Cancer Center. Program in Genes and Development, The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, TX.

Daniel L. Levy. Department of Molecular Biology, University of Wyoming, Laramie, WY.

Gufa Lin. Stem Cell Institute, University of Minnesota, Minneapolis, MN. School of Life

Science and Technology, Tongji University, Shanghai, China.

Angus M. MacNicol. Department of Neurobiology and Developmental Sciences, University of Arkansas for Medical Sciences, Little Rock, AR.

Pierre D. McCrea. Department of Biochemistry and Molecular Biology, The University of Texas MD Anderson Cancer Center. Program in Genes and Development, The University of Texas Graduate School of Biomedical Sciences at Houston, Houston, TX.

Rachel K. Miller. Department of Biochemistry and Molecular Biology, The University of Texas MD Anderson Cancer Center. Department of Pediatrics, University of Texas Medical School at Houston, Houston, TX.

Mauricio Moreno. Department of Cell and Molecular Biology, Centre for Aging and Regeneration, Millennium Nucleus in Regenerative Biology, Faculty of Biological Sciences, Pontificia Universidad Catolica de Chile, Santiago, Chile.

Hitoshi Morita. National Institute for Basic Biology, Okazaki, Japan. Institute of Science and Technology Austria, Klosterneuburg, Austria.

William A. Muñoz. Department of Biochemistry and Molecular Biology, The University of Texas MD Anderson Cancer Center, Houston, TX. Stowers Institute for Medical Research, Kansas City, MO.

Francesco Orsini. Physics Department, Università degli Studi di Milano, Milan, Italy.

Andrea Pasini. Institute of Developmental Biology of Marseille, CNRS, Aix-Marseille Université, Marseille, France.

Louis Du Pasquier. Zoology and Evolutionary Biology, University of Basel, Basel, Switzerland.

Tomas Pieler. Institute for Developmental Biochemistry, Göttingen, Germany.

Rafał P. Piprek. Department of Comparative Anatomy, Institute of Zoology, Jagiellonian University, Kraków, Poland.

Nicolas Pollet. CNRS, Institute of Systems and Synthetic Biology, Université d'Evry-Val-d'Essonne. Genopole, Centre d'Exploration et de Recherche Fonctionnelle Amphibiens Poissons, Evry, France.

Amy K. Sater. Department of Biology and Biochemistry, University of Houston, Houston, TX.

Thibault Scalvenzi. CNRS, Institute of Systems and Synthetic Biology, Université d'Evry-Val-d'Essonne, Evry, France.

Pierluigi Scerbo. Institute of Developmental Biology of Marseille, CNRS, Aix-Marseille Université, Marseille, France.

Makoto Suzuki. National Institute for Basic Biology, Okazaki, Japan.

Jean-Pierre Tassan. CNRS and University of Rennes 1, Institute of Genetics & Development of Rennes (IGDR), UMR 6290, Faculty of Medicine, Rennes, France.

Karina Tapia. Department of Cell and Molecular Biology, Centre for Aging and Regeneration, Millennium Nucleus in Regenerative Biology, Faculty of Biological Sciences, Pontificia Universidad Catolica de Chile, Santiago, Chile.

Aurore Thélie. CNRS, Institute of Systems and Synthetic Biology, Université d'Evry-Val-d'Essonne, Evry, France.

Virginie Thomé. Institute of Developmental Biology of Marseille, CNRS, Aix-Marseille Université, Marseille, France.

Naoto Ueno. National Institute for Basic Biology, Okazaki, Japan.

Léna Vouillot. CNRS, Institute of Systems and Synthetic Biology, Université d'Evry-Val-d'Essonne, Evry, France.

Preface

The purpose of writing *Xenopus Development* was to provide a comprehensive review of the current knowledge on the most popular amphibian model in developmental biology. The pioneering research by John Gurdon on nuclear transfer and nuclear remodeling in *Xenopus laevis* was awarded the Nobel Prize in Medicine or Physiology in 2012. This is perhaps the first time that research on the *Xenopus* model has been recognized with the highest scientific award. Recent sequencing of the *Xenopus tropicalis* genome allows combining the classical developmental biology observations and experiments carried out on *X. laevis* with modern genetic and genomic studies of *X. tropicalis*. This is a unique situation in modern developmental biology, with two different but closely related species being used for different purposes and being studied using different approaches, thereby allowing the results to be automatically merged and easily extrapolated. Availability of these data sets will have an enormous impact on the general application of the *Xenopus* model system. At present, there are two *Xenopus* resource centers, one in the US and one in the UK, which offer training in the use of *Xenopus* as an experimental model system. Both the *X. laevis* and *X. tropicalis* models have the potential to be used more frequently in the future and will certainly deliver novel and exciting information in the field of developmental biology.

The book is divided into four parts: Section I – Oocyte and Early Embryo (Chapters 1–5); Section II – Midblastula Transition, Gastrulation, and Neurulation (Chapters 6–9); Section III – Metamorphosis and Organogenesis (Chapters 10–15); and Section IV – Novel Techniques and Approaches (Chapters 16–20). This arrangement allows presenting the novel discoveries in the field of *Xenopus* developmental biology in a systematic manner and focusing on the methodological aspects of *Xenopus* research. We are now witnessing an explosive development of novel methods, approaches, and techniques, which pave the way to explore new areas of research for scientific discoveries. Researchers in the field can benefit from these circumstances and make use of this unique opportunity.

Most importantly, we have managed to gather in this book outstanding contributors who have provided an excellent historical perspective as well as described the state of the art in the field of their expertise.

Last, but not least, there has not been a book dedicated to *Xenopus* since the 2000 Cold Spring Harbor Lab Press laboratory manual, and we hope that the current volume will fill this void successfully.

Malgorzata Kloc
Houston, USA
Jacek Z. Kubiak
Rennes, France

Section I

Oocyte and Early Embryo

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

Xenopus Development, First Edition. Edited by Malgorzata Kloc and Jacek Z. Kubiak.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

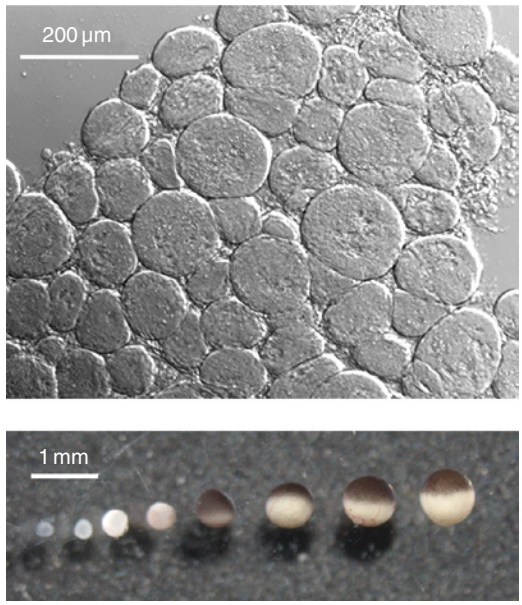


Figure 1.1 Oocytes of *X. tropicalis*. The top panel shows the range of oocyte sizes found in an ovary from an immature frog (3.5 cm snout to vent). At this stage, most oocytes have diameters under 100 µm. The lower panel shows oocytes of different sizes, obtained from the ovary of a mature female. Such ovaries also contain smaller oocytes like those shown in the upper panel. Photo courtesy of Zehra Nizami.

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 Macgregor 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

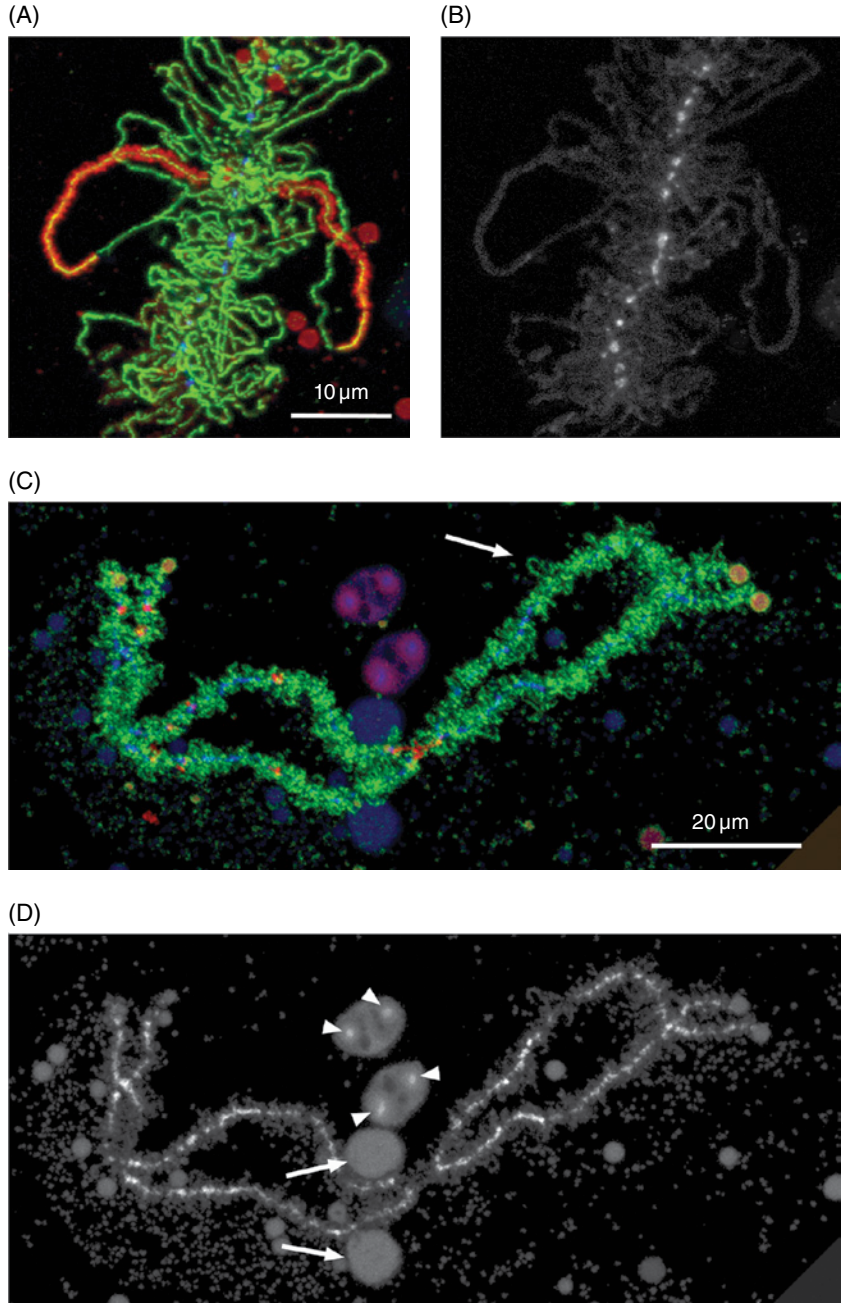


Figure 1.2 LBCs of the newt *Notophthalmus viridescens* (A and B) and *X. tropicalis* (C and D). (A) A short segment of an LBC stained with antibodies against pol II (green) and the RNA-binding protein CELF1 (red) (Morgan 2007). The axes of all loops appear as diffraction-limited green lines, because 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.

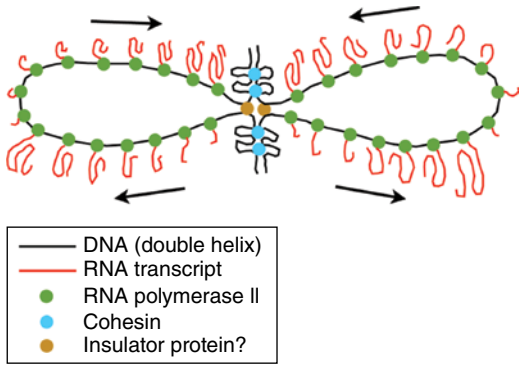


Figure 1.3 Highly stylized diagram of LBC structure. Transcriptionally active sister chromatids extend laterally from the main axis of the chromosome, which consists of regions where transcriptionally inactive sisters are closely paired and associated with cohesins (Austin et al. 2009). Loops can consist of one or more TUs, which may have either the same or opposite polarities on the same loop. RNA polymerase II molecules are packed closely along the DNA axis of each loop and elongating RNA transcripts are attached to them. The transcripts are associated with various proteins, including splicing factors (not shown here). It is not known what holds the bases of the loops together. One possibility is that insulators or similar molecules that define transcriptionally active regions of chromatin are involved. To see a color version of this figure; see Plate 2.

visible at the light optical level. By electron microscopy, these regions usually appear as a **single** fiber about 10nm 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

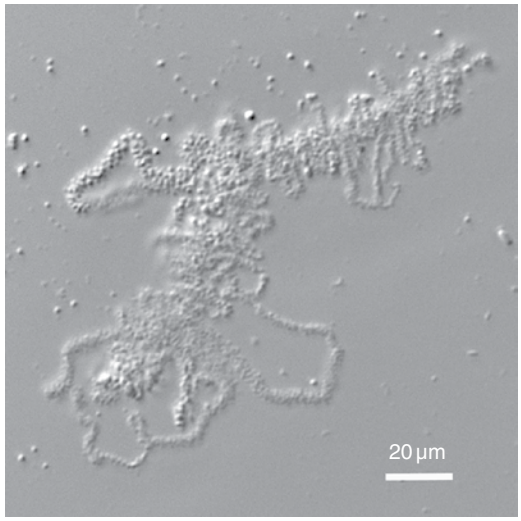


Figure 1.4 An LBC consisting of a single unpaired chromatid. This LBC was formed when a sperm head of *X. laevis* was injected into the GV of the newt *N. viridescens*. Individual chromatids derived from the sperm begin transcribing shortly after injection, eventually forming giant chromosomes similar to the endogenous LBCs. Because the *X. laevis* chromatids do not replicate in the GV, the LBCs formed from them consist of single chromatids and the transcription loops are unpaired.

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

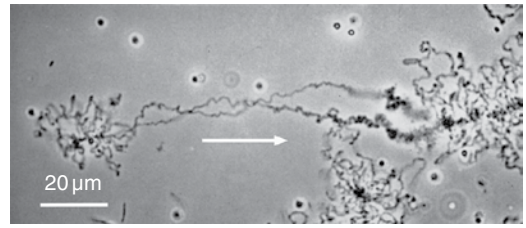


Figure 1.5 A dlb in a chromosome of the newt *N. viridescens*. Such bridges can be formed by stretching a chromosome with microneedles, but they also occur by accident when LBCs are prepared for microscopical examination. Note the polarity of the loops, which allows one to determine the direction of transcription (arrow) relative to the chromosome as a whole.

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).

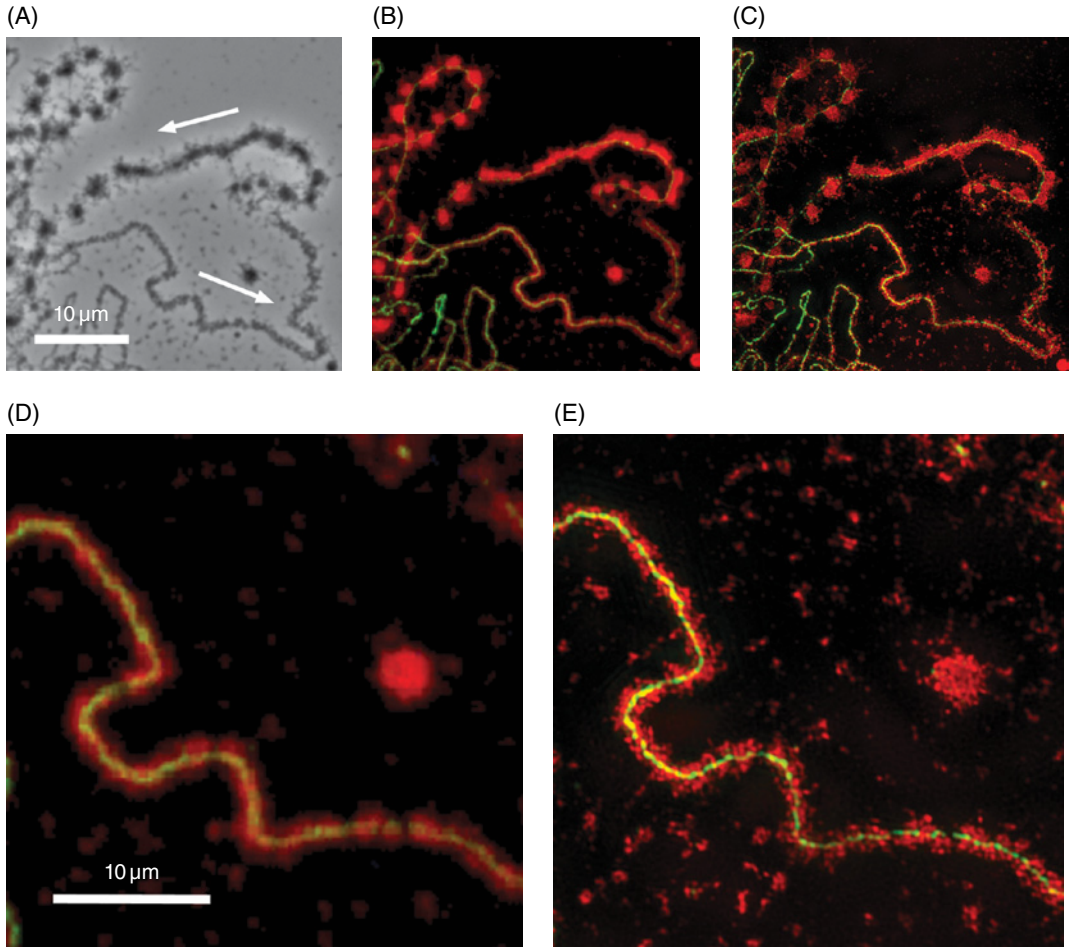


Figure 1.6 Images of a loop from the newt *N. viridescens*. (A) The entire loop imaged by phase contrast microscopy. The pronounced thin-to-thick polarity of the RNP matrix signifies the direction of transcription (arrows). (B) A confocal image of the same loop after immunostaining with mAb H14 against phosphorylated pol II (green) and mAb Y12 against symmetrical dimethylarginine, an epitope found on several splicing snRNPs (red). Green pol II stain is evident at the thin end of the loop but is obscured by the heavy mAb Y12 stain along most of the loop. (C) Image of the same loop taken by structured illumination superresolution microscopy. (D) Confocal image of the thin end of the loop at higher magnification. (E) The same loop imaged by structured illumination microscopy. Pol II now appears as a green line of nearly uniform width along the length of the loop. The red RNP matrix is resolved into a series of small particles about 50 nm in diameter. The superresolution images were taken on a DeltaVision OMX structured illumination microscope by Sidney Shaw and James Powers, Department of Biology, Indiana University. To see a color version of this figure, see Plate 3.

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 10^7 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,

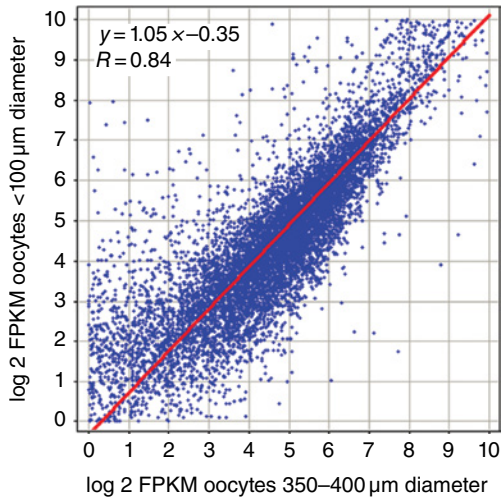


Figure 1.7 Similarity of transcriptomes from *X. tropicalis* oocytes less than 100 µm diameter and oocytes that have reached 350–400 µm diameter, approximately half their final size. Shown here are the log 2 FPKM scores for approximately 9700 different transcripts. The slope of approximately 1.0 and the high correlation (R = 0.84) show that transcripts are stored at similar relative concentrations from the earliest to midstages of oogenesis. Transcripts from fully mature oocytes are similar (not shown here).

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 LBCs. Some years ago, we published relatively crude 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

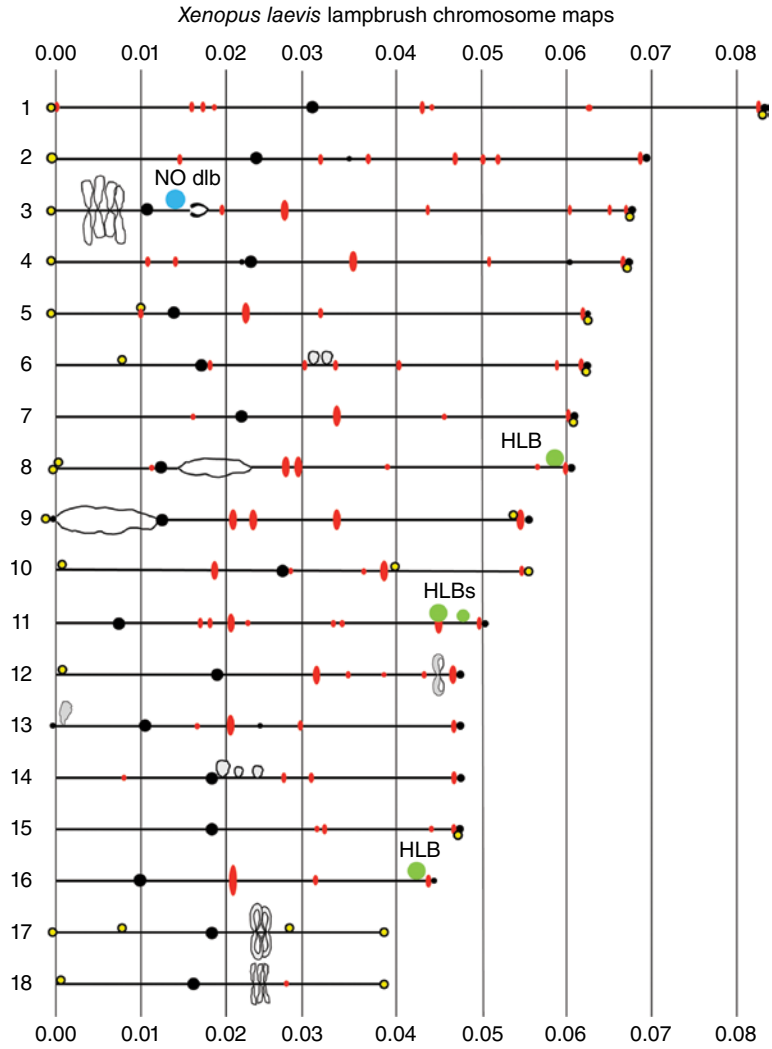


Figure A1.1 Cytological maps of the 18 LBCs of *X. laevis*, based on the analysis of 41 complete or nearly complete spread preparations. Lengths are given as fraction of the total length of all chromosomes. The numbering system is the same as that given in Murphy et al. (2002), differing slightly from 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.

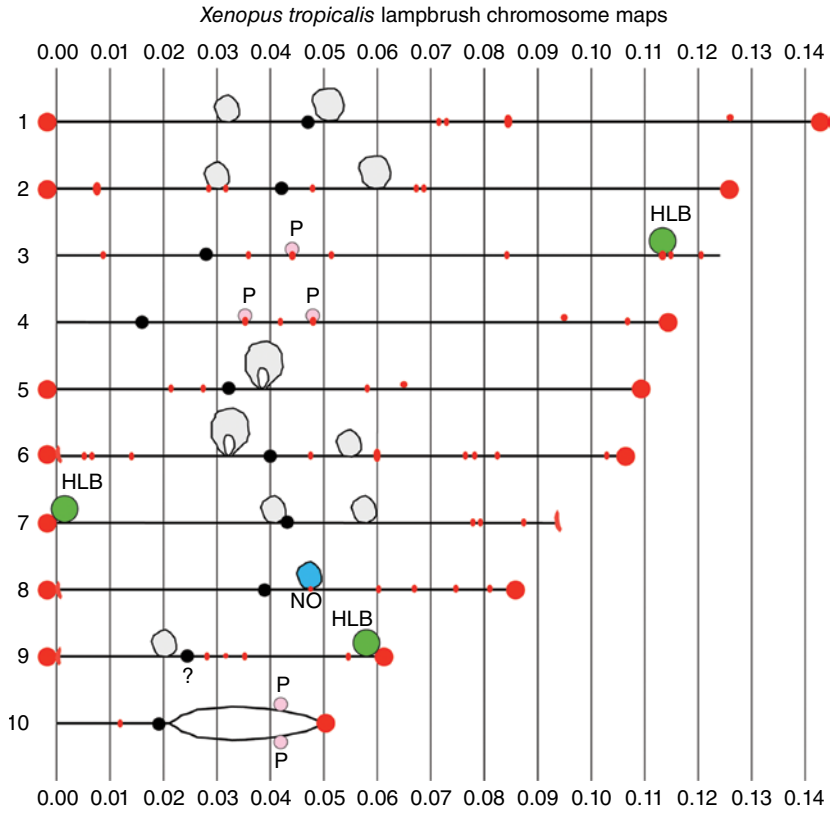


Figure A1.2 Cytological maps of the 10 LBCs of *X. tropicalis*, based on the analysis of 29 complete or nearly complete spread preparations. Lengths are given as fraction of the total length of all chromosomes. Centromere positions (large solid circles) were determined from a subset of 10 preparations 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).

National Institute of General Medical Sciences of the National Institutes of Health under award number R01 GM33397. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health. JGG is an American Cancer Society Professor of Developmental Genetics.

References

Austin C, Novikova N, Guacci V, Bellini M (2009) Lampbrush chromosomes enable study of cohesin dynamics. Chromosome Res 17:165–184.

Berger S, Menzel D, Traub P (1994) Chromosomal architecture in giant premeiotic nuclei of the green alga *Acetabularia*. Protoplasma 178:119–128.

Brown DD, Littna E (1964) RNA synthesis during the development of *Xenopus laevis*, the South African clawed toad. J Mol Biol 8:669–687.

Brown DD, Dawid IB (1968) Specific gene amplification in oocytes. Science 160:272–280.

Brown DD, Wensink PC, Jordan E (1971) Purification and some characteristics of 5S DNA from *Xenopus laevis*. Proc Natl Acad Sci U S A 68:3175–3179.

Callan HG (1954) Recent work on the structure of cell nuclei. in **Fine Structure of Cells: Symposium of the VIIIth Congress in Cell Biology**, Leiden 1954, International Union of Biological Sciences Publ, Series B, pp. 89–109. Noordhoff, Groningen.

Callan HG (1957) The lampbrush chromosomes of *Sepia officinalis* L., *Anilocra physodes* L. and *Scyllium catulus* Cuv. and their structural relationship to the lampbrush chromosomes of amphibia. Pubbl Staz Zool Napoli 29:329–346.

Callan HG (1963) The nature of lampbrush chromosomes. Int Rev Cytol 15:1–34.

Callan HG (1986) Lampbrush Chromosomes. Springer-Verlag, Berlin.

Callan HG, Macgregor HC (1958) Action of deoxyribonuclease on lampbrush chromosomes. Nature (Lond) 181:1479–1480.

Callan HG, Lloyd L (1960) Lampbrush chromosomes of crested newts *Triturus cristatus* (Laurenti). Philos Trans R Soc Lond B Biol Sci 243:135–219.

Callan HG, Gall JG, Berg CA (1987) The lampbrush chromosomes of *Xenopus laevis*: preparation, identification, and distribution of 5S DNA sequences. Chromosoma (Berlin) 95:236–250.

Callan HG, Gall JG, Murphy C (1988) The distribution of oocyte 5S, somatic 5S and 18S + 28S rDNA sequences in the lampbrush chromosomes of *Xenopus laevis*. Chromosoma (Berlin) 97:43–54.

Callan HG, Gall JG, Murphy C (1991) Histone genes are located at the sphere loci of *Xenopus* lampbrush chromosomes. Chromosoma (Berlin) 101: 245–251.

Casimir CM, Gates PB, Ross-Macdonald PB, Jackson JF, Patient RK, Brockes JP (1988) Structure and expression of a newt cardio-skeletal myosin gene. Implications for the C value paradox. J Mol Biol 202:287–296.

Davidson EH (1986) Gene Activity in Early Development. Academic Press, Orlando.

Deryusheva S, Krasikova A, Kulikova T, Gaginskaya E (2007) Tandem 41-bp repeats in chicken and Japanese quail genomes: FISH mapping and transcription analysis on lampbrush chromosomes. Chromosoma 116:519–530.

Diaz MO, Gall JG (1985) Giant readthrough transcription units at the histone loci on lampbrush chromosomes of the newt *Notophthalmus*. Chromosoma 92:243–253.

Diaz MO, Barsacchi-Pilone G, Mahon KA, Gall JG (1981) Transcripts from both strands of a satellite DNA occur on lampbrush chromosome loops of the newt *Notophthalmus*. Cell 24:649–659.

Eckmann CR, Jantsch MF (1999) The RNA-editing enzyme ADAR1 is localized to the nascent ribonucleoprotein matrix on *Xenopus* lampbrush chromosomes but specifically associates with an atypical loop. J Cell Biol 144:603–615.

Gaginskaya E, Kulikova T, Krasikova A (2009) Avian lampbrush chromosomes: a powerful tool for exploration of genome expression. Cytogen Gen Res 124:251–267.

Gall JG (1954) Lampbrush chromosomes from oocyte nuclei of the newt. J Morphol 94:283–352.

Gall JG (1956) On the submicroscopic structure of chromosomes. Brookhaven Symp Biol 8:17–32.

Gall JG (1958) Chromosomal differentiation. in A Symposium on the Chemical Basis of Development (eds. WD McElroy, B Glass), pp. 103–135. Johns Hopkins Press, Baltimore.

Gall JG (1963) Kinetics of deoxyribonuclease action on chromosomes. Nature (Lond) 198:36–38.

Gall JG (1968) Differential synthesis of the genes for ribosomal RNA during amphibian oogenesis. Proc Natl Acad Sci U S A 60:553–560.

Gall JG (1981) Chromosome structure and the C-value paradox. J Cell Biol 91:3s–14s.

Gall JG (2012) Are lampbrush chromosomes unique to meiotic cells? Chromosome Res 20:905–909.

Gall JG, Callan HG (1962) H^3 uridine incorporation in lampbrush chromosomes. Proc Natl Acad Sci U S A 48:562–570.

Gall JG, Murphy C (1998) Assembly of lampbrush chromosomes from sperm chromatin. Mol Biol Cell 9:733–747.

Gall JG, Diaz MO, Stephenson EC, Mahon KA (1983) The transcription unit of lampbrush chromosomes. in Gene Structure and Regulation in Development (eds. S Subtelny, F Kafatos), pp. 137–146. Alan R. Liss, New York.

Gardner EJ, Nizami ZF, Talbot CC, Jr., Gall JG (2012) Stable intronic sequence RNA (sisRNA), a new class of noncoding RNA from the oocyte nucleus of *Xenopus tropicalis*. Genes Dev 26:2550–2559.

Giles KE, Gowher H, Ghirlando R, Jin C, Felsenfeld G (2010) Chromatin boundaries, insulators, and long-range interactions in the nucleus. Cold Spring Harb Symp Quant Biol 75:79–85.

Gregory TR (2006) Animal Genome Size Database. http://www.genomesize.com (accessed on November 8, 2013).

Hamkalo BA, Miller OL, Jr. (1973) Electron-microscopy of genetic activity. Annu Rev Biochem 42:379–396.

Hellsten U, Harland RM, Gilchrist MJ, et al. (2010) The genome of the western clawed frog *Xenopus tropicalis*. Science 328:633–636.

Hess O (1971) Lampenbürstenchromosomen. in Handbuch der allgemeinen Pathologie (eds. H-W Altmann, F Büchner, H Cottier, et al.), pp. 215–281. Springer-Verlag, Berlin.

Liu JL, Gall JG (2012) Induction of human lampbrush chromosomes. Chromosome Res 20: 971–978.

Macgregor HC (2012) Chromomeres revisited. Chromosome Res 20:911–924.

Macgregor HC, Andrews C (1977) The arrangement and transcription of 'middle repetitive' DNA

sequences on lampbrush chromosomes of *Triturus*. Chromosoma 63:109–126.

Miller OL, Jr. (1966) Structure and composition of peripheral nucleoli of salamander oocytes. J Natl Cancer Inst Monogr 23:53–66.

Miller OL, Jr., Hamkalo BA (1972) Visualization of RNA synthesis on chromosomes. Int Rev Cytol 33:1–25.

Morgan GT (2002) Lampbrush chromosomes and associated bodies: new insights into principles of nuclear structure and function. Chromosome Res 10:177–200.

Morgan GT (2007) Localized co-transcriptional recruitment of the multifunctional RNA-binding protein CELF1 by lampbrush chromosome transcription units. Chromosome Res 15:985–1000.

Mott MR, Callan HG (1975) An electron-microscope study of the lampbrush chromosomes of the newt *Triturus cristatus*. J Cell Sci 17:241–261.

Murphy C, Wang Z, Roeder RG, Gall JG (2002) RNA polymerase III in Cajal bodies and lampbrush chromosomes of the *Xenopus* oocyte nucleus. Mol Biol Cell 13:3466–3476.

Nizami ZF, Gall JG (2012) Pearls are novel Cajal body-like structures in the *Xenopus* germinal vesicle that are dependent on RNA pol III transcription. Chromosome Res 20:953–969.

Peacock WJ (1965) Chromosome replication. J Natl Cancer Inst Monogr 18:101–131.

Penrad-Mobayed M, El Jamil A, Kanhoush R, Perrin C (2009) Working map of the lampbrush chromosomes of *Xenopus tropicalis*: a new tool for cytogenetic analyses. Dev Dyn 238:1492–1501.

Perkowska E, Macgregor HC, Birnstiel ML (1968) Gene amplification in the oocyte nucleus of mutant and wild-type *Xenopus laevis*. Nature (Lond) 217:649–650.

Rückert J (1892). Zur Entwickelungsgeschichte des Ovarialeies bei Selachiern. Anat Anz 7:107–158.

Scheer U (1981) Identification of a novel class of tandemly repeated genes transcribed on lampbrush chromosomes of *Pleurodeles waltlii*. J Cell Biol 88:599–603.

Scheer U, Franke WW, Trendelenburg MF, Spring H (1976) Classification of loops of lampbrush chromosomes according to the arrangement of transcriptional complexes. J Cell Sci 22:503–519.

Schmid M, Nanda I, Hoehn H, et al. (2005) Second report on chicken genes and chromosomes 2005. Cytogenet Genome Res 109:415–479.

Simeoni I, Gilchrist MJ, Garrett N, Armisen J, Gurdon JB (2012) Widespread transcription in an amphibian oocyte relates to its reprogramming activity on transplanted somatic nuclei. Stem Cells Dev 21:181–190.

Smith JJ, Putta S, Zhu W, et al. (2009) Genic regions of a large salamander genome contain long introns and novel genes. BMC Genomics 10:19.

Solovei IV, Joffe BI, Gaginskaya ER, Macgregor HC (1996) Transcription on lampbrush chromosomes of a centromerically localized highly repeated DNA in pigeon (*Columba*) relates to sequence arrangement. Chromosome Res 4:588–603.

Spradling AC (1993) Developmental genetics of oogenesis. in The Development of *Drosophila melanogaster* (eds. M Bate, A Martinez-Arias), pp. 1–70. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Spring H, Scheer U, Franke WW, Trendelenburg MF (1975) Lampbrush-type chromosomes in the primary nucleus of the green alga *Acetabularia mediterranea*. Chromosoma 50:25–43.

Stephenson EC, Erba HP, Gall JG (1981) Histone gene clusters of the newt *Notophthalmus* are separated by long tracts of satellite DNA. Cell 24: 639–647.

Tomlin SG, Callan HG (1951) Preliminary account of an electron microscope study of chromosomes from newt oocytes. Q J Microsc Sci 92:221–224.

Varley JM, Macgregor HC, Erba HP (1980a) Satellite DNA is transcribed on lampbrush chromosomes. Nature (Lond) 283:686–688.

Varley JM, Macgregor HC, Nardi I, Andrews C, Erba HP (1980b) Cytological evidence of transcription of highly repeated DNA sequences during the lampbrush stage in *Triturus cristatus carnifex*. Chromosoma 80:289–307.

Voronina E, Wessell GM (2003) The regulation of oocyte maturation. Curr Top Dev Biol 58:53–110.

Wallace H, Birnstiel ML (1966) Ribosomal cistrons and the nucleolar organizer. Biochim Biophys Acta 114:296–310.

Wells DE, Gutierrez L, Xu Z, et al. (2011) A genetic map of *Xenopus tropicalis*. Dev Biol 354:1–8.

2 RNA Localization during Oogenesis in *Xenopus laevis*

James O. Deshler
Directorate for Biological Sciences, National Science Foundation, Arlington, VA

Abstract: The polarized distribution of mRNA is a wide-spread mechanism for regulating cell differentiation and cell function. *Xenopus* oocytes have served as a wonderful model system to investigate the mechanism(s) underlying this process. Here, a summary of major findings in the *Xenopus* oocyte system is presented, and these findings are compared with findings in other species and cell types. A model is presented that suggests RNA localization elements form secondary structural elements comprised of distinct RNA strands from two or more localizing mRNA molecules. In this model, these intermolecular RNA structures play a role in recruiting critical proteins required for the localization process. Since this mechanism is likely to regulate the spatial expression patterns of thousands of proteins encoded in a single genome, future work should focus on advanced algorithm development to identify these and other types of nonprotein-coding RNA regulatory elements that play a major role in establishing diverse phenotypes from specific genotypes.

Xenopus oocytes as a model system for exploring RNA localization

The generation of polarized distributions of specific RNAs, proteins, and subcellular organelles is a key step toward organizing a cell. This spatial and temporal aspect of regulation contributes significantly to cell type-specific functions in all organisms. The specific localization of distinct mRNAs is a widespread mechanism for generating polarity in both somatic and germ cells and has been studied extensively in highly polarized cells, such as oocytes, neurons, and epithelial cells where the process of establishing mRNA polarity is most amenable to experimental investigation. The primary role for mRNA localization is to establish localized protein synthesis from distinct mRNAs at particular subcellular locations where proteins are required for specific cellular functions and exclude them from regions where they are not needed or may be deleterious. One example of this is the localized synthesis of proteins at neuronal synapses which can be hundreds of microns away from the nucleus in the cell bodies where mRNAs are synthesized. The local synthesis of distinct proteins at synapses

Xenopus Development, First Edition. Edited by Malgorzata Kloc and Jacek Z. Kubiak.
© 2014 John Wiley & Sons, Inc. Published 2014 by John Wiley & Sons, Inc.

is thought to play a critical role in synaptic plasticity, long-term memory, as well as neurological disorders (Richter and Klann 2009; Liu-Yesucevitz et al. 2011).

Egg cells, like neuronal cells, also have a high degree of polarity and organization that is required to support the formation of an embryo as soon as fertilization occurs. These female germ cells of *Drosophila melanogaster* and *Xenopus laevis* have both been utilized extensively to investigate the mechanisms of RNA localization and the establishment of cell polarity because they are amenable to distinct types of experimental investigation. In both species, as in most animals, primordial germ cells are set aside early during embryogenesis as a source of stem cells that will differentiate into eggs or sperm in females or males, respectively. As primordial germ cells differentiate into oogonia and then oocytes in the ovary, they initiate meiosis, but arrest their cell cycle in the first meiotic prophase at which time they begin the process of oogenesis to form an egg. During oogenesis, these meiotic cells have the maximum copy number of each gene, and segments of genome that encode the ribosomal RNA genes are amplified to accommodate the high demand for protein synthesis in the growing oocytes. In *Xenopus*, this process takes 9–12 months but is on the order of just a few days in *Drosophila*. For a comparative description of this biological process in vertebrate and invertebrate animal models, including *Drosophila* and *Xenopus*, the reader is encouraged to read a review by Saffman and Lasko (1999). During oogenesis, oocytes accumulate yolk protein from the mother, but also generate highly organized patterns of mRNA localization and consequent protein expression. Sometimes the resulting polarized pattern of protein expression is visible to the naked eye. For example, fully grown *Xenopus* oocytes are over 1 mm in diameter and have pigment granules in the cortex of their animal hemisphere, making one half of the oocyte quite dark in appearance. Cells that acquired these pigment granules during early development migrate around the embryo, surrounding it completely later in development. The opposite hemisphere is referred to as the vegetal hemisphere. It has no pigment and appears light in color.

While arrested in the prophase of meiosis I, *Xenopus* oocytes progress through six characterized stages of growth, and a mixture of stage I–VI oocytes is present in the adult female ovary. Stage I oocytes are transparent and are 50–100 µm in diameter. As oocytes grow and accumulate yolk protein, they become opaque during stage II of oogenesis (100–450 µm diameter). Pigment granules form at the surface of the animal side of oocytes during the later stage III of oogenesis (450–600 µm diameter) and continue to increase in the animal hemisphere until the final stage VI of oogenesis (1200–1300 µm diameter) (Dumont 1972). Many RNAs have been discovered that localize to the vegetal pole and vegetal cortex of *Xenopus* oocytes. This process occurs primarily during stages I–III of oogenesis. Those RNAs that begin to localize in stage I oocytes, such as Xcat-2, first accumulate at a structure called the Balbiani body or mitochondrial cloud which is a large structure adjacent to one side of the nucleus and thus first defines the animal–vegetal axis of the growing oocyte (Figure 2.1). Some RNAs, such as Xcat-2, are targeted with somewhat more specificity to the germ plasm within the mitochondrial cloud, causing these RNAs to be segregated to primordial germ cells during early development (Kloc et al. 2000). The mitochondrial cloud, along with the associated early-pathway RNAs, migrates from its region near the nucleus of stage I oocytes to the vegetal cortex during stage II of oogenesis and remains at the vegetal pole through stage VI. RNAs that localize to the vegetal pole during the so-called "late pathway", such as Vg1, are distributed throughout the cytoplasm of stage I oocytes and begin their localization during stage II at which point they localize to a wedge-shaped structure just behind the early-pathway RNAs at the vegetal pole (Figure 2.1). These RNAs continue to localize to the vegetal cortex during stages III and IV of oogenesis. By stage IV of oogenesis, most of the Vg1 is localized throughout the vegetal cortex, whereas early-pathway RNAs remain in the cortex at the vegetal pole. The two best-characterized late-pathway RNAs, Vg1 and VegT, encode proteins that act synergistically (Agius et al. 2000) to specify the mesoderm during early embryogenesis

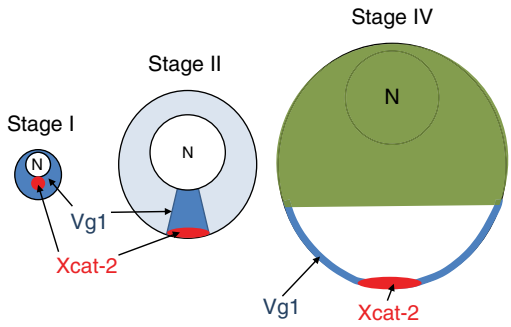


Figure 2.1 Distribution of early- and late-pathway RNAs in stage I–IV oocytes. On the left is a stage I oocyte showing the nucleus (N), the Vg1 mRNA distributed throughout the cytoplasm (blue), and the Xcat-2 localized to the Balbiani body or mitochondrial cloud adjacent to the nucleus (red). By stage II, the mitochondrial cloud and early-pathway RNAs have moved to the vegetal cortex, whereas late-pathway RNAs, such as Vg1 (blue), begin to localize to a wedge-shaped structure between the nucleus (N) and the early-pathway RNAs at the vegetal pole. A stage IV oocyte is shown on the right with a pigmented animal hemisphere at the top and Vg1 (blue) distributed through most of the vegetal cortex. Xcat-2 (red) and other early RNAs remain in the vegetal cortex but mostly at the vegetal pole. The oocytes are drawn to relative scale with the stage I oocyte being approximately 100 μM in diameter. The process of growing from a stage I to stage IV oocyte takes months in an adult female. For a comprehensive book of protocols and high-quality photos of different-staged oocytes, the reader is referred to volume 36 of *Methods in Cell Biology* (O'Keefe et al. 1991). To see a color version of this figure, see Plate 4.

(Kessler and Melton 1995; Joseph and Melton 1998; Zhang et al. 1998). Stage II oocytes are probably the best for studying the localization process because only at this stage will the early- and late-pathway injected exogenous RNAs adopt their relative localization patterns that most closely mimic their endogenous counterparts (Kloc et al. 1996) with only 18–36 h of culturing post injection. The molecular mechanism underlying this process of sorting and localizing mRNAs to the vegetal cortex will be the focus of this chapter.

In order to explore the mechanisms that mediate mRNA localization, it is important to consider the advantages and disadvantages of the various model systems employed to study the process. For example, a plethora of genetic manipulations are available in the *Drosophila* system and have been used successfully to identify and characterize proteins required for mRNA localization and transport in *Drosophila* oocytes and embryos. Through an elegant application of molecular, genetic, and developmental approaches available only in *Drosophila*, it has been shown that ectopic mislocalization of a single posterior mRNA, nanos, to the anterior end of an oocyte is sufficient to generate an entire posterior body structure resulting in a bipolar embryo (Gavis and Lehmann 1992). This fascinating result demonstrates that the polarized distribution of just a single upstream factor can be sufficient to establish all downstream patterning of a developing embryo, at least in this system. Insights into both the importance and mechanism of mRNA localization gained from the *Drosophila* system have been enormous and are summarized in recent review articles (Becalska and Gavis 2009; Lasko 2011). One potential limitation of the *Drosophila* system, however, is that from an evolutionary perspective, the patterning observed in developing *Drosophila* embryos is highly derived, such that specific orthologous or homologous mRNA localization pathways in distantly related animals have not yet been identified in oocytes and may not exist, even though many of the core RNA binding proteins and molecular motors are shared between species. This is one reason investigators have studied mRNA localization in other models, such as *Xenopus* oocytes, where genetic manipulations are not possible, but in which other types of experimental approaches are available and have revealed key insights into the mRNA localization process of vertebrates. Important advantages of the *Xenopus* oocyte model system include the ability to prepare cellular extracts from individually staged oocytes, to prepare undiluted cytoplasmic extracts that maintain associations that are sensitive to dilution, to microinject known quantities of labeled and unlabeled RNAs for *in vivo* competition experiments, to perform live imaging of RNAs being localized, and to immunoprecipitate proteins and/or RNAs presumably associated with RNA localization complexes.

Previous reviews have described numerous mRNAs that become localized to the vegetal pole during stages I–IV of oogenesis in *Xenopus* (King et al. 2005; Kloc and Etkin

2005). In this review, I will focus on distinguishing features of and recent findings about the mRNA localization process that directs RNAs toward the vegetal pole of growing *Xenopus* oocytes. In addition, questions for future research in this system will be addressed with the expectation that further exploration into these areas will help to inform studies of the mRNA localization process in *Xenopus* as well as other species across the phylogenetic tree. A few RNAs have also been discovered that localize to the animal pole and appear to interact with some of the vegetal pathway localization factors (Snedden et al. 2013). However, little else is known about the mechanism of their localization, and they will not be discussed further in this chapter.

Cis-elements and the role of short repeated motifs

The first mRNA localization element (LE) to be mapped in *Xenopus* is located in a 340-nucleotide (nt) fragment of the approximately 1200-nt 3′-untranslated region (UTR) of the Vg1 mRNA (Mowry and Melton 1992). This fragment is both necessary and sufficient to localize to the vegetal pole when injected into stage III/IV oocytes and cultured for 2–3 days. This has turned out to be a trend in that mRNA LEs reside in the 3′-UTR of most localized mRNAs throughout various species. Subsequent characterization of the Vg1 LE showed that there were short five- to nine-nt interspersed perfect repeat sequences that seemed to be more important for localization (Deshler et al. 1997) than other regions of the Vg1 LE when subjected to a comprehensive deletion analysis (Gautreau et al. 1997). The biggest surprise resulting from these studies was that the deletion of the smallest repeat, UUCAC, repeated five times in the Vg1 LE, led to the biggest reduction in localization when compared to other repeated sequences that are longer or larger deletions of the LE that don't contain repeated motifs (Gautreau et al. 1997). Since the UUCAC motif and other short sequences are required for localization and serve as binding sites for proteins that were identified by their ability to bind specifically to RNA LEs, these RNA binding proteins were also thought to be involved in the RNA localization process (Deshler et al. 1997, 1998). However, tandem arrays of these individual motifs fail to localize in isolation when injected into *Xenopus* oocytes (Deshler et al. 1998; Lewis et al. 2004), so it was postulated that combinations of motifs interact with a set of RNA binding proteins to form a ribonucleoprotein (RNP) complex that is competent to localize (Bubunenko et al. 2002; Lewis et al. 2004). This idea emerged through studies of the Vg1 and VegT mRNAs, both of which localize later in oogenesis than the so-called early-pathway RNAs, such as Xcat-2 or Xlsirt (Kloc and Etkin 1995).

The idea that a combination of distinct short RNA interspersed repeated sequence motifs interacts with their cognate RNA binding proteins to form a localization-competent RNP complex is a reasonable explanation for the role of these short motifs. However, this view became more complicated when a few early-pathway mRNAs were examined in detail. Xcat-2 is one of the best-characterized mRNAs that localizes to the mitochondrial cloud of early stage I oocytes before reaching the vegetal pole (Figure 2.1). Several groups have shown that when injected into later-stage oocytes, Xcat-2 is perfectly capable of localizing directly to the vegetal pole during the Vg1 or "late pathway" (Zhou and King 1996; Hudson and Woodland 1998; Allen et al. 2003), and *in vivo* competition assays show that the Xcat-2 LE competes for Vg1 localization factors more efficiently than the Vg1 LE does itself (Choo et al. 2005). Moreover, labeled Xcat-2 LE localizes much faster during later stages of oogenesis than the Vg1 LE (Choo et al. 2005), as does the Xlsirt early-pathway RNA when coinjected simultaneously with Vg1 into stage II oocytes (Kloc et al. 1996). These and other data, such as the fact that the Xcat-2 LE recruits Kinesin II (Betley et al. 2004), show quite convincingly that the Xcat-2 LE interacts extremely well with the Vg1 mRNA localization machinery and can utilize the late pathway even though endogenous Xcat-2 localizes much earlier than Vg1. Confusion arises from the discovery of a short motif, UGCAC, that is repeated six times in the approximately 230-nt Xcat-2 LE (Betley et al. 2002) and is absolutely required

for localization of the Xcat-2 LE at any stage. In addition, the UGCAC deletion mutant fails to compete for Vg1 localization machinery during the late pathway (Choo et al. 2005). Thus, from a functional sense, the UGCAC motif in the Xcat-2 LE is analogous to the UUCAC motif in the Vg1 LE. Furthermore, UGCAC and UUCAC motifs are at least partially interchangeable between the Xcat-2 and Vg1 LEs with regard to their ability to specify localization (Chang et al. 2004). A dilemma then arises when trying to explain why the UUCAC motif is a specific binding site for the Vg1 LE binding protein (Vera/Vg1RBP), whereas UGCAC is not (Deshler et al. 1998; Choo et al. 2005). This leads us to question the original interpretation of the correlation between UUCAC binding Vera/Vg1RBP and localization of vegetal RNAs, which inferred that UUCAC motifs promote localization by serving as binding sites for Vera/Vg1RBP (Deshler et al. 1998; Bubunenko et al. 2002; Kwon et al. 2002). In fact, recent work has shown that a dominant-negative RNA binding-deficient form of Vera/Vg1RBP fails to inhibit the localization of the Vg1 LE, suggesting that direct binding of Vera/Vg1RBP to the Vg1 LE RNA is not required for RNA localization of Vg1 (Rand and Yisraeli 2007). Together, these investigations raise the possibility that UUCAC motifs are more similar to UGCAC motifs and promote localization through some other mechanism yet to be identified. In fact, a situation such as this exists in the *Drosophila* field where 13 IMP binding motifs exist in the oskar 3′-UTR; IMP is a *Drosophila* homolog of Vera/Vg1RBP. In this system, the IMP binding sites are required for proper localization of *oskar* mRNA and for localization of the IMP protein with oskar at the posterior pole. Thus, IMP binding to the IMP binding motifs is required for its own localization to the posterior pole. However, IMP is not required for the localization of oskar to the posterior pole of *Drosophila* oocytes (Munro et al. 2006). Thus, IMP binding motifs must promote the localization of the *oskar* mRNA to the posterior pole through means other than serving as binding sites for the *Drosophila* Vera/Vg1RBP homolog, IMP.

Based on the findings just described, it is important to consider that UUCAC and IMP binding sites are similar to UGCAC motifs and promote localization through a mechanism that does not require binding to Vera/Vg1RBP or IMP, respectively. What might such a mechanism be? Two scenarios seem most likely and are not mutually exclusive. In the first scenario, these motifs could simply be binding other RNA binding proteins that promote localization, but have not yet been detected using biochemical methods available in the *Xenopus* system. Along these lines, members of my laboratory spent a lot of time trying to identify RNA binding proteins that specifically recognize UGCAC motifs using a variety of biochemical methods in the hope of finding a new key factor required for the process. No such protein was ever identified. This is a negative result and consequently was never published.

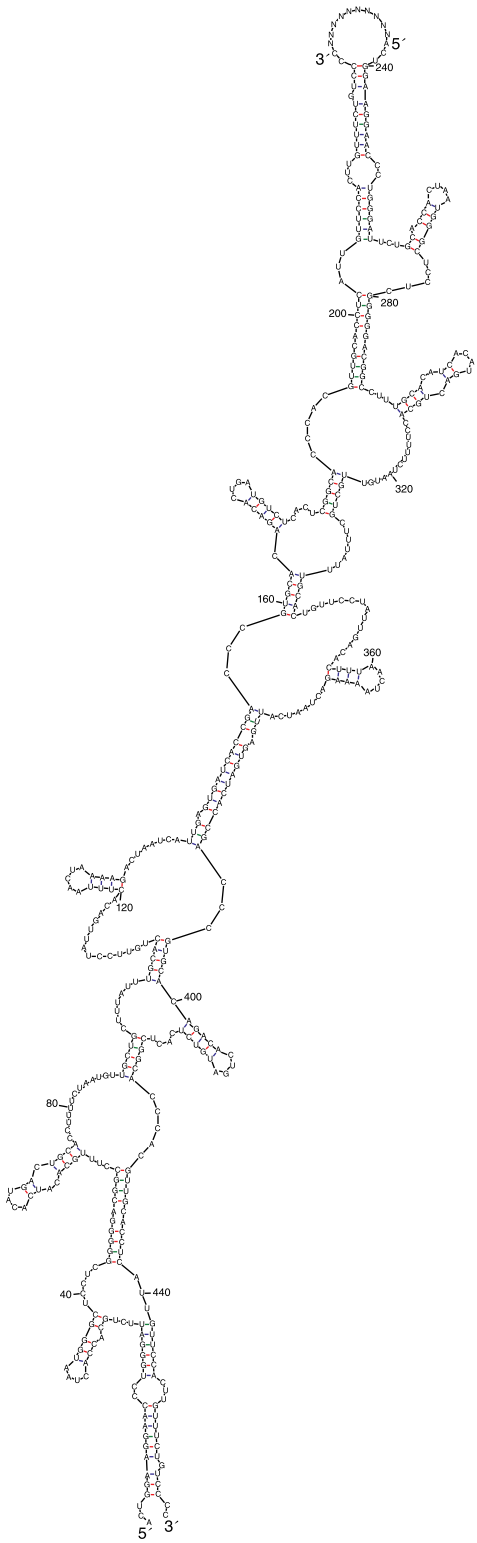
Another possibility is that UUCAC, IMP binding sites, and/or UGCAC motifs are simply the building blocks or evolutionary signatures of higher-order RNA structures that promote localization through the interactions with localization machinery that recognizes secondary and/or tertiary RNA structures. Conceptually, this is an attractive scenario because it is known that the highly conserved double-stranded RNA binding protein Staufen is involved in the localization of RNAs to the vegetal cortex of *Xenopus* oocytes (Yoon and Mowry 2004). Staufen is required for the localization of mRNAs to both the anterior and posterior pole of *Drosophila* oocytes (St Johnston et al. 1991) and recognizes complex structures in the bicoid LE that consist of stem–loop structures and intermolecular base-pairing interactions that support the formation of dimers and/or multimers (Ferrandon et al. 1997). Thus, Staufen is a general RNA localization factor required for localization of multiple mRNAs to different locations in a cell, and it is known to recognize high-order RNA structures *in vivo* in a selective fashion in *Drosophila*. Even so, biochemical assays have failed to detect specific binding of Staufen protein to RNA localization sequences in any model system. Furthermore, specific Staufen binding sites and higher-order RNA structures required for localization have not been identified in vertebrates. Therefore, the identification of double-stranded segments of

RNA LEs in the *Xenopus* system would represent an important step toward understanding how Staufen mediates RNA localization in these vertebrate cells.

The most effective *a priori* method for identifying RNA secondary structures that exist *in vivo* is to use phylogenetic comparisons of optimal (energetically most favorable) and suboptimal RNA secondary structures predicted for orthologous RNA sequences by RNA folding programs such as MFOLD. There is an approximately 90% chance that the biologically relevant RNA secondary structure for a single RNA sequence that forms *in vivo* will exist within the set of suboptimal structures that are within 10% of the free energy of the optimal structure predicted by MFOLD (Pace et al. 1989). Irrelevant secondary structures can generally be eliminated by comparing suboptimal structures predicted for two different, but orthologous, RNA sequences. Such an approach was used to identify the relatively complicated RNA secondary structure in the approximately 645-nt RNA LE of the *Drosophila* bicoid 3′-UTR (MacDonald 1990; Seeger and Kaufman 1990) which shows approximately 65% nt identity in alignments of *D. melanogaster* and *Drosophila pseudoobscura* bicoid sequences.

In an attempt to perform a similar analysis in the *Xenopus* system, I focused on the CAC-rich Xcat-2 LE (Betley et al. 2002) because it has extremely robust localization activity when compared to the Vg1 LE side by side (Choo et al. 2005), and it is much shorter, which limits the complexity of possible double-stranded RNA structures predicted through computational analysis. In addition, the ΔUGCAC localization-defective mutant is well characterized and could serve as a control for any structures emerging from this analysis. The sets of suboptimal structures predicted for the approximately 227 nt Xcat-2 RNA localization sequence of *X. laevis* or *Xenopus borealis* which show 89% nt identity when aligned to each other share no common secondary structural elements. One potential concern about this comparison was that too many common suboptimal structures would be identified given the high degree of sequence identity, making identification of the correct structure unlikely. Surprisingly, however, no common optimal or suboptimal RNA secondary structures were found even though both share the six UGCAC motifs required for localization (Betley et al. 2002; Chang et al. 2004) in addition to their high overall nt identity (*X. borealis* actually has one additional UGCAC motif). Both MFOLD and PFOLD, designed to find secondary structures common to more than one sequence, were used in this analysis (data not shown).

As mentioned earlier, previous work has shown that intermolecular RNA base pairing that supports the dimerization and/or multimer formation of the bicoid RNA LE (Ferrandon et al. 1997) mediates its specific binding to Staufen protein *in vivo* (Wagner et al. 2001) which is required for localization of bicoid RNA (St Johnston et al. 1991). Even though no common secondary or "hairpin" structures in orthologous Xcat-2 sequences could be identified, it was still possible that dimerization domains could exist in the Xcat-2 localization sequence. To identify regions of the Xcat-2 LE that have the potential to form intermolecular RNA base pairs, two copies of the sequence were linked together in tandem and analyzed with MFOLD. This was done for the *X. laevis* and *X. borealis* sequences, and both showed the same basic result: their LEs are predicted to form extensive intermolecular base-pairing interactions (Figure 2.2). When one copy of the *X. laevis* and one copy of the *X. borealis* MCLE sequences are fused in tandem, the ability to form this intermolecular structure is lost, no matter which sequence is entered first into the folding program (data not shown). This finding suggests that as the *X. laevis* and *X. borealis* Xcat-2 genes evolved, their LEs maintained an ability to form dimers. The *X. laevis* region of intermolecular RNA base-pairing potential consists of 80 intermolecular base pairs and only 20 intramolecular base pairs. Importantly, the UGCAC localization-defective deletion mutant (Betley et al. 2002) is not predicted to form such extensive intermolecular base pairing (data not shown). This ability to form intermolecular stretches of double-stranded RNA was also observed when several ascidian CAC-rich RNA LEs were analyzed (data not shown). Strikingly, when the fastest or most efficient human CAC-rich RNA LE we have identified in humans (Syntaxin1B2) (Andken



et al. 2007) is analyzed in a similar fashion by MFOLD, it is also predicted to contain a significant intermolecular base-pairing region (data not shown). Together, these findings provide strong phylogenetic evidence for dimerization and/or multimerization domains within functional CAC-rich RNA LEs. A somewhat related analysis showed that the ability to form extended double-stranded stretches of RNA correlated with localization activity for the noncoding Xlsirt RNA (Allen et al. 2003).

While the ability to form intermolecular dimers may be shared between the CAC-rich RNA LEs of Syntaxin1B2 and Xcat-2 and the bicoid RNA LE in *Drosophila*, there are major differences between the intermolecular interactions formed by bicoid and the vertebrate RNAs. Dimerization of bicoid is mediated by two discontinuous segments of only four or five base pairs via extremely dynamic RNA–RNA interactions (Ferrandon et al. 1997) referred to as RNA kissing reactions. RNA kissing is widely known to regulate a number of genetic processes (Eguchi et al. 1991; Gerhart et al. 1998) and can involve as few as two base pairs (Eguchi and Tomizawa 1991; Kim and Tinoco 2000). The stem-loop structures that have been proven through their evolutionary conservation position these kissing nt in the loop conformation that promotes the intermolecular pairing. The putative dimerization domains we have identified in vertebrate genes do not have conserved stem-loop structures and contain much more extensive intermolecular

Figure 2.2 Intermolecular base pairing potential of the Xcat-2 RNA LE. Two tandem copies of the Xcat-2 MCLE connected with 10 **N** were analyzed with MFOLD. The resulting structure is shown that has extensive intermolecular base-pairing potential that would support the formation of dimers or multimers *in vivo*. Evidence for this structure comes from the fact that mutations that reduce intermolecular base pairing impair localization but are rescued by compensatory mutations in *trans* (data not shown). For Xcat-2, nucleotides 403–610 were used, but we resequenced the DNA since a predicted restriction enzyme site from the NCBI sequence (Acc#X72340) was absent, and we identified a sequencing error that significantly affected the predicted extent of intermolecular base pairs. The sequence in this figure is the corrected sequence.