A Conversation with Marisa Bartolomei

Interviewer: Beth Moorefield

Senior Editor, Nature Structural and Molecular Biology

Marisa Bartolomei is a Professor in the Department of Cell and Developmental Biology and Co-Director of the Epigenetics Institute, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania.

Beth Moorefield: You study a unique process of gene expression in mammalian cells known as genomic imprinting, which directs expression specifically from either maternal or paternal alleles. I thought we could speak about the function of the imprinted genes and the mechanisms that govern their regulation.

Dr. Bartolomei: Imprinting is a mammalian phenomenon, and it affects ~100–200 genes. It's nicely conserved in mammals, which gives us the opportunity to use mouse as a good model to study imprinting in humans. These genes have very important processes in growth, but they also have functions in postnatal energy homeostasis, in behavior, and in other processes. So, when these genes are missing or defective, you end up with very broad changes—broad sorts of imprinting disorders—if there are defects in humans. That would include Beckwith–Wiedemann or Silver–Russell syndromes—those are growth imprinting disorders—or Angelman and Prader–Willi syndromes, and those are neurobehavioral disorders. So they have a broad range of functions, and absence of these genes causes these disorders, which is why we want to really understand their regulation.

Beth Moorefield: Are the maternal or paternal alleles imprinted with equal frequency, or is there a bias toward one or the other?

Dr. Bartolomei: These imprinted genes are regulated by regions that we call "imprinting control regions." These imprinting control regions are discrete elements in the genome that experience epigenetic modifications, so they have DNA methylation that's put on either during male or female gametogenesis. If you look at these regions, there are many more of these methylated imprinting-control regions that come from the female germline, and just a few in the male germline, so a lot of the action seems to be happening in the female germline. That said, these imprinted genes are found in large clusters through the genome and they seem to be pretty well equally represented as maternally expressed imprinted genes and paternally expressed imprinted genes. There doesn't seem to be a bias in that sense.

Beth Moorefield: How would they distinguish either of these alleles? How are they identified as maternal or paternal so that they're differentially recognized by the transcriptional machinery?

Dr. Bartolomei: That's the question. We know that these differential epigenetic modifications that are put on in the germline are what helps us to say, "If this comes from the maternal allele, either express or repress off the maternal allele," or modifications that are put on in the paternal germline help us to recognize something as being paternal and either expressed or repressed. These germline modifications are actually key to allowing the somatic cells to say, "Maternal, express; paternal, repress," or vice versa. These modifications that are put on in the germline are really the key to the imprinting. When something is perturbed, that's when you see dysregulation of imprinted genes, and that can occur in the germline, or it also can occur postfertilization.

Beth Moorefield: Do each of the alleles share common regulatory elements themselves?

Dr. Bartolomei: When imprinted genes were first identified ~25 years ago, there actually were three imprinted genes that were published in the same year: IGF-2 (Insulin-like Growth Factor 2), Insulin-like Growth Factor 2 Receptor, and H19, which encodes a noncoding RNA. It was thought that these imprinted genes were going to contain this primary sequence that was identified either in the male or the female germline and that was what was going to be the key to imprinting. In fact, we thought there was going to be a "silver bullet": "We're going to identify the imprinting box!"… like the TATA box, there was an "imprinting box." It was going to be a simple sequence. But there is no such simple sequence, so maybe there's something more complex that we just haven't figured out with secondary structure or some other kind of sequence. Each imprinted region, each imprinted gene, seems to have a different type of control sequence that does different things and we really don't know what that specific signal is that says, "This is going to be an imprinted gene and this is not going to be an imprinted gene." We

Published by Cold Spring Harbor Laboratory Press; doi: 10.1101/sqb.2017.82.034447

have little bits and pieces, but the absolute mechanism hasn't been worked out completely.

Beth Moorefield: Are there actually affected areas? The ones that are differentially modified, are they always in the same position within these alleles or within the different genes themselves?

Dr. Bartolomei: That gets a little bit to the mechanism. The real central piece of imprinting, these imprinting control regions or ICRs, those are the ones that have the methylation put on either in the male or female germline and it's maintained as long as imprinting is maintained. There're two regions that people where these ICRs can be found. One is intergenic—so, in between genes—and the other is actually in promoters. Often, it's in the promoters of long noncoding RNAs. One of the first uses that was found for long noncoding RNAs was that their transcription drove imprinting genes in *cis*. The methylation is in the promoter and if it's highly methylated, the long noncoding RNA is not made. If it's unmethylated, it is made. That's an example of how these imprinting control regions are not only giving you maternal- or paternal-like differentiation, but also functioning to do something.

The intergenic imprinting control regions are a little different. For the ones that I study, H19 and IGF-2, the imprinting control region is actually an insulator region and it binds CTCF. When it's unmethylated, CTCF binds and it helps to confer the imprinting of the region. When it's methylated, CTCF can't bind, so there's no insulator formed. These ICRs have these two sorts of broad functions. There are other mechanisms involved, but these are the two main functions.

Beth Moorefield: If you have the differential pattern, how does it actually get propagated to the daughter cells upon division?

Dr. Bartolomei: When the methylation is put on in the germline, it's through the de novo DNA methyltransferases. It's likely that differential chromatin modifications are involved too, but the propagation is through the maintenance methyltransferase, DNMT1. It's actually one of the first experiments done by the Jaenisch laboratory showing that in the absence of DNMT1, there was failure to maintain imprinting. As the cells divide, DNMT1 migrates with the replication fork and puts the methylation on the newly replicated strands so that the imprints are maintained. The key for imprinting is that it's epigenetically maintained. It's maintained throughout the life of the organism, in most cases.

Beth Moorefield: Not only maintenance but also the modifications have to be reversed to permit germline development: How is that accomplished?

Dr. Bartolomei: The mammalian embryo is very interesting in that there are two major times in development when there is this very large-scale reprogramming that occurs. The first time is after fertilization. The gametes come in with their own methylation patterns, and there is reprogramming as these cells are going to become pluripotent. At that time, imprints are maintained. There is something very special about them that enables that differential methylation to be maintained.

With mammals, as the embryo develops, the germline is set aside from the somatic cells. At early postimplantation, some cells that are recruited from the somatic cells, based on their location, are going to go on to form the germline. What happens there is then there is a second time of reprogramming. In that case, everything that will be reprogrammed is reprogrammed, including imprinting control regions. That's because if you're taking the germline from somatic cells that have maternal imprints and paternal imprints and you're going to become a germ cell in a female, you want to erase the maternal and paternal imprints and put on the maternal imprints; if you're in a male, you erase those and put on paternal imprints. That reprogramming occurs as the primordial germ cells are being specified and are replicating and migrating to enter the genital ridge.

Because a lot of methylation is lost at that time—DNMT1 is down-regulated and moved outside the nucleus for the most part—it was originally thought that what would happen is that the cells would replicate and there wouldn't be maintenance methylation. You would have a passive loss of methylation: Things would be diluted out, they enter the genital ridge, and they're remethylated. We know now that sometimes the demethylation is faster. In the last 10 years, a whole new mechanism of demethylation has been described through the TET enzymes. It's a family of three enzymes that will oxidize methylcytosines—TET1, -2, and -3—TET1 and -2 are expressed at the time when the methylation imprints are being erased. What we, and others, have shown now is that in the absence of TET1, some germ cells fail to erase all of the DNA methylation imprints; others do it fine. It seems that active and passive methylation are working together to reprogram, to erase imprints. In some pools of germ cells, we see a lot of remaining methylation; some, we see hardly any. We don't really understand what's going on in that case. The more you dig, the more mysteries you find in this field.

It's really an incredibly exciting field, a lot to be learned. That's still the basis of on-going research. Why is it? Why is it that we need two mechanisms, an active and a passive? Are they redundant? Are they helping each other? Are there some regions of the genome that are just so densely methylated, have such high affinity for DNMTs, that you need this extra mechanism in place to demethylate? We don't know yet, and we'd like to really learn that.

Beth Moorefield: It's not clear yet whether or not the residual methylation is actually precluding the de novo methylation events or what the functional interactions might be there?

Dr. Bartolomei: We do know that if we look at TET mutants that come from female and we look at maternally methylated regions that we would expect to be methylated in oocytes, they're methylated fine, but if we look at paternally methylated regions that should be *un*methylated

in oocytes, they are not. Sometimes they're still methylated; sometimes they're not. It doesn't look like residual methylation—whatever might be happening—is preventing de novo methylation from occurring. That's only looking at early in development and later in development. We don't know what's happening in between. If we could live-image cells… There's always this "on-paper experiment" you can think about where you could follow a cell and see demethylation occurring at every different stage. We can't do anything like that yet.

Beth Moorefield: Is there any evidence that environmental factors can influence any of these activities and actually alter patterns?

Dr. Bartolomei: That's another area of active investigation in my lab and lots of other labs. There's a whole field called Developmental Origins of Health and Disease. The idea there is that insults in utero to an embryo can be maintained or remembered and expressed later on in development. We're very interested in this hypothesis. A lot of it was originally defined by David Barker, and it was shown epidemiologically during the Dutch Hunger Winter when there was minimal caloric availability to people at that time. Women who had first-trimester babies during those times of minimal calories, these children were born and 50 years later had very high levels of heart disease and metabolic diseases. A lot of that was epidemiologically shown, but it has now been moved to animal models.

We've been very interested in this idea that very early gestational exposures may be the most dramatic. That's because that's when there's a lot of reprogramming going on. We've looked at endocrine disrupters and some assisted reproductive technologies where there're exposures in the early embryos. We use imprinted genes because for us they're our canary in the coal mine. We've shown with some of these exposures there are defects in the ability to maintain the methylation imprints or maintain other kinds of methylation, and there're defects in other kinds of reprogramming. Environmental perturbations during that early time in development can have a lasting influence on offspring. We're trying to look at that in more detail. It's another area of very active investigation as people have the capacity now to profile what's happening in the early embryo. Before we couldn't do that but now we can take small numbers of cells, look at DNA methylation, look at chromatin modifications, compare normal to environmental exposure of some kind. Science is such that we can do a lot of things that we hoped we could do 20 years ago.

A Conversation with Steve Henikoff

Interviewer: Richard Sever

Assistant Director, Cold Spring Harbor Laboratory Press

Steve Henikoff is a Member, Basic Sciences Division of the Fred Hutchinson Cancer Center and an Investigator at the Howard Hughes Medical Institute.

Richard Sever: You work with a veritable alphabet soup of centromere-associated proteins. Could you tell us a bit about the history of centromeres?

Dr. Henikoff: Walter Fleming introduced the term "chromatin" in 1882, so chromatin actually precedes the rediscovery of Mendel's laws, precedes the beginning of genetics. What he called chromatin was in these drawings of chromosome segregation in, I think it was newt: beautiful drawings of chromosomes that were being pulled to the poles at mitosis. What was particularly striking was that you could see everything. You saw the centrosome; you saw the spindle fibers that were holding onto the chromosomes. He was very careful to show a little darkening for each of the connections between the spindle fiber and the chromosome. He actually saw centromeres, or heterochromatin as we call it now, pericentric regions, whatever. He saw them and he colored them in. So actually, centromeres were discovered before the chromosome theory of heredity. Centromeres were really the first genetic loci, after Mendel's work.

The beginning of the molecular study of centromeres was this paper from Louise Clark and John Carbon where they showed that if you take a little piece of DNA that comes from the genetically mapped centromere and put it on a plasmid, you make a chromosome that will segregate normally during both mitosis and meiosis. That proved that that piece of DNA was the centromere. I remember when that paper came out, it was quite revolutionary. I thought, "Oh! They're all gonna be that way." We were studying heterochromatin in *Drosophila* at the time, and I thought it would really be very simple because in there among all those repeated sequences that we can't make any sense out of, there's going to be some little magic sequence. But that was not to be.

A lot of progress was made studying the budding yeast centromere, and even fission yeast turned out to be different. It turned out that nearly all multicellular organisms have centromeres that are highly repetitive tandem sequences. There was a lot of skepticism at the time that centromeres were just going to be homogeneous alpha satellite repeats, for example—which turns out to be what our centromeres are composed of. Other people thought it might be other satellites, until there was the experiment by Hunt Willard and his colleagues showing that you can actually make artificial centromeres. It's the equivalent of the yeast experiment, but you can do it with human centromeres. But you needed long arrays of these alpha satellite sequences, so that was a problem.

Then there was the fact that in 2000 or 2001, the draft human genome was published. It's still a gap at all of our centromeres. We haven't assembled it yet, because it's so homogeneous you can't do much with it. It's been difficult to study. But a lot of progress has been made in understanding centromeres, particularly in yeast. That's our best example.

Richard Sever: How long is that dedicated sequence in budding yeast?

Dr. Henikoff: It's one hundred twenty base pairs. It varies a little bit between chromosomes; there are sixteen chromosomes. They have canonical sequences that comprise the first element, which is about eight base pairs long, and that's a binding site for a transcription factor. On the other end, there is the binding site for a kinetochore-specific complex that recruits the centromeric nucleosome that sits in the middle, but there are only about eighty base pairs for that. People have been wondering, how do you get a nucleosome at eighty base pairs? This has been a bit of a controversy. It's pretty well established that it's the smaller particles. They're only four histones in there, probably, because that's all that you can really wrap eighty base pairs around.

I don't think that's been so much of a problem. Where it's really turned out to be a problem are the satellite centromeres: the ones in humans, plants, and most higher eukaryotes.

Richard Sever: These "satellite centromeres" are the alpha satellite sequences? What exactly are those?

Dr. Henikoff: The alpha satellite is a one hundred seventy or so base pair unit that gets repeated multiple times. It's almost homogeneous, but importantly, there are these higher order repeats. In our genome, for example, the majority of the alpha satellites for most of our chromosomes is a fundamental dimeric unit. These units might only be about 60% or so identical, one to the next, so

Published by Cold Spring Harbor Laboratory Press; doi: 10.1101/sqb.2017.82.034785

they're really very different from each other. But then you go to the next one, where that dimer gets repeated, et cetera, in the same orientation. It's very homogeneous, but then sometimes there's enough divergence that you'll see that there are higher orders above that: there's dimers, there's tetramers, et cetera. Usually, it's even-numbered ones for most of the chromosomes, but three of our chromosomes have an odd number. Five is the basic number. So it gets very complicated from chromosome to chromosome. That's been one of the problems.

Mouse centromeres are much simpler. They only have two different satellite repeats. One of them, the minor, has a one hundred twenty base pair repeat unit, and the major has a two hundred thirty-four base repeat unit, and they don't see higher order structures. The higher order structures are something that you find in primates but you don't find it in some others. We've been looking a lot into *Arabidopsis*; plants have the same kind of structure. They have a 178 base pair repeat unit, but they don't have higher order structures.

I can go through all these examples, but the one thing that's in common is that there are these highly repetitive sequences that are there, and that makes for an interesting challenge to study them. But it also asks the question: Why is it like that?

Richard Sever: Given this is such a fundamental thing that's shared by anything that needs to segregate its sister chromatids—which is everything—why is it all so different?

Dr. Henikoff: We have a hypothesis that we refer to as centromere meiotic drive. It only occurs in female meiosis. Female meiosis is asymmetric. Of the four products of meiosis, only one of them will get chosen to go to the next generation. There's the egg pole, and then there's the pole that will give you the polar bodies. In other words, only one of the four products will move to the egg pole. It turns out that there is a reorientation process, because there's a competition between centromeres from the maternal side, and centromeres from the paternal side.

Richard Sever: So we're back to selfish genes.

Dr. Henikoff: It's absolutely selfish. We'd actually proposed that centromere drive is a process in which there's a reorientation. There's a competition between the maternal and the paternal side and because of that competition centromeres are competing to make it into the egg pole. They're the most selfish elements that can be. They're just repeat sequences, and so the bigger centromeres in some organisms make it to the egg pole. That was shown very nicely in work from Ben Black and from Mike Lampson, who showed that this actually occurs in mouse, and they worked out a lot of the details of the process. So, we think that the "why" is that centromeres are competing, and they're rapidly evolving because of it. It's a Darwinian process of cheating in female meiosis.

Richard Sever: The next level, which is all of this alphabet soup of proteins coming in, is probably more conserved.

Dr. Henikoff: No, actually. It turns out it's an arms race. Harmit Malik, when he was a postdoc in the lab, discovered that CENP-A—centromere protein A—which is the centromeric histone, which is absolutely essential to everything…

Richard Sever: This is the thing that replaces the regular histone H…

Dr. Henikoff: Right. The H3 variant—we were working on in *Drosophila*, but it turns out to be more general—itself was showing rapid evolution. What Malik discovered was that you see a large excess of replacement, over synonymous, changes. We know about arms races where this can occur between a pathogen and the host immune system. A virus' coat protein will mutate to evade immune surveillance, and then the host immune system will clobber it, so you get this arms race going, and it builds up changes on the surface of the viral coat part. We see the same kind of process must have been occurring for the centromeric histone. But how could that be? That's going to be deleterious, and you'll be dead, because every cell division's going to die.

We found that for CENP-A, for CENP-C, basic centromeric proteins are undergoing an arms race, and female meiosis is the only time when you can actually have an arms race. We call that the centromere paradox: Why is it that centromeric sequences are evolving so rapidly, both in the size of the centromeres—your Y chromosome and mine might differ by an order of magnitude in size; they differ tremendously—yet they're doing the most basic process we know of in genetics, which is chromosome segregation, which is very stable. How could that be? We think it's this arms race of female meiosis. That was fifteen or so years ago, and I think it's holding up pretty well.

Richard Sever: You mentioned CENP-A, and you've noted there are others: CENP-B, -C, -TWSX, -Z. This is a giant protein complex on top of the DNA that links to the kinetochore or attaches to microtubules, correct?

Dr. Henikoff: That's right. It's on the DNA. The DNA wraps around the nucleosome part; it also wraps around CENP-TWSX. There's a complex of histone fold proteins that the DNA wraps around. What's important about that is there're two connections to the outer kinetochore. There's been some question as to where those connections come from, at the chromatin level. One connection is CENP-T, and one is CENP-C. That's where we get into alphabet soup. CENP-T has partners: W, S, and X. CENP-A has partners: H4, H2A, and H2B. If you think about it like that, you've got two parts that connect. What we show is that they're all actually part of one coherent complex. We showed this in a couple different ways, including a new method that we call "CUT&RUN" [Cleavage Under Targets and Release Using Nuclease].

Richard Sever: Everybody was familiar with ChIP-seq, where you bind your antibody onto the protein and you sonicate away all DNA. This seems like a much cleaner way to get a better signal.

Dr. Henikoff: This is a method that was introduced in 2004 by Ulrich Laemmli, however it really hadn't been used since until we started working on this and improving the protocol. We're starting out with intact cells, as opposed to breaking everything up like you do for ChIP. Everything stays intact. We then permeabilize them in such a way that we can get antibodies to go in and find their targets in the chromatin, say, for a transcription factor or a centromere protein. Then we add a fusion between protein A and micrococcal nuclease. Protein A will bind to the immunoglobulin G of the antibody, so therefore you're tethering this micrococcal nuclease, which is our standard tool for looking at chromatin.

Once it's tethered, it has some very nice properties. We can put it at low temperature so it'll just stay in place during the reaction. Micrococcal nuclease requires calcium, so it's not going to do anything until you add the calcium. You add the calcium, let it go for a few seconds or an hour or something, it doesn't seem to matter. It will cut out on both sides the particle, and then it'll float out through the nuclear pores and we just take the supernatants and extract the DNA, and it's very simple. Our current protocol takes about half a day. It could even take just a couple hours.

Being clean is important when you're looking at something like a centromere where there's only one per chromosome, or transcription factors where you have one here that has a footprint of about twenty base pairs and another here of twenty base pairs, and you've got ten kilobases in between. Basically, when you chew up the whole DNA like you do with ChIP, you're going to have a certain probability of every little piece of DNA getting sequenced, and that's going to give you a lot of background. It might be a low background, but you'll have to do a lot of sequencing in order to see those peaks. With CUT&RUN, because we leave the vast majority of the DNA behind on the beads and only sequence what gets released and floats out, the background really goes way down, such that we only have to sequence about a tenth as deeply. And because the antibody is not looking at the whole cell contents that get ground up when you do ChIP but only seeing the intact cells and the nucleus, it's only seeing the surfaces. That means that the antibody binding's usually very efficient. For a histone modification, we can get down to a hundred cells, so it should be good for low cell numbers. It's all new, but a lot of people are trying it, and I've been getting good reports. I think it's going to catch on.

Also, the ChIP method is destructive because at some point, before you add the antibody, you solubilize everything and you can lose some parts. That's what we're finding: No matter how you solubilize it, whether you chew everything up with micrococcal nuclease or grind it up, it's going to cause some damage. With CUT&RUN, the antibody's added when everything is intact. In fact, it's added at the time that you permeabilize the cell. They were live, and we get them as quickly as possible, and everything is intact. Because of that, I think we're getting a more accurate picture. What we see is quite compatible with what we've seen with ChIP, but I think it shows us that what we'd been looking at were, sort of, eroded particles. Actually, the intact particles are much larger than what we thought typically: over one hundred eighty base pairs over this dimeric repeat unit of three hundred and forty. Because they're larger, they can accommodate a lot of the alphabet soup of proteins that we see there, including the ones I've just talked about. We can see that they're all there.

A Conversation with Tim Stearns

Interviewer: Jan Witkowski

Cold Spring Harbor Laboratory

Tim Stearns is the Frank Lee and Carol Hall Professor and a Professor of Genetics at Stanford University.

Jan Witkowski: What are some of the key points about centriole structure and regulation that are of research interest at the moment?

Dr. Stearns: The centrosomes are segregated in mitosis and that is one of the interesting things about them. If you look through history, the realization that they are important organelles in cells happened more than a hundred years ago, but a modern understanding of what they do and how you ensure that each cell gets one at the beginning of the cell cycle has only occurred more recently. They, like chromosomes, are segregated on the mitotic spindle. So, by analogy to chromosomes, that means they have to be duplicated exactly once per cell cycle. The duplication and the segregation of centrosomes and the centrioles within them is really something that you can think about in parallel with thinking about chromosomes.

Jan Witkowski: A centriole looks almost as though it buds off from… I was going to say its parent, but that's not right. The original?

Dr. Stearns: We do use the word "mother," that's true.

Jan Witkowski: How does that process happen? It comes off at right angles.

Dr. Stearns: It's really quite remarkable. For DNA, we've known for a long time—and the original notion of this came from the Watson and Crick double-strand helix—that you replicate DNA by pulling it apart and then use those strands as templates to create new strands. For the centriole, it's been a mystery as to how it duplicates because you don't have that obvious mechanism that comes from the structure. What you do have is the centriole—which is a proteinaceous structure, not nucleic acid—and it has this very characteristic structure, the centriole, which has this beautiful ninefold symmetry and other proteins attach to that. That duplicates just like chromosomes, once every cell cycle.

As you say, it looks somewhat like a new centriole grows as a bud on the side of an existing one. What does happen is that when you initiate centriole formation in a cell in the G_1 phase of the cell cycle, each centriole grows exactly one from its side. The new one grows at a right angle from the side of the existing one. The mother centriole would be the existing one and a new daughter grows from the side of the mother. That's true in many organisms, so it's really a quite well conserved aspect of how duplication works. And yet, why it works that way, why you only form one new centriole on the side of a mother is still not clear. Even though we know a lot about the molecules involved in duplication, the reason it's constrained to occur exactly there and the way you constrain it to occurring exactly once, are still things that are active areas of work.

Jan Witkowski: What are the molecules that make up a centriole?

Dr. Stearns: The structure of the centriole itself, when you look at the classical image that you can find in textbooks or on Google Images, is this beautiful ninefold symmetric microtubule-based structure, where the microtubules are in triplets. This is unusual. You only find these triplets where there's an A tubule and then a B and a C tubule that grow on the side of the A tubule. The centriole is the only place you find that. Both the ninefold symmetry and having these triplets are things that really define the centriole.

The ninefold symmetry comes from a protein called SAS-6. There's been a lot of really quite beautiful work on this that shows that there's intrinsic behavior of the protein that results in that ninefold symmetry. The SAS-6 protein forms ninefold symmetric oligomers both in vitro and, presumably, in vivo and that's the basis for this ninefold symmetry of the centriole, which is absolutely conserved from all eukaryotes that have centrioles, and presumably in the last common ancestor of all eukaryotes. Those microtubules that this cartwheel that SAS-6 forms are attached to are formed from α- and β-tubulin just like all microtubules in the cell, except that, again, you never find triplet microtubules in any other setting.

For a long time it's been thought that there must be other proteins: δ-tubulin and ε-tubulin are members of the tubulin family but are not the canonical members of the ones that make microtubules, and the specialized tubulin, γ-tubulin, that nucleates microtubules, or caps them at least. These are different.

Published by Cold Spring Harbor Laboratory Press; doi: 10.1101/sqb.2017.82.035360

Jan Witkowski: The A, B, and C tubules use the same tubulin molecules?

Dr. Stearns: As regular microtubules, yes. The bulk of those A, B, and the C tubules are made up of α- and β-tubulin, which make this heterodimer that polymerizes to make microtubules. It seems that those centriolar microtubules in this triplet form are made largely of α- and β-tubulin. I say "largely" just to leave open the possibility that there's something else that might make the triplet microtubules special.

Jan Witkowski: Why don't they go on to make a quadruplet?

Dr. Stearns: That is a great question. Why is it limited to being a triplet? That is not clear. It's not clear why they don't ever go beyond being triplets.

Jan Witkowski: All the things that link the nine, the bundles to keep them in a cylindrical form: Are there things between those triplets?

Dr. Stearns: Yes. Part of what keeps the structure intact is this cartwheel that you see when you do electron microscopy; that's really the only way you can see it in its full glory in terms of its structure. That's when you see the ninefold symmetry of this cartwheel at the center of the centriole. Again, there are spokes that radiate out from that and they contact these microtubules that are arranged in ninefold symmetry. Other proteins that are part of that have been seen in the electron microscope but we don't know the identity of them. There's something called the A–C linker, which links adjoining triplet microtubules, but no component has ever been identified for those yet.

Jan Witkowski: You're using genetic tools to dissect structure.

Dr. Stearns: Right. It's one of the revolutions of cell biology: the ability, using CRISPR–Cas9, to manipulate the genomes of mammalian cells. This ability really started with zinc fingers, which were hard to use and then TALENs, which were easier. But now CRISPR–Cas9 makes it tremendously easy to manipulate the genome of mammalian cells. Not as easy as many people think. If you're interested in turning mammalian cells into the equivalent of a yeast cell, that's not really feasible, but this ability to make mutations and then assess the phenotype is key to modern cell biology.

Jan Witkowski: How have you been using CRISPR–Cas9 in your work?

Dr. Stearns: These proteins, δ-tubulin and ε-tubulin, which are variant members of the tubulin family are not in all eukaryotes, but they're in most that have centrioles and cilia. We have been working on these proteins for a long time. We identified them—and others did too—in the late '90s. Susan Dutcher first identified δ-tubulin and then we identified ε-tubulin and more recently we, and others, identified ζ-tubulin: the last member of the tubulin superfamily. I can say that conclusively because we have a lot of sequenced genomes, so it's clear that there are no other members of the tubulin superfamily. Because of the conservation of them, we know that the last common ancestor of eukaryotes had all six of these. Some have been lost in various branches of the eukaryotic tree, but if you look at human cells, we have α, β, γ—which all eukaryotes have—and then we have δ- and ε-tubulin.

We've known this for a long time. The limitation has been, how do you determine the function of proteins in mammalian cells without the ability to manipulate the genome? RNAi, of course, was one of the tools that came along and made some forms of manipulation of gene expression possible. We tried that with δ- and ε-tubulin but it never worked very well. Our thought about that is that there's really a very small number of molecules of these proteins, such that depleting the RNA, even to the levels you typically get with good RNAi, was never sufficient to yield a phenotype. Now, we can make real deletions using Cas9 that allow one to have a defined genotype with a predicted null mutation homozygous form and then assess the phenotype of the cells.

Jan Witkowski: When you knock out these, what happens to the cells? First, what happens to the centriole? Then what's the effect on the cell?

Dr. Stearns: Using Cas9, we tried hard to knock out both δ-tubulin and ε-tubulin and could never identify cell lines where they were clearly knocked out. It turns out that the reason for that is because cells that have a problem with their centrioles, if they are p53-plus, p53 senses a problem and the cells arrest in G_1 of the cell cycle. We didn't realize at the beginning of our experiment that that was true, but we did soon afterward from the work of others. Once we started doing the experiments in p53-minus cell lines, then we could easily get null mutations. Right away, that hinted that there was going to be something wrong with centrioles that was being sensed by p53. Sure enough, when we looked closely at the phenotype of these mutant cell lines—and both the δ-tubulin and ε-tubulin null mutants have very similar phenotypes—the phenotype at the level of the centrioles is that the structure of the centrioles that do exist is different. They only have singlet microtubules instead of triplet microtubules. Remember, those triplets are one of the key features of what makes a centriole special. These lack triplet microtubules and only have singlets.

There's another aspect to it that caught our eye immediately, which was that if you look at a population of cells that are null mutants for these genes, there's a very interesting distribution of centrioles in these cells. A normal cell has two centrioles in the G_1 phase of the cell cycle. They duplicate prior to mitosis; you have four. Then they segregate on the spindle and you get a daughter cell with two centrioles again. In this case, about half of the cells had zero centrioles: none detectable whatsoever. The other half of the cells had more than four centrioles: five, six, seven, eight. Right away again, that suggested that there was something very odd about centriole duplication in these cells.

We determined that what was going on, is that that half having centrioles and half having no centrioles is what you see when you look at an asynchronous population, with

cells at all stages of the cell cycle. If you look at specific stages of the cell cycle, you see that in G_1 there are zero centrioles. They form in S Phase. They persist through mitosis and then they disintegrate at some point between mitosis and the next interphase. Actually, probably the start of the next interphase.

It's one of the interesting things about centrioles. Just as we talked about the seeming budding that occurs when you grow a new centriole from the side of an existing one, because that's such a common feature of centriole duplication people have long assumed that you had to have an existing centriole to grow a new one. We know that that is not true because it's been shown many different ways that for many cells from animals and other organisms too, you can create new centrioles de novo with no preexisting centriole.

Jan Witkowski: So, you've got these two populations in asynchronous cultures, but if you followed a single cell, you'd find that it at one stage had none and at another stage had several?

Dr. Stearns: Exactly. If you look at a single cell as it progressed through the cell cycle, you start with zero, and you would create centrioles de novo with no existing centrioles. Those would persist until mitosis but then they would… I use the word "disintegrate," although actually seeing that happen is not something that we've been able to do, but that's our interpretation of the results.

Jan Witkowski: Are there cells that never have any centrioles? Can cells survive and grow and multiply without centrioles?

Dr. Stearns: It is true that centrioles duplicate exactly once per cell cycle and they participate in the mitotic spindle. Because of their close association with the spindle, people had always assumed that they were required to make the spindle. That is clearly not true, because you can force mammalian cells to lose their centrioles, either genetically or now with a small molecule inhibitor of a required kinase, and in both of those cases the cells will be able to grow without centrioles. So, no centrioles, no centrosome, but they can do mitosis. They have to be p53-minus though, because if you don't have centrosomes, there is a problem—probably with mitosis—that p53 senses. If you get rid of p53, the cells can divide just fine with no centrosomes.

Jan Witkowski: p53 having that sort of oversight absolutely makes a lot of sense in terms of trying to avoid chromosomal abnormalities and such.

Dr. Stearns: Right.

Jan Witkowski: You've shown by knocking out these tubulins that they are associated or involved in important aspects of centrioles. Do you know where they are in the centriole?

Dr. Stearns: We've been unable to really get convincing localization of these proteins. It's not for lack of trying. We, and others, have tried to see where in the cell these proteins are. For cell biologists, that's one of the key clues to what a protein does. It's been difficult either with antibodies or with tagged proteins of various sorts. I think it's because there're very few molecules of these proteins in cells. That's an assessment based not only on failure to localize the proteins but also on this small amount of RNA and the very low level of protein as assessed with antibodies on western blots. If you look at the centriole structure, where these proteins might fit into that where they could help to make or to stabilize the triplet microtubules, there's several ways you could imagine very low-abundance proteins playing critical roles in that structure.

Jan Witkowski: You mentioned that you've not really been able to follow the disintegration or destruction of these centrioles that end. What do you think is going on? Have you tried messing around with the ubiquitination system to see if you can force them through?

Dr. Stearns: We've gotten close to seeing the disintegration. We certainly can see it by light microscopy, but centrioles are very small structures and light microscopy doesn't really get you to the level you need to be to tell what's going on. By electron microscopy, which we'd done in collaboration with Jadranka Loncarek, we can see that as cells enter mitosis the centrioles elongate, even these aberrant centrioles. They don't have triplet microtubules; they only have singlet microtubules so they're narrower, but they do elongate to the full length. In the electron microscope, it's clear that there's something wrong with them. Having singlets only, it seems that they don't recruit some proteins that are parts of what a maturing centriole would have. They have gaps in the structure. We believe that what is resulting in disintegration is the depolymerization of these singlet microtubules. Thus, the triplet microtubules are really required to be stable. That's what provides the long-term stability of these triplet microtubules, which persist for a very long time. These microtubules don't turn over, unlike most microtubules in the cell.

Jan Witkowski: Centrioles are also a part of cilia?

Dr. Stearns: Yes.

Jan Witkowski: Cilia also get reformed after division?

Dr. Stearns: They do. Most cells in our body that are in G_0 phase—so they're doing their thing as a differentiated cell—will have a primary cilium projecting from the surface from the cell. The microtubule component of that cilium, which is the core structure, is contiguous with the centriole, so it grows from the end of the centriole. When one says that the centriole—or it's often called the basal body in this context—is part of the cilium, it literally is. It is part of the microtubule structure of the cilium. That grows every cell cycle after mitosis in a cycling cell and again, is present in most differentiated cells.

Jan Witkowski: Your "funny" ones don't get involved in the cilium?

Dr. Stearns: We've never seen one with a cilium. That, in part, is because we've also seen that they never have these appendages, which are structures that are added onto the

centriole to functionalize it. In the case of making a cilium, they allow it to interact with the plasma membrane. That interaction is critical to creating the ciliary compartment and you have to do that to be able to traffic molecules specifically into the ciliary compartment to make the cilium. These centrioles never have that ability to attach to the membrane and we never see them associated with a cilium, so that's another defect that these centrioles have.

Jan Witkowski: So those cells never develop cilia?

Dr. Stearns: Right. These null cells never develop cilia, which would clearly be a bad thing for the organism because the primary cilium, at least in vertebrates or mammals, is absolutely required for development.

Jan Witkowski: If I could grant you one wish, what would you really like as your next bit of knowledge?

Dr. Stearns: I think the most useful thing would be to know exactly where those proteins are because when you're working at the level of trying to relate protein function to structure as we are, you really need to be able to say where these proteins—which could be structural components of this complex thing, the centriole—are in the structure. That will inform models for how they actually work. Speculatively, without any evidence yet, this linker between the triplet microtubules—the "A–C linker" it's called, because it links the A tubule to the C tubule of the next triplet—could be a structure that these δ- and ε-tubulins are either a component of, or required to make. The absence of that structure might explain the instability of the centriole. That's just a hypothesis and we need to have data about that, but the localization of the proteins would be a great start for that.